M3-4-5 A16 Notes: Geometric Mechanics I

Professor Darryl D Holm   Fall Term 2013
Imperial College London   d.holm@ic.ac.uk
http://www.ma.ic.ac.uk/~dholm/
Course meets $\boxed{\text{T, W, Th 11am @ Hux 658}}$

Text for the course M3-4-5 A16:

Geometric Mechanics I: Symmetry and Dynamics (aka GM1)
by Darryl D Holm, World Scientific: Imperial College Press, Singapore, Second edition (2011).
ISBN 978-1-84816-195-5

Contents

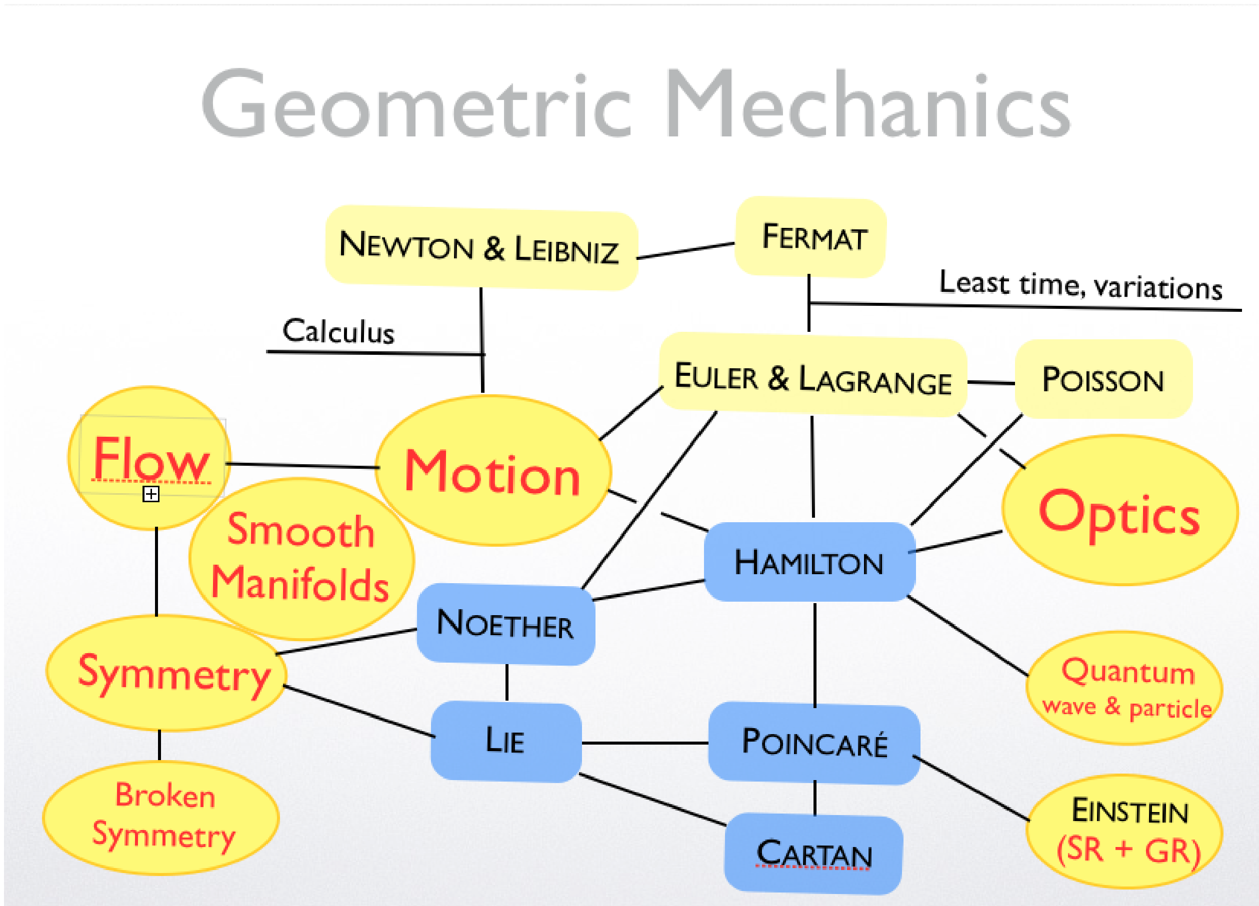
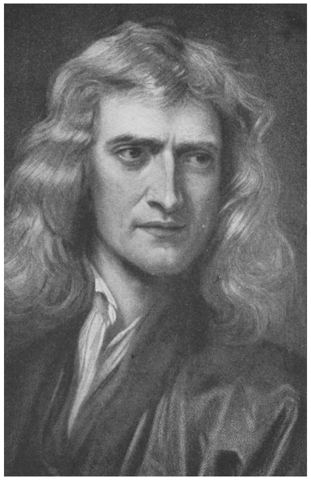


Figure 1: Geometric Mechanics has involved many great mathematicians!

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

Background reading: Chapter 2, [Ho2011GM1].

2 Newton



Isaac Newton

Briefly stated, Newton's three laws of motion in an inertial frame are:

1. **Law of Inertia** An object in uniform motion (constant velocity) will remain in uniform motion unless acted upon by a force.
2. **Law of Acceleration** Mass times acceleration equals force.
3. **Law of Reciprocal Action** To every action there is an equal and opposite reaction.

Newton's Law of Inertia may be regarded as the definition of an inertial frame. Newton also introduced the following definitions of space, time and motion. These definitions are needed to formulate and interpret Newton's three laws governing particle motion in an inertial frame.

Definition

2.1 (Space, time, motion)**.**

- ***Space*** *is three-dimensional.* ***Position*** *in space is located at vector coordinate* $\mathbf{r} \in \mathbb{R}^3$*, with* ***length*** $|\mathbf{r}| = \langle \mathbf{r}\,,\,\mathbf{r}\,\rangle^{1/2}$ *defined by the* ***metric pairing*** *denoted* $\langle\,\cdot\,,\,\cdot\,\rangle : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$.
- ***Time*** *is one-dimensional. A moment in time occurs at* $t \in \mathbb{R}$.
- ***Motion of a single particle*** *in space (*$\mathbb{R}^3$*, fixed orientation) is continuously parameterised by time* $t \in \mathbb{R}$ *as a* ***trajectory***

$$\mathcal{M} : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3\,, \quad (\mathbf{r}_0, t) \to \mathbf{r}(t)\,,$$

which maps initial points $\mathbf{r}_0 = \mathbf{r}(0)$ *in* $\mathbb{R}^3$ *into curves* $\mathbf{r}(t) \in \mathbb{R}^3$ *parameterised by time* $t \in \mathbb{R}$.

- ***Velocity*** *is the tangent vector at time t to the particle trajectory,* $\mathbf{r}(t)$, $d\mathbf{r}/dt := \dot{\mathbf{r}} \in T\mathbb{R}^3 \simeq \mathbb{R}^3 \times \mathbb{R}^3$ *with coordinates* $(\mathbf{r}, \dot{\mathbf{r}})$.
- ***Acceleration*** *measures how the velocity (tangent vector to trajectory) may also change with time,* $\mathbf{a} := \dot{\mathbf{v}} = \ddot{\mathbf{r}} \in TT\mathbb{R}^3 \simeq \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3$ *with coordinates* $(\mathbf{r}, \dot{\mathbf{r}}, \ddot{\mathbf{r}})$.

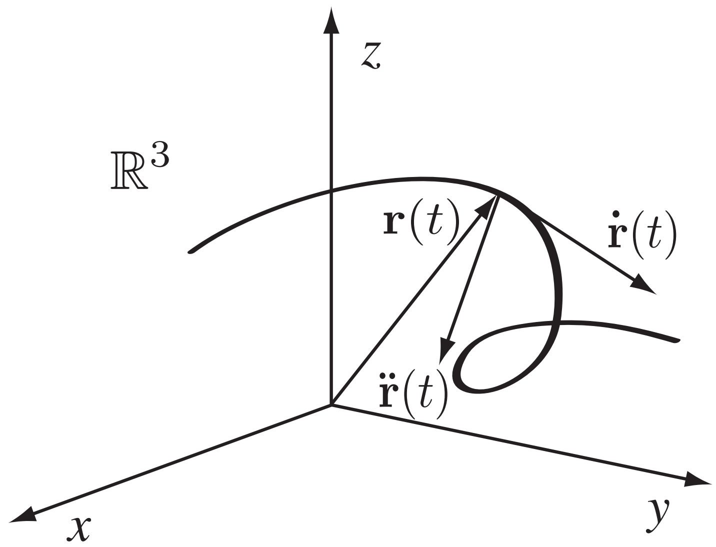


Figure 2: Position $\mathbf{r}(t) \in \mathbb{R}^3 \times \mathbb{R}$, velocity $\dot{\mathbf{r}}(t) \in T\mathbb{R}^3 \times \mathbb{R}$ and acceleration $\ddot{\mathbf{r}}(t) \in TT\mathbb{R}^3 \times \mathbb{R}$ along a trajectory of motion governed by Newton's Second Law, $m\ddot{\mathbf{r}} = \mathbf{F}(\mathbf{r}, \dot{\mathbf{r}})$.

- ***Motion of*** N ***particles*** *is defined by the one-parameter map,*

$$\mathcal{M}_N : \mathbb{R}^{3N} \times \mathbb{R} \to \mathbb{R}^{3N} .$$

2.1 Inertial frames and Galilean transformations

Definition

2.2 (Galilean transformations).

Linear transformations of reference location $(\mathbf{r}_0)$, *origin of time* t_0, *orientation* O *or state of uniform translation at constant velocity* $(\mathbf{v}_0 t)$ *are called* ***Galilean group transformations***. *They have a matrix representation given by*

$$\begin{bmatrix} O & \mathbf{v}_0 & \mathbf{r}_0 \\ \mathbf{0} & 1 & t_0 \\ \mathbf{0} & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{r} \\ t \\ 1 \end{bmatrix} = \begin{bmatrix} O\mathbf{r} + \mathbf{v}_0 t + \mathbf{r}_0 \\ t + t_0 \\ 1 \end{bmatrix} \tag{2.1}$$

Definition

2.3 (Group). *A* ***group*** G *is a set of elements that possesses a binary product (multiplication),* $G \times G \to G$, *such that the following properties hold:*

- *The product* gh *of* g *and* h *is associative, that is,* $(gh)k = g(hk)$.
- *A unique identity element exists,* $e:\ eg = g$ *and* $ge = g$, *for all* $g \in G$.
- *The inverse operation exists,* $G \to G$, *so that* $gg^{-1} = g^{-1}g = e$.

Definition

2.4 (Lie group). *A* ***Lie group*** *is a group that depends smoothly on a set of* n *parameters. That is, a Lie group is both a group and a smooth manifold (a smooth space that is everywhere locally isomorphic to* $\mathbb{R}^n$*), for which the group operation is given by composition of smooth invertible functions.*

Proposition

2.5 (Lie group property). *Galilean transformations form a Lie group, modulo reflections (which are discrete operations).*

Remark

2.6 (Parameters of Galilean transformations). *The Galilean group in three dimensions* $G(3)$ *has ten parameters*

$$(O \in SO(3)\,,\ \mathbf{r}_0 \in \mathbb{R}^3\,,\ \mathbf{v}_0 \in \mathbb{R}^3\,,\ t_0 \in \mathbb{R}\,)\,.$$

Remark

2.7 (Matrix representation of $G(3)$). *The formula for group composition $G(3) \times G(3) \to G(3)$ may be represented by matrix multiplication from the left as*

$$\begin{pmatrix} \tilde{O} & \tilde{\mathbf{v}}_0 & \tilde{\mathbf{r}}_0 \\ \mathbf{0} & 1 & \tilde{t}_0 \\ \mathbf{0} & 0 & 1 \end{pmatrix} \begin{pmatrix} O & \mathbf{v}_0 & \mathbf{r}_0 \\ \mathbf{0} & 1 & t_0 \\ \mathbf{0} & 0 & 1 \end{pmatrix} = \begin{pmatrix} \tilde{O}O & \tilde{O}\mathbf{v}_0 + \tilde{\mathbf{v}}_0 & \tilde{O}\mathbf{r}_0 + \tilde{\mathbf{v}}_0 t_0 + \tilde{\mathbf{r}}_0 \\ \mathbf{0} & 1 & \tilde{t}_0 + t_0 \\ \mathbf{0} & 0 & 1 \end{pmatrix}.$$

Definition

2.8 (Subgroup). *A* ***subgroup*** *is a subset of a group whose elements also satisfy the defining properties of a group.*

Exercise. List the subgroups of the Galilean group that do not involve time. ★

Answer. The subgroups of the Galilean group that are independent of time consist of

- Spatial translations $g_1(\mathbf{r}_0)$ acting on $\mathbf{r}$ as $g_1(\mathbf{r}_0)\mathbf{r} = \mathbf{r} + \mathbf{r}_0$.
- Proper rotations $g_2(O)$ with $g_2(O)\mathbf{r} = O\mathbf{r}$ where $O^T = O^{-1}$ and $\det O = +1$. This subgroup is called $SO(3)$, the **special orthogonal group** in three dimensions.
- Rotations and reflections $g_2(O)$ with $O^T = O^{-1}$ and $\det O = \pm 1$. This subgroup is called $O(3)$, the **orthogonal group** in three dimensions.
- Spatial translations $g_1(\mathbf{r}_0)$ with $\mathbf{r}_0 \in \mathbb{R}^3$ compose with proper rotations $g_2(O) \in SO(3)$ acting on a vector $\mathbf{r} \in \mathbb{R}^3$ as

  $$E(O, \mathbf{r}_0)\mathbf{r} = g_1(\mathbf{r}_0)g_2(O)\mathbf{r} = O\mathbf{r} + \mathbf{r}_0\,,$$

  where $O^T = O^{-1}$ and $\det O = +1$. This subgroup is called $SE(3)$, the **special Euclidean group** in three dimensions. Its **action** on $\mathbb{R}^3$ is written abstractly as $SE(3) \times \mathbb{R}^3 \to \mathbb{R}^3$.
- Spatial translations $g_1(\mathbf{r}_0)$ compose with proper rotations and reflections $g_2(O)$, as $g_1(\mathbf{r}_0)g_2(O)$ acting on $\mathbf{r}$. This subgroup is called $E(3)$, the **Euclidean group** in three dimensions. ▲

Remark

2.9. *Spatial translations and rotations do not commute in general. That is, $g_1 g_2 \neq g_2 g_1$, unless the direction of translation and axis of rotation are collinear.*

Remark

2.10 (Group structure of Galilean transformations)**.** *The Galilean group is a semidirect-product Lie group, which may be written as*

$$G(3) = SE(3) \circledS \mathbb{R}^4 = \Big(SO(3) \circledS \mathbb{R}^3 \Big) \circledS \mathbb{R}^4 \,.$$

That is, the subgroup of Euclidean motions consisting of rotations and Galilean velocity boosts $(O, \mathbf{v}_0) \in SE(3)$ acts homogeneously on the subgroups of space and time translations $(\mathbf{r}_0, t_0) \in \mathbb{R}^4$ which commute with each other.

2.2 Matrix representation of $SE(3)$

The special Euclidean group in three dimensions $SE(3)$ acts on a position vector $\mathbf{r} \in \mathbb{R}^3$ by

$$E(O, \mathbf{r}_0)\mathbf{r} = O\mathbf{r} + \mathbf{r}_0 \,.$$

A 4×4 **matrix representation** of this action may be found by noticing that its right-hand side arises in multiplying the matrix times the *extended* vector $(\mathbf{r}, 1)^T$ as

$$\begin{pmatrix} O & \mathbf{r}_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{r} \\ 1 \end{pmatrix} = \begin{pmatrix} O\mathbf{r} + \mathbf{r}_0 \\ 1 \end{pmatrix} .$$

Therefore we may identify a group element of $SE(3)$ with a 4×4 matrix,

$$E(O, \mathbf{r}_0) = \begin{pmatrix} O & \mathbf{r}_0 \\ 0 & 1 \end{pmatrix} .$$

The group $SE(3)$ has six parameters. These are the angles of rotation about each of the three spatial axes by the orthogonal matrix $O \in SO(3)$ with $O^T = O^{-1}$ and the three components of the vector of translations $\mathbf{r}_0 \in \mathbb{R}^3$.

The **group composition law** for $SE(3)$ is expressed as

$$\begin{aligned} E(\tilde{O}, \tilde{\mathbf{r}}_0) E(O, \mathbf{r}_0)\mathbf{r} &= E(\tilde{O}, \tilde{\mathbf{r}}_0)(O\mathbf{r} + \mathbf{r}_0) \\ &= \tilde{O}(O\mathbf{r} + \mathbf{r}_0) + \tilde{\mathbf{r}}_0 \,, \end{aligned}$$

with $(O\,,\,\tilde{O}) \in SO(3)$ and $(\mathbf{r}\,,\,\tilde{\mathbf{r}}_0) \in \mathbb{R}^3$. This formula for group composition may be represented by matrix multiplication from the left as

$$E(\tilde{O},\tilde{\mathbf{r}}_0)E(O,\mathbf{r}_0) = \begin{pmatrix} \tilde{O} & \tilde{\mathbf{r}}_0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} O & \mathbf{r}_0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \tilde{O}O & \tilde{O}\mathbf{r}_0 + \tilde{\mathbf{r}}_0 \\ 0 & 1 \end{pmatrix},$$

which may also be expressed by simply writing the top row,

$$(\tilde{O}\,,\,\tilde{\mathbf{r}}_0)(O\,,\,\mathbf{r}_0) = (\tilde{O}O\,,\,\tilde{O}\mathbf{r}_0 + \tilde{\mathbf{r}}_0)\,.$$

The identity element (e) of $SE(3)$ is represented by

$$e = E(I,\mathbf{0}) = \begin{pmatrix} I & \mathbf{0} \\ 0 & 1 \end{pmatrix},$$

or simply $e = (I,\mathbf{0})$. The inverse element is represented by the matrix inverse

$$E(O,\mathbf{r}_0)^{-1} = \begin{pmatrix} O^{-1} & -O^{-1}\mathbf{r}_0 \\ 0 & 1 \end{pmatrix}.$$

In this matrix representation of $SE(3)$, one checks directly that

$$\begin{aligned} E(O,\mathbf{r}_0)^{-1}E(O,\mathbf{r}_0) &= \begin{pmatrix} O^{-1} & -O^{-1}\mathbf{r}_0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} O & \mathbf{r}_0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} I & \mathbf{0} \\ 0 & 1 \end{pmatrix} = (I,\mathbf{0}) = e\,. \end{aligned}$$

In the shorter notation, the inverse may be written as

$$(O\,,\,\mathbf{r}_0)^{-1} = (O^{-1}\,,\,-O^{-1}\mathbf{r}_0)$$

and $O^{-1} = O^T$ since the 3×3 matrix $O \in SO(3)$ is orthogonal.

Exercise. Compute the matrix representation of the inverse Galilean group transformations. ★

Definition

2.11 (Uniform rectilinear motion).
*Coordinate systems related by Galilean transformations are said to be in **uniform rectilinear motion** relative to each other.*

Proposition

2.12 (Existence of inertial frames).
*Following Newton, we assume the existence of a preferred reference frame, which Newton called Absolute Space and with respect to which he formulated his laws. Coordinate systems in uniform rectilinear motion relative to Absolute Space are called **inertial frames**.*

Proposition

2.13 (Principle of Galilean relativity).
*The laws of motion are independent of reference location, time, orientation, or state of uniform translation at constant velocity. Hence, these laws are **invariant** under Galilean transformations. That is, the laws of motion must have the same form in any inertial frame.*

2.3 Newton's Laws

The definitions of space, time, motion, uniform rectilinear motion and inertial frames provide the terms in which Newton wrote his three laws of motion. The first two of these may now be written more precisely as [KnHj2001]:

(#1) **Law of Inertia** An object in uniform rectilinear motion relative to a given inertial frame remains so, unless acted upon by an external force.

(#2) **Law of Acceleration** When acted upon by a prescribed external force, $\mathbf{F}$, an object of mass m accelerates according to $m\ddot{\mathbf{r}} = \mathbf{F}(\mathbf{r}, \dot{\mathbf{r}})$ relative to a given inertial frame.

Remark

2.14. *For several particles, Newton's Law #2 determines the motion resulting from the prescribed forces $\mathbf{F}_j$ as*

$$m_j\ddot{\mathbf{r}}_j = \mathbf{F}_j(\mathbf{r}_k - \mathbf{r}_l, \dot{\mathbf{r}}_k - \dot{\mathbf{r}}_l)\,, \quad \textit{with} \;\; j,k,l = 1,2,\dots,N, \;\; \textit{(no sum)}\,.$$

*This force law is independent of reference location, time or state of uniform translation at constant velocity. It will also be independent of reference orientation and thus it will be **Galilean invariant**, provided the forces $\mathbf{F}_j$ transform under rotations and parity reflections as*

$$m_j O\ddot{\mathbf{r}}_j = O\mathbf{F}_j = \mathbf{F}_j\Big(O(\mathbf{r}_k - \mathbf{r}_l), O(\dot{\mathbf{r}}_k - \dot{\mathbf{r}}_l)\Big)\,,$$

for any orthogonal transformation O. (The inverse of an orthogonal transformation is its transpose, $O^{-1} = O^T$. Such transformations include rotations and reflections. They preserve both lengths and relative orientations of vectors.)

Exercise. Prove that orthogonal transformations preserve both lengths and relative orientations of vectors.

★

Newton's Law #3 applies to closed systems.

Definition

2.15 (Closed system)**.**
*A system of N material points with masses m_j at positions $\mathbf{r}_j$, $j = 1, 2, \dots, N$, acted on by forces $\mathbf{F}_j$ is said to be **closed** if*

$$\mathbf{F}_j = \sum_{k\neq j} \mathbf{F}_{jk} \quad \textit{where} \quad \mathbf{F}_{jk} = -\mathbf{F}_{kj}\,. \tag{2.2}$$

Remark

2.16. *Newton's law of gravitational force applies to closed systems, since*

$$\mathbf{F}_{jk} = \frac{\gamma\, m_j m_k}{|\mathbf{r}_{jk}|^3}\mathbf{r}_{jk}\,, \quad \textit{where} \quad \mathbf{r}_{jk} = \mathbf{r}_j - \mathbf{r}_k\,, \tag{2.3}$$

with gravitational constant γ.

Exercise. Prove that Newton's law of motion

$$m_j \ddot{\mathbf{r}}_j = \sum_{k \neq j} \mathbf{F}_{jk}\,, \tag{2.4}$$

with gravitational forces $\mathbf{F}_{jk}$ in (2.3) is Galilean invariant. ★

(#3) Law of Reciprocal Actions For closed mechanical systems, **action equals reaction**. That is,

$$\mathbf{F}_{jk} = -\mathbf{F}_{kj}\,. \tag{2.5}$$

Corollary

2.17 (Action, reaction, momentum conservation). *For two particles,* ***action equals reaction*** *implies* $\dot{\mathbf{p}}_1 + \dot{\mathbf{p}}_2 = 0$*, for* $\mathbf{p}_j = m_j \mathbf{v}_j$ *(no sum on* j*).*

Proof. For two particles, $\dot{\mathbf{p}}_1 + \dot{\mathbf{p}}_2 = m_1 \dot{\mathbf{v}}_1 + m_2 \dot{\mathbf{v}}_2 = \mathbf{F}_{12} + \mathbf{F}_{21} = 0\,.$ □

2.4 Dynamical quantities

Definition

2.18 (Dynamical quantities).
The following dynamical quantities are often useful in characterising particle systems:

- *Kinetic energy,* $K = \frac{1}{2} m |\mathbf{v}|^2\,;$
- *Momentum,* $\mathbf{p} = \partial K / \partial \mathbf{v} = m\mathbf{v}\,;$
- *Moment of inertia,* $\mathbb{I} = m|\mathbf{r}|^2 = m\Big\langle \mathbf{r}\,,\,\mathbf{r} \Big\rangle\,;$

- *Centre of mass of a particle system,* $\mathbf{R}_{CM} = \sum_j m_j \mathbf{r}_j / \sum_k m_k$ *;*
- *Angular momentum,* $\mathbf{J} = \mathbf{r} \times \mathbf{p}$ *.*

Proposition

2.19 (Total momentum of a closed system)**.**
Let $\mathbf{P} = \sum_j \mathbf{p}_j$ *and* $\mathbf{F} = \sum_j \mathbf{F}_j$*, so that* $\dot{\mathbf{p}}_j = \mathbf{F}_j$*. Then* $\dot{\mathbf{P}} = \mathbf{F} = 0$ *for a closed system. Thus, a closed system conserves its total momentum* $\mathbf{P}$*.*

Proof. As for the case of two particles, sum the motion equations and use the definition of a closed system to verify conservation of its total momentum. □

Corollary

2.20 (Uniform motion of centre of mass)**.**
The ***centre of mass*** *for a closed system is defined as*

$$\mathbf{R}_{CM} = \sum_j m_j \mathbf{r}_j / \sum_k m_k \,.$$

Thus, the centre of mass velocity is

$$\mathbf{V}_{CM} = \dot{\mathbf{R}}_{CM} = \sum_j m_j \mathbf{v}_j / M = \mathbf{P}/M \,,$$

where $M = \sum_k m_k$ *is the total mass of the* N *particles.*

For a closed system,

$$\dot{\mathbf{P}} = 0 = M \ddot{\mathbf{R}}_{CM} \,,$$

so the centre of mass for a closed system is in uniform motion. Thus, it defines an inertial frame called the ***centre-of-mass frame****.*

Proposition

2.21 (Work rate of a closed system)**.**
Let

$$K = \frac{1}{2} \sum_j m_j |\mathbf{v}_j|^2$$

be the total kinetic energy of a closed system. Its time derivative is

$$\frac{dK}{dt} = \sum_j \langle m_j \dot{\mathbf{v}}_j \,,\, \mathbf{v}_j \rangle = \sum_j \langle \mathbf{F}_j \,,\, \mathbf{v}_j \rangle \,.$$

This expression defines the rate at which the forces within the closed system perform work.

Definition

2.22 (Conservative forces)**.**
*The forces $\mathbf{F}_j(\mathbf{r}_1, \dots \mathbf{r}_N)$ for a closed system of N particles are **conservative**, if*

$$\begin{aligned} \sum_j \Big\langle \mathbf{F}_j \,,\, d\mathbf{r}_j \Big\rangle &= -\, dV(\mathbf{r}_1, \dots, \mathbf{r}_N) \\ &:= \sum_j \Big\langle -\, \frac{\partial V}{\partial \mathbf{r}_j} \,,\, d\mathbf{r}_j \Big\rangle \,, \end{aligned} \tag{2.6}$$

*where dV is the **differential** of a smooth function $V : \mathbb{R}^N \to \mathbb{R}$ which called the **potential**, or **potential energy**.*

Remark

2.23. *For conservative forces the potential is independent of the particle velocities. That is, $\partial V / \partial \dot{\mathbf{r}}_j = 0$ for all $j = 1, \dots, N$.*

Proposition

2.24 (Energy conservation)**.**
If the forces are conservative, then the total energy

$$E = K + V$$

is a constant of motion for a closed system.

Proof. This result follows from the definition of work rate of a closed system, so that

$$\frac{dK}{dt} = \sum_j \Big\langle \mathbf{F}_j \,,\, \frac{d\mathbf{r}_j}{dt} \Big\rangle = -\, \frac{dV}{dt} \,, \tag{2.7}$$

for conservative forces. □

Proposition

2.25 (A class of conservative forces)**.**
Forces that depend only on relative distances between pairs of particles are conservative.

Proof. Suppose $\mathbf{F}_{jk} = f_{jk}(|\mathbf{r}_{jk}|)\mathbf{e}_{jk}$ with $\mathbf{e}_{jk} = \mathbf{r}_{jk}/|\mathbf{r}_{jk}| = \partial|\mathbf{r}_{jk}|/\partial\mathbf{r}_{jk}$ (no sum). In this case,

$$
\begin{aligned}
\langle \mathbf{F}_{jk}\,,\, d\mathbf{r}_{jk}\rangle &= \langle f_{jk}(|\mathbf{r}_{jk}|)\mathbf{e}_{jk}\,,\, d\mathbf{r}_{jk}\rangle \\
&= f_{jk}(|\mathbf{r}_{jk}|)\, d|\mathbf{r}_{jk}| \\
&= -\,dV_{jk} \quad \text{where} \quad V_{jk} = -\int f_{jk}(|\mathbf{r}_{jk}|)d|\mathbf{r}_{jk}|\,.
\end{aligned}
$$

□

Example

2.26 (Conservative force)**.**
The gravitational, or Coulomb force between particles j and k satisfies

$$
\mathbf{F}_{jk} = \frac{\gamma\, m_j m_k}{|\mathbf{r}_{jk}|^3}\mathbf{r}_{jk} = -\frac{\partial V_{jk}}{\partial \mathbf{r}_{jk}}\,,
$$

where

$$
V_{jk} = \frac{\gamma\, m_j m_k}{|\mathbf{r}_{jk}|} \quad \textit{and} \quad \mathbf{r}_{jk} = \mathbf{r}_j - \mathbf{r}_k\,,
$$

so it is a conservative force.

Proposition

2.27 (Total angular momentum)**.**
A closed system of interacting particles conserves its total angular momentum $\mathbf{J} = \sum_i \mathbf{J}_i = \sum_i \mathbf{r}_i \times \mathbf{p}_i$.

Proof. As for the case of two particles, one computes the time derivative for the sum,

$$
\dot{\mathbf{J}} = \sum_i \dot{\mathbf{J}}_i = \sum_i \dot{\mathbf{r}}_i \times \mathbf{p}_i + \sum_i \mathbf{r}_i \times \dot{\mathbf{p}}_i = \sum_{i,j} \mathbf{r}_i \times \mathbf{F}_{ij}\,,
$$

since $\dot{\mathbf{p}}_i = \sum_j \mathbf{F}_{ij}$ in the absence of external forces and $\dot{\mathbf{r}}_i \times \mathbf{p}_i$ vanishes. Rewriting this as a sum over pairs $i < j$ yields

$$\dot{\mathbf{J}} = \sum_{i,j} \mathbf{r}_i \times \mathbf{F}_{ij} = \sum_{i<j} (\mathbf{r}_i - \mathbf{r}_j) \times \mathbf{F}_{ij} = \mathbf{T}\,.$$

That is, the total angular momentum is conserved, provided the ***total torque*** $\mathbf{T}$ vanishes in this equation. When $\mathbf{T}$ vanishes, the total angular momentum $\mathbf{J}$ is conserved. □

Corollary

2.28 (Conserving total angular momentum)**.**
In particular, Proposition 2.27 implies that total angular momentum $\mathbf{J}$ *is constant for a closed system of particles interacting via central forces, for which force* $\mathbf{F}_{ij}$ *is parallel to the inter-particle displacement* $(\mathbf{r}_i - \mathbf{r}_j)$*.*

2.5 Newtonian form of free rigid rotation

Definition

2.29. *In* ***free rigid rotation****, a set of points undergoes rotation about its centre of mass and the pairwise distances between the points all remain fixed.*

A system of coordinates in free rigid motion is stationary in a rotating orthonormal basis. This rotating orthonormal basis is given by

$$\mathbf{e}_a(t) = O(t)\mathbf{e}_a(0)\,, \quad a = 1, 2, 3, \tag{2.8}$$

in which $O(t)$ is an orthogonal 3×3 matrix, so that $O^{-1} = O^T$. The three unit vectors $\mathbf{e}_a(0)$ with $a = 1, 2, 3$, denote an orthonormal basis for fixed reference coordinates. This basis may be taken as being aligned with any choice of fixed spatial coordinates at the initial time, $t = 0$. For example one may choose an initial alignment so that $O(0) = Id$.

Each point $\mathbf{r}(t)$ in rigid motion may be represented in either coordinate basis as

$$\begin{aligned} \mathbf{r}(t) &= r^A(t)\mathbf{e}_A(0)\,, \quad \text{in the fixed (spatial) basis,} && (2.9)\\ &= r^a\mathbf{e}_a(t)\,, \quad \text{in the rotating (body) basis,} && (2.10) \end{aligned}$$

and the *fixed* components r^a relative to the rotating basis satisfy $r^a = \delta^a_A r^A(0)$ for the choice that the two bases are initially aligned. (Otherwise, δ^a_A is replaced by an orthogonal 3×3 matrix describing the initial rotational misalignment.) The fixed basis $\mathbf{e}_a(0)$ is called the

spatial frame and the rotating basis $\mathbf{e}_a(t)$ is called the **body frame**. The components of vectors in the spatial frame are related to those in the body frame by the mutual rotation of their axes in (2.8) at any time. In particular,

$$\mathbf{e}_a(0) = O^{-1}(t)\mathbf{e}_a(t)\,, \quad a = 1,2,3. \tag{2.11}$$

Lemma

2.30. *The velocity* $\dot{\mathbf{r}}(t)$ *of a point* $\mathbf{r}(t)$ *in free rigid rotation depends* linearly *on its position relative to the centre of mass.*

Proof. In particular, $\mathbf{r}(t) = r^a O(t)\mathbf{e}_a(0)$ implies

$$\dot{\mathbf{r}}(t) = r^a\dot{\mathbf{e}}_a(t) = r^a\dot{O}(t)\mathbf{e}_a(0) =: r^a\dot{O}O^{-1}(t)\mathbf{e}_a(t) =: \widehat{\omega}(t)\mathbf{r}\,, \tag{2.12}$$

which is linear in $\mathbf{r}(t)$. □

Being orthogonal, the matrix $O(t)$ satisfies $OO^T = Id$ and one may compute that

$$\begin{aligned} 0 = (OO^T)^{\cdot} \;&=\; \dot{O}O^T + O\dot{O}^T \\ &=\; \dot{O}O^T + (\dot{O}O^T)^T \\ &=\; \dot{O}O^{-1} + (\dot{O}O^{-1})^T \\ &=\; \widehat{\omega} + \widehat{\omega}^T\,. \end{aligned}$$

This computation implies the following.

Lemma

2.31 (Skew symmetry).
The matrix $\widehat{\omega}(t) = \dot{O}O^{-1}(t)$ *in (2.12) is skew symmetric. That is,*

$$\widehat{\omega}^T = -\,\widehat{\omega}\,.$$

Definition

2.32 (Hat map for the angular velocity vector).
The skew symmetry of $\hat{\omega}$ *allows one to introduce the corresponding* ***angular velocity vector*** $\boldsymbol{\omega}(t) \in \mathbb{R}^3$ *whose components* $\omega_c(t)$*, with* $c = 1,2,3$*, are given by*

$$(\dot{O}O^{-1})_{ab}(t) = \widehat{\omega}_{ab}(t) = -\,\epsilon_{abc}\,\omega_c(t)\,. \tag{2.13}$$

Equation (2.13) defines the ***hat map****, which is an isomorphism between* 3×3 *skew-symmetric matrices and vectors in* $\mathbb{R}^3$*.*

According to this definition, one may write the matrix components of $\widehat{\omega}$ in terms of the vector components of $\boldsymbol{\omega}$ as

$$\widehat{\omega} = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}, \tag{2.14}$$

and the velocity in space of a point at $\mathbf{r}$ undergoing rigid body motion is found by the matrix multiplication $\widehat{\omega}(t)\mathbf{r}$ as

$$\dot{\mathbf{r}}(t) =: \widehat{\omega}(t)\mathbf{r} =: \boldsymbol{\omega}(t) \times \mathbf{r}\,. \tag{2.15}$$

Hence, the velocity of free rigid motion of a point displaced by $\mathbf{r}$ from the centre of mass is a rotation in space of $\mathbf{r}$ about the time-dependent angular velocity vector $\boldsymbol{\omega}(t)$.

Exercise. Compute the kinetic energy of free rigid motion about the centre of mass of a system of N points of mass m_j, with $j = 1, 2, \ldots, N$, located at distances $\mathbf{r}_j$ from the centre of mass, as

$$K = \frac{1}{2}\sum_j m_j|\dot{\mathbf{r}}_j|^2 = \frac{1}{2}\sum_j m_j|\boldsymbol{\omega} \times \mathbf{r}_j|^2 = \frac{1}{2}\boldsymbol{\omega}\cdot\mathbb{I}\boldsymbol{\omega}\,.$$

Use this expression to define the moment of inertia $\mathbb{I}$ of the system of N particles in terms of their masses and distances from the centre of mass. ★

Definition

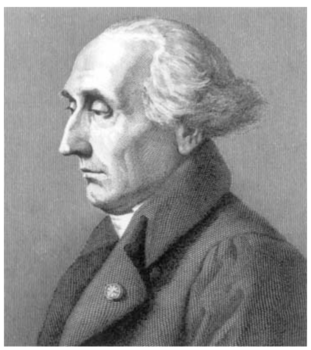
2.33 (Angular momentum of rigid rotation)**.**
*The **angular momentum** is defined as the derivative of the kinetic energy with respect to angular velocity. In the present case, this definition produces the linear relation,*

$$\begin{aligned} \mathbf{J} = \frac{\partial K}{\partial \boldsymbol{\omega}} &= -\sum_{j=1}^{N} m_j \mathbf{r}_j \times \Big(\mathbf{r}_j \times \boldsymbol{\omega}\Big) \\ &= \sum_{j=1}^{N} m_j \Big(|\mathbf{r}_j|^2 Id - \mathbf{r}_j \otimes \mathbf{r}_j\Big)\boldsymbol{\omega} \\ &=: \mathbb{I}\,\boldsymbol{\omega}\,, \end{aligned} \tag{2.16}$$

where $\mathbb{I}$ is the moment of inertia tensor defined by the kinetic energy of free rotation.

Exercise. Show that the definition of angular momentum $\mathbf{J} = \partial K/\partial \boldsymbol{\omega}$ in (2.16) recovers the previous one $\mathbf{J} = \mathbf{r} \times \mathbf{p}$ for a single rotating particle. ★

3 Lagrange


Lagrange

In 1756, at the age of 19, Lagrange sent a letter to Euler in which he proposed the solution to an outstanding problem dating from antiquity. The isoperimetric problem solved in Lagrange's letter may be stated as follows: Among all closed curves of a given fixed perimeter in the plane, which curve maximises the area that it encloses? The circle does this. However, Lagrange's method of solution was more important than the answer. Lagrange's solution to the isoperimetric problem problem laid down the principles for the **calculus of variations** and perfected results which Euler himself had introduced.

Lagrange used the calculus of variations to re-formulate Newtonian mechanics as the **Euler-Lagrange** equations. These equations are covariant: they take the same form in any coordinate system. Specifically, the Euler-Lagrange equations appear in the same form in coordinates on any **smooth manifold**, that is, on any space that admits the operations of calculus in local coordinates. This formulation, called Lagrangian mechanics, is also the language in which mechanics may be extended from finite to infinite dimensions.

3.1 Basic definitions for manifolds

Definition

3.1 (Smooth manifold).
*A **smooth manifold** M is a set of points together with a finite (or perhaps countable) set of subsets $U_\alpha \subset M$ and 1-to-1 mappings $\phi_\alpha : U_\alpha \to \mathbb{R}^n$ such that*

1. $\bigcup_\alpha U_\alpha = M$.
2. *For every nonempty intersection $U_\alpha \cap U_\beta$, the set $\phi_\alpha(U_\alpha \cap U_\beta)$ is an open subset of $\mathbb{R}^n$ and the 1-to-1 mapping $\phi_\beta \circ \phi_\alpha^{-1}$ (called the **transition function**) is a smooth function on $\phi_\alpha(U_\alpha \cap U_\beta)$.*

Remark

3.2. *As a practical matter, a smooth manifold of dimension k is a space that is locally isomorphic to $\mathbb{R}^k$ and admits calculus operations in its local coordinates. The most common examples of smooth manifolds are smooth curves on the plane (e.g., the circle $x^2 + y^2 = 1$) or curves and surfaces in three-dimensional Euclidean space $\mathbb{R}^3$. Riemann's treatment of the sphere S^2 in $\mathbb{R}^3$ is a famous example of how to show that a set of points defines a manifold.*

Example

3.3 (Stereographic projection of $S^2 \to \mathbb{R}^2$).
The unit sphere S^2 may be defined as a surface in $\mathbb{R}^3$ given by the set of points satisfying

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

The spherical polar angle θ and azimuthal angle ϕ are often used as coordinates on S^2. However, the angle ϕ cannot be defined uniquely at the North and South poles, where $\theta = 0$ and $\theta = \pi$, respectively. Riemann's treatment used Ptolemy's stereographic projection to define two overlapping subsets that satisfied the defining properties of a smooth manifold.

Let $U_N = S^2 \backslash \{0, 0, 1\}$ and $U_S = S^2 \backslash \{0, 0, -1\}$ be the subsets obtained by deleting the North and South poles of S^2, respectively. The stereographic projections ϕ_N and ϕ_S from the North and South poles of the sphere onto the equatorial plane, $z = 0$, are defined respectively by

$$\phi_N : U_N \to \xi_N + i\eta_N = \frac{x + iy}{1 - z} = e^{i\phi} \cot(\theta/2),$$

$$\text{and} \qquad \phi_S : U_S \to \xi_S + i\eta_S = \frac{x - iy}{1 + z} = e^{-i\phi} \tan(\theta/2).$$

The union of these two subsets covers S^2. On the overlap of their projections, the coordinates $(\xi_N, \eta_N) \in \mathbb{R}^2$ and $(\xi_S, \eta_S) \in \mathbb{R}^2$ are related by

$$(\xi_N + i\eta_N)(\xi_S + i\eta_S) = 1\,.$$

According to Definition 3.1 these two properties show that $S^2 \in \mathbb{R}^3$ is a smooth manifold.

Exercise. Prove the formulas above for the complex numbers $\xi_N + i\eta_N$ and $\xi_S + i\eta_S$ in the stereographic projection. For this, it may be useful to start with the stereographic projection for the circle. ★

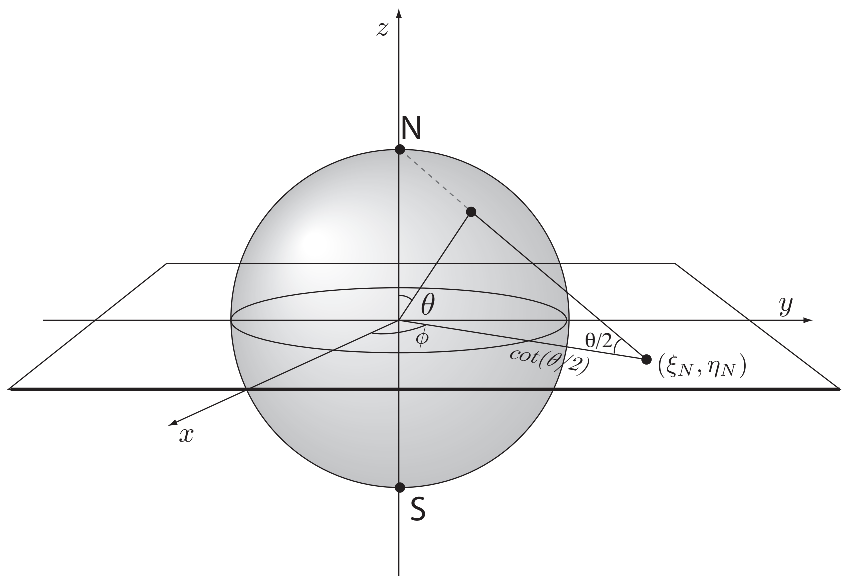


Figure 3: In the stereographic projection of the Riemann sphere onto the complex plane from the *North* pole, complex numbers lying outside (resp., inside) the unit circle are projected from points in the upper (resp., lower) hemisphere.

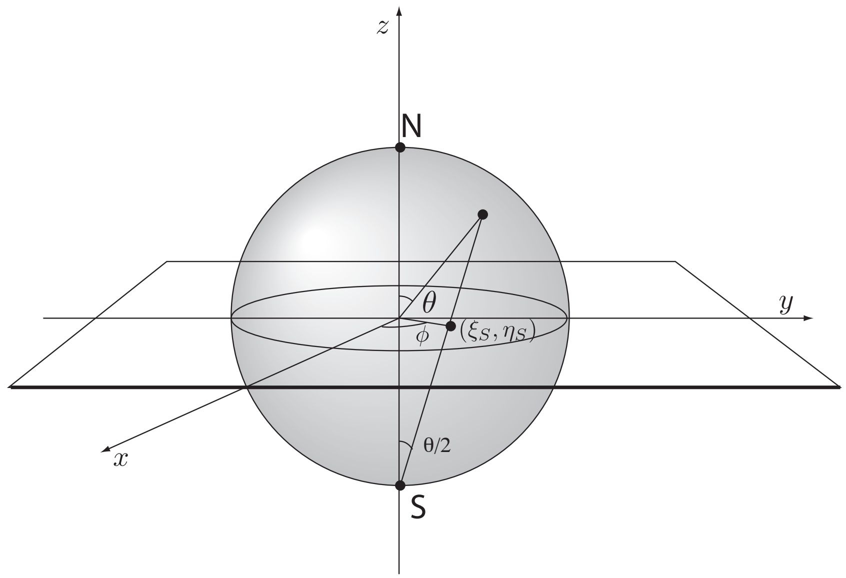


Figure 4: In the stereographic projection of the Riemann sphere onto the complex plane from the *South* pole, complex numbers lying outside (resp., inside) the unit circle are also projected from points in the upper (resp., lower) hemisphere.

Definition

3.4 (Submersion)**.**

A subspace $M \subset \mathbb{R}^n$ may be defined by the intersections of level sets of k smooth relations $f_i(x) = 0$,

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

with $\det(\partial f_i/\partial x^a) \neq 0$*,* $a = 1, 2, \dots, n$*, so that the gradients* $\partial f_i/\partial x^a$ *are linearly independent. Such a subspace M defined this way is called a* ***submersion*** *and has dimension* $\dim M = n - k$*.*

Remark

3.5. *A submersion is a particularly convenient type of smooth manifold. As we have seen,the unit sphere*

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

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

Exercise. Prove that all submersions are submanifolds. (For assistance, see Lee [Le2003].) ★

Remark

3.6 (Foliation of a manifold)**.** *A **foliation** looks locally like a decomposition of the manifold as a union of parallel submanifolds of smaller dimension. For example, the manifold* $\mathbb{R}^3/\{0\}$ *may be foliated by spheres* S^2*, which make up the **leaves** of the foliation. As another example, the two-dimensional* $\mathbb{R}^2$ *leaves of a book in* $\mathbb{R}^3$ *are enumerated by a (one-dimensional) page number.*

Definition

3.7 (Tangent space to level sets)**.**
Suppose the set

$$M = \big\{x \in \mathbb{R}^n \big| f_i(x) = 0, \ i = 1, \dots, k\big\}$$

with linearly independent gradients $\partial f_i / \partial x^a$, $a = 1, 2, \dots, n$, *is a smooth manifold in* $\mathbb{R}^n$. *The* tangent space *at each* $x \in M$, *is defined by*

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

*Note: in this expression we introduce the **Einstein summation convention**. That is, repeated indices are to be summed over their range.*

Remark

3.8. *The tangent space is a linear vector space.*

Example

3.9 (Tangent space to the sphere in $\mathbb{R}^3$)**.** *The sphere S^2 is the set of points $(x, y, z) \in \mathbb{R}^3$ solving $x^2 + y^2 + z^2 = 1$. The tangent space to the sphere at such a point (x, y, z) is the plane containing vectors (u, v, w) satisfying $xu + yv + zw = 0$.*

Definition

3.10 (Tangent bundle)**.** *The* tangent bundle *of a smooth manifold M, denoted by TM, is the smooth manifold whose underlying set is the disjoint union of the tangent spaces to M at the points $q \in M$; that is,*

$$TM = \bigcup_{q \in M} T_q M$$

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

Example

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

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

Remark

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

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

Example

3.13 (Tangent bundle TS^1 of the circle S^1).
The tangent bundle of the unit circle parameterised by an angle θ may be imagined in three dimensions as the union of the circle with a one-dimensional vector space of line vectors (the velocities $\dot{\theta}$) sitting over each point on the circle, shown in Figure 5.

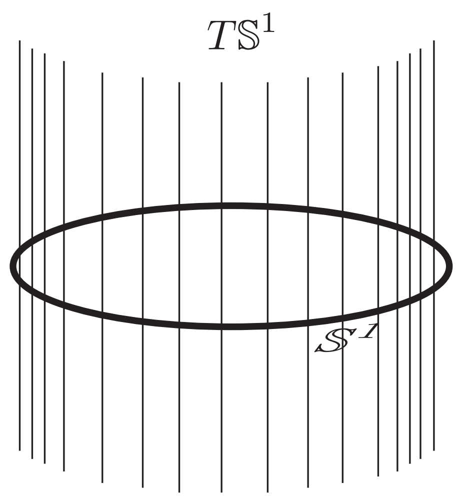


Figure 5: The tangent bundle TS^1 of the circle S^1 with coordinates $(\theta, \dot{\theta})$ is the union of the circle with a one-dimensional vector space of line vectors (the angular velocities).

Definition

3.14 (Vector fields).
A ***vector field*** *X on a manifold M is a map $X : M \to TM$ that assigns a vector $X(q)$ at every point $q \in M$. The real vector space of vector fields on a manifold M is denoted by $\mathfrak{X}(M)$.*

Definition

3.15. *A* ***time-dependent vector field*** *is a map*

$$X : M \times \mathbb{R} \to TM$$

such that $X(q,t) \in T_qM$ for each $q \in M$ and $t \in \mathbb{R}$.

Definition

3.16 (Integral curves)**.**
*An **integral curve** of vector field $X(q)$ with initial condition q_0 is a differentiable map $q :]t_1, t_2[\to M$ such that the open interval $]t_1, t_2[$ contains the initial time $t = 0$, at which $q(0) = q_0$ and the tangent vector coincides with the vector field*

$$\dot{q}(t) = X(q(t))$$

for all $t \in]t_1, t_2[$.

Remark

3.17. *In what follows we shall always assume we are dealing with vector fields that satisfy the conditions required for their integral curves to exist and be unique.*

Definition

3.18 (Vector basis)**.**
As in (3.4) a vector field $\dot{q}$ is defined by the components of its directional derivatives in the chosen coordinate basis, so that, for example,

$$\dot{q} = \dot{q}^a \frac{\partial}{\partial q^a} \quad \textit{(\textbf{vector basis})}. \tag{3.1}$$

In this vector basis, the vector field $\dot{q}$ has components $\dot{q}^a$, $a = 1, \dots, K$.

Definition

3.19 (Fibres of the tangent bundle)**.**
*The velocity vectors $(\dot{q}^1, \dot{q}^2, \dots, \dot{q}^K)$ in the tangent spaces T_qM to M at the points $q \in M$ in the tangent bundle TM are called the **fibres** of the bundle.*

Definition

3.20 (Dual vector space).
Any finite dimensional vector space V possesses a dual vector space of the same dimension. The dual space V^* consists of all linear functions $V \to \mathbb{R}$. The dual to the tangent space T_qM is called the ***cotangent space*** *T_q^*M to the manifold M at a point $q \in M$.*

Definition

3.21 (Dual basis).
The differential $df \in T_q^*M$ of a smooth real function $f: M \to \mathbb{R}$ is expressed in terms of the basis dq^b, $b = 1, \dots, K$, that is ***dual*** *to $\partial/\partial q^a$, $a = 1, \dots, K$, as*

$$df(q) = \frac{\partial f}{\partial q^b}\, dq^b \quad (\textit{\textbf{dual basis}}). \tag{3.2}$$

That is, the linear function $df(q) : T_qM \to \mathbb{R}$ lives in the space T_q^*M dual to the vector space T_qM.

Definition

3.22 (Contraction).
The operation $\lrcorner$ of ***contraction*** *between elements of a vector basis and its dual basis is defined in terms of a* ***nondegenerate symmetric pairing***

$$\langle \,\cdot\,,\,\cdot\, \rangle : T_qM \times T_q^*M \to \mathbb{R},$$

as the bilinear relation

$$\left\langle \frac{\partial}{\partial q^b}\,,\, dq^a \right\rangle := \frac{\partial}{\partial q^b} \lrcorner\, dq^a = \delta^a_b\,,$$

where δ^a_b is the Kronecker delta. That is, $\delta^a_b = 0$ for $a \neq b$ and $\delta^a_b = 1$ for $a = b$.

Definition

3.23 (Directional derivative).
The ***directional derivative*** *of a smooth function $f: M \to \mathbb{R}$ along the vector $\dot q \in T_qM$ is defined as*

$$\dot q \lrcorner\, df = \langle\, \dot q\,,\, df(q)\,\rangle = \left\langle\, \dot q^b \frac{\partial}{\partial q^b}\,,\, \frac{\partial f}{\partial q^a} dq^a \,\right\rangle = \frac{\partial f}{\partial q^a}\, \dot q^a = \frac{d}{dt} f(q(t))\,.$$

Definition

3.24 (Cotangent space to a smooth manifold).
The space of differentials $df(q)$ of smooth functions f defined on a manifold M at a point $q \in M$ forms a dual vector space called the ***cotangent space*** *of M at $q \in M$ which is denoted as T_q^*M.*

Definition

3.25 (Cotangent bundle of a manifold).
The disjoint union of cotangent spaces to M at the points $q \in M$ given by

$$T^*M = \bigcup_{q \in M} T_q^*M \tag{3.3}$$

is a vector space called the ***cotangent bundle*** *of M and is denoted as T^*M.*

Remark

3.26 (Covariant versus contravariant vectors).
Historically, the components of vector fields were called ***contravariant*** *while the components of differential one-forms were called* ***covariant****. The covariant and contravariant components of vectors and tensors are distinguished by their coordinate transformation properties under changes of vector basis and dual basis for a change of coordinates $q \to y(q)$. For example, the components in a new vector basis are*

$$\dot{q} \;\;=\;\; \Big(\dot{q}^a \frac{\partial y^b}{\partial q^a}\Big)\frac{\partial}{\partial y^b} =: \dot{y}^b \frac{\partial}{\partial y^b} = \dot{y}\,,$$

while the components in a new dual basis are

$$df(q) \;\;=\;\; \Big(\frac{\partial f}{\partial q^b}\,\frac{\partial q^b}{\partial y^a}\Big)\,dy^a =: \frac{\partial f}{\partial y^a}\,dy^a = df(y)\,.$$

Thus, as every physicist learns about covariant and contravariant vectors and tensors,

"By their transformations shall ye know them."
– A. Sommerfeld (private communication, O. Laporte)

Exercise. Consider the following mixed tensor

$$T(q) = T^{abc}_{ijk}(q)\frac{\partial}{\partial q^a}\otimes\frac{\partial}{\partial q^b}\otimes\frac{\partial}{\partial q^c}\otimes dq^i\otimes dq^j\otimes dq^k,$$

in which $\otimes$ denotes direct (or, tensor) product. How do the components of the mixed tensor T transform under a change of coordinates $q \to y(q)$?

That is, write the components of $T(y)$ in the new basis in terms of the Jacobian matrix for the change of coordinates and the components $T^{abc}_{ijk}(q)$ of $T(q)$. ★

3.2 Euler-Lagrange equation on a manifold

3.2.1 Motion on a K-dimensional submanifold of $\mathbb{R}^{3N}$

Consider the motion of N particles undergoing conservative motion on a smooth K-dimensional manifold $M \subset \mathbb{R}^{3N}$.

Let $q = (q^1, q^2, \dots q^K)$ be coordinates on the manifold M, which is defined as $\mathbf{r}_j = \mathbf{r}_j(q^1, q^2, \dots, q^K)$, with $j = 1, 2, \dots, N$. Consequently, a velocity vector $\dot{q}(t)$ tangent to a path $q(t)$ in the manifold M at point $q \in M$ induces a velocity vector in $\mathbb{R}^{3N}$ by

$$\dot{\mathbf{r}}_j(t) = \sum_{a=1}^{K}\frac{\partial\mathbf{r}_j}{\partial q^a}\,\dot{q}^a(t) \quad \text{for} \quad j = 1, 2, \dots N\,. \tag{3.4}$$

Remark

3.27. *The $2K$ numbers $q^1, q^2, \dots, q^K, \dot{q}^1, \dot{q}^2, \dots, \dot{q}^K$, provide a local coordinate system for T_qM, the tangent space to M at $q \in M$.*

Remark

3.28 (Generalised coordinates).
*The choice of coordinates q is arbitrary up to a **reparametrisation map** $q \to Q \in M$ with $\det(\partial Q/\partial q) \neq 0$. For this reason, the $\{q\}$ are called **generalised coordinates**.*

Theorem

3.29 (Euler-Lagrange equation).
Newton's law for conservative forces gives

$$\sum_{j=1}^{N} \Big\langle (m_j \ddot{\mathbf{r}}_j - \mathbf{F}_j)\,,\, d\mathbf{r}_j(q) \Big\rangle \Big|_{\mathbf{r}(q)} = \sum_{a=1}^{K} \Big\langle \big[\, L \,\big]_{q^a}\,,\, dq^a \Big\rangle\,,$$

where

$$\big[\, L \,\big]_{q^a} := \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^a} - \frac{\partial L}{\partial q^a}\,,$$

for particle motion tangent to a manifold $M \subset \mathbb{R}^{3N}$ with generalised coordinates q_a and for conservative forces. Here, the quantity

$$L(q,\dot{q}) := T(q,\dot{q}) - V(\mathbf{r}(q))$$

*is called the **Lagrangian** and $T(q,\dot{q})$ is the particle kinetic energy on M, namely,*

$$T(q,\dot{q}) \;=\; \frac{1}{2}\sum_{j=1}^{N} m_j |\mathbf{v}_j|^2 = \frac{1}{2}\sum_{j=1}^{N} m_j |\dot{\mathbf{r}}_j(q)|^2 = \frac{1}{2}\sum_{j=1}^{N} m_j \Big| \sum_{a=1}^{K} \frac{\partial \mathbf{r}_j}{\partial q^a}\dot{q}^a \Big|^2$$

The proof of the theorem proceeds by assembling the formulas for constrained acceleration and work rate in the next two lemmas, obtained by direct computations using Newton's law for conservative forces.

By the way, we have suspended the summation convention on repeated indices for a moment to avoid confusion between the two different types of indices for the particle label and for the coordinate components on the manifold TM.

Lemma

3.30 (Constrained acceleration formula).
The induced kinetic energy $T(q,\dot{q})$ on the manifold M satisfies

$$\sum_{j=1}^{N} \Big\langle m_j \ddot{\mathbf{r}}_j\,,\, d\mathbf{r}_j(q) \Big\rangle \Big|_{\mathbf{r}(q)} = \sum_{a=1}^{K} \Big\langle \frac{d}{dt}\frac{\partial T}{\partial \dot{q}^a} - \frac{\partial T}{\partial q^a}\,,\, dq^a \Big\rangle\,. \tag{3.5}$$

Proof. The constrained acceleration formula follows from differentiating $T(q,\dot q)$ to obtain

$$\frac{\partial T}{\partial \dot q^a} = \sum_{j=1}^N m_j \Big(\sum_{b=1}^K \frac{\partial \mathbf{r}_j}{\partial q^b} \dot q^b \Big) \cdot \frac{\partial \mathbf{r}_j}{\partial q^a} \,,$$

and

$$\frac{d}{dt}\frac{\partial T}{\partial \dot q^a} = \sum_{j=1}^N m_j \ddot{\mathbf{r}}_j \cdot \frac{\partial \mathbf{r}_j}{\partial q^a} + \frac{\partial T}{\partial q^a} \,.$$

$\square$

Lemma

3.31 (Work rate formula).
Forces $\mathbf{F}_j(\mathbf{r}_1,\dots,\mathbf{r}_N)$ *evaluated on the manifold* M *satisfy*

$$\sum_{j=1}^N \Big\langle \mathbf{F}_j \,,\, d\mathbf{r}_j(q) \Big\rangle \Big|_{\mathbf{r}(q)} = -\,dV\big(\mathbf{r}(q)\big) = -\sum_{a=1}^K \Big\langle \frac{\partial V\big(\mathbf{r}(q)\big)}{\partial q^a} \,,\, dq^a \Big\rangle \,.$$

Proof. This Lemma follows, because the forces $\mathbf{F}_j(\mathbf{r}_1,\dots,\mathbf{r}_N)$ are conservative. $\square$

The proof of Theorem 3.29 now proceeds by assembling the formulas for constrained acceleration and work rate in the previous two lemmas, as a direct calculation.

Proof.

$$\begin{aligned} &\sum_{j=1}^N \Big\langle m_j \ddot{\mathbf{r}}_j - \mathbf{F}_j \,,\, d\mathbf{r}_j(q) \Big\rangle \Big|_{\mathbf{r}(q)} \\ &\quad = \sum_{a=1}^K \Big\langle \frac{d}{dt}\frac{\partial T(q,\dot q)}{\partial \dot q^a} - \frac{\partial\big(T(q,\dot q) - V(\mathbf{r}(q))\big)}{\partial q^a} \,,\, dq^a \Big\rangle \,. \end{aligned}$$

$\square$

Corollary

3.32 (Newton $\simeq$ Lagrange).
*Newton's law for the motion of N particles on a K-dimensional manifold $M \subset \mathbb{R}^{3N}$ defined as $\mathbf{r}_j = \mathbf{r}_j(q^1, q^2, \dots q^K)$, with $j = 1, 2, \dots N$, is equivalent to the **Euler-Lagrange equations**,*

$$\big[\, L \,\big]_{q^a} := \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^a} - \frac{\partial L}{\partial q^a} = 0\,, \tag{3.6}$$

for motion tangent to the manifold M and for conservative forces with Lagrangian

$$L(q, \dot{q}) := T(q, \dot{q}) - V\big(\mathbf{r}(q)\big)\,. \tag{3.7}$$

Proof. This corollary of Theorem 3.29 follows by independence of the differential basis elements dq^a in the final line of its proof. □

Theorem

3.33 (Hamilton's principle of stationary action).
*The **Euler-Lagrange equation**,*

$$\big[\, L \,\big]_{q^a} := \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^a} - \frac{\partial L}{\partial q^a} = 0\,, \tag{3.8}$$

*follows from stationarity of the **action**, S, defined as the integral over a time interval $t \in (t_1\,,\, t_2)$*

$$S = \int_{t_1}^{t_2} L(q, \dot{q})\, dt\,. \tag{3.9}$$

*Then **Hamilton's principle of stationary action**,*

$$\delta S = 0\,, \tag{3.10}$$

implies $[\, L \,]_{q^a} = 0$, for variations δq^a that are tangent to the manifold M and which vanish at the endpoints in time.

Proof. Notation in this proof is simplified by suppressing superscripts a in q^a and only writing q. The meaning of the variational derivative in the statement of Hamilton's principle is the following. Consider a family of C^2 curves $q(t,s)$ for $|s| < \varepsilon$ satisfying $q(t,0) = q(t)$, $q(t_1,s) = q(t_1)$, and $q(t_2,s) = q(t_2)$ for all $s \in (-\varepsilon,\varepsilon)$. The ***variational derivative*** of the action S is defined as

$$\delta S = \delta \int_{t_1}^{t_2} L(q(t),\dot{q}(t))dt := \frac{d}{ds}\bigg|_{s=0} \int_{t_1}^{t_2} L(q(t,s),\dot{q}(t,s))dt\,. \tag{3.11}$$

Differentiating under the integral sign, denoting

$$\delta q(t) := \frac{d}{ds}\bigg|_{s=0} q(t,s)\,, \tag{3.12}$$

and integrating by parts produces

$$\begin{aligned}\delta S &= \int_{t_1}^{t_2} \Big(\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta \dot{q}\Big)\,dt \\ &= \int_{t_1}^{t_2} \Big(\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\Big)\delta q\,dt + \left[\frac{\partial L}{\partial \dot{q}}\delta q\right]_{t_1}^{t_2},\end{aligned} \tag{3.13}$$

where one exchanges the order of derivatives by using $q_{st} = q_{ts}$ so that $\delta\dot{q} = \frac{d}{dt}\delta q$. Vanishing of the variations at the endpoints $\delta q(t_1) = 0 = \delta q(t_2)$ then causes the last term to vanish, which finally yields

$$\delta S = \int_{t_1}^{t_2} \Big(\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta \dot{q}\Big)\,dt = \int_{t_1}^{t_2} \Big(\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\Big)\delta q\,dt\,.$$

The action S is stationary $\delta S = 0$ for an arbitrary C^1 function $\delta q(t)$ if and only if the Euler-Lagrange equations (3.6) hold, that is, provided $[\,L\,]_q = 0$. □

Corollary

3.34 (Noether's theorem)**.** *Each smooth symmetry of the Lagrangian in Hamilton's principle implies a conservation law for its Euler–Lagrange equation [KoSc2001].*

Proof. In Hamilton's principle $\delta S = 0$ for $S = \int_{t_1}^{t_2} L(\mathbf{q},\dot{\mathbf{q}})\,dt$ as in (3.9), the Lagrangian L has a symmetry if it is *invariant* under the transformations $\mathbf{q}(t,0) \to \mathbf{q}(t,\varepsilon)$. In this case, stationarity $\delta S = 0$ under the infinitesimal variations defined in (3.12) follows because of this invariance of the Lagrangian, even if these variations did *not* vanish at the endpoints in time. The variational calculation (3.13) in vector notation

$$0 = \delta S = \int \underbrace{\left(\frac{\partial L}{\partial \mathbf{q}} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{q}}} \right)}_{\text{Euler–Lagrange}} \cdot \delta \mathbf{q} \, dt + \underbrace{\left[\frac{\partial L}{\partial \dot{\mathbf{q}}} \cdot \delta \mathbf{q} \right]_{t_1}^{t_2}}_{\text{Noether}} \tag{3.14}$$

then shows that along the solution paths of the Euler–Lagrange equation (3.8) any smooth symmetry of the Lagrangian L implies

$$\left[\frac{\partial L}{\partial \dot{\mathbf{q}}} \cdot \delta \mathbf{q} \right]_{t_1}^{t_2} = 0 \, .$$

Thus, the quantity $\delta \mathbf{q} \cdot (\partial L / \partial \dot{\mathbf{q}})$ is a constant of the motion (i.e., it is constant along the solution paths of the Euler–Lagrange equation) whenever $\delta S = 0$, because of the symmetry of the Lagrangian L in the action $S = \int L \, dt$. □

Exercise. What does Noether's theorem imply for symmetries of the action principle given by $\delta S = 0$ for the following action?

$$S = \int_{t_1}^{t_2} L(\dot{\mathbf{q}}(t)) dt \, .$$

★

Answer. In this case, $\partial L / \partial \mathbf{q} = 0$. This means the action S is invariant under translations $\mathbf{q}(t) \to \mathbf{q}(t) + \boldsymbol{\varepsilon}$ for any constant vector $\boldsymbol{\varepsilon}$. Setting $\delta \mathbf{q} = \boldsymbol{\varepsilon}$ in the Noether theorem associates conservation of any component of $\mathbf{p} := (\partial L / \partial \dot{\mathbf{q}})$ with invariance of the action under spatial translations in that direction. For this case, the conservation law also follows immediately from the Euler–Lagrange equation (3.8). ▲

3.3 Geodesic motion on Riemannian manifolds

The kinetic energy in Theorem 3.29 may be rewritten as

$$
\begin{aligned}
T(q,\dot q) &= \frac{1}{2}\sum_{j=1}^{N} m_j \sum_{a,b=1}^{K} \left(\frac{\partial \mathbf{r}_j}{\partial q^a}\cdot\frac{\partial \mathbf{r}_j}{\partial q^b}\right)\dot q^a \dot q^b \\
&=: \frac{1}{2}\sum_{j=1}^{N} m_j \sum_{a,b=1}^{K} (g_j(q))_{ab}\,\dot q^a \dot q^b ,
\end{aligned}
$$

which defines the quantity $(g_j(q))_{ab}$. For $N=1$, this reduces to

$$
T(q,\dot q) = \frac{1}{2} m\, g_{ab}(q)\,\dot q^a \dot q^b ,
$$

where we now reinstate the summation convention; that is, we again sum repeated indices over their range, which in this case is $a,b = 1,2,\dots,K$, where K is the dimension of the manifold M.

3.3.1 Free particle motion in a Riemannian space

The Lagrangian for the motion of a free particle of unit mass is its kinetic energy, which defines a **Riemannian metric** on the manifolm M (i.e., a nonsingular positive symmetric matrix depending smoothly on $q \in M$) that in turn yields a **norm** $\|\cdot\| : TM \to \mathbb{R}_+$, by

$$
L(q,\,\dot q) = \frac{1}{2}\dot q^b g_{bc}(q)\dot q^c =: \frac{1}{2}\|\dot q\|^2 \ \geq 0\,. \tag{3.15}
$$

The Lagrangian in this case has partial derivatives given by,

$$
\frac{\partial L}{\partial \dot q^a} = g_{ac}(q)\dot q^c \quad \text{and} \quad \frac{\partial L}{\partial q^a} = \frac{1}{2}\frac{\partial g_{bc}(q)}{\partial q^a}\dot q^b \dot q^c\,.
$$

Consequently, its Euler-Lagrange equations $[\,L\,]_{q^a} = 0$ are

$$
\begin{aligned}
[\,L\,]_{q^a} &:= \frac{d}{dt}\frac{\partial L}{\partial \dot q^a} - \frac{\partial L}{\partial q^a} \\
&= g_{ae}(q)\ddot q^{\,e} + \frac{\partial g_{ae}(q)}{\partial q^b}\dot q^b \dot q^e - \frac{1}{2}\frac{\partial g_{be}(q)}{\partial q^a}\dot q^b \dot q^e = 0\,.
\end{aligned}
$$

Symmetrising the coefficient of the middle term and contracting with co-metric g^{ca} satisfying $g^{ca}g_{ae} = \delta^c_e$ yields

$$\ddot{q}^{\,c} + \Gamma^c_{be}(q)\dot{q}^b\dot{q}^e = 0\,, \tag{3.16}$$

with

$$\Gamma^c_{be}(q) = \frac{1}{2}g^{ca}\left[\frac{\partial g_{ae}(q)}{\partial q^b} + \frac{\partial g_{ab}(q)}{\partial q^e} - \frac{\partial g_{be}(q)}{\partial q^a}\right], \tag{3.17}$$

in which the Γ^c_{be} are the **Christoffel symbols** for the Riemannian metric g_{ab}. These Euler-Lagrange equations are the **geodesic equations** of a free particle moving in a Riemannian space.

Exercise. Calculate the induced metric and Christoffel symbols for the sphere $x^2 + y^2 + z^2 = 1$, written in polar coordinates $(\theta.\phi)$ with $x + iy = e^{i\phi}\sin\theta$, $z = \cos\theta$. ★

Exercise. Calculate the Christoffel symbols when the metric in the Lagrangian takes the form in equation (??) for Fermat's principle; namely

$$L(\mathbf{q},\,\dot{\mathbf{q}}) = \frac{1}{2}n^2(\mathbf{q})\,\dot{q}^b\delta_{bc}\dot{q}^c\,, \tag{3.18}$$

in *Euclidean* coordinates $\mathbf{q} \in \mathbb{R}^3$ with a prescribed index of refraction $n(\mathbf{q})$. ★

3.3.2 Geodesic motion on the 3×3 special orthogonal matrices

A three-dimensional spatial rotation is described by multiplication of a spatial vector by a 3×3 special orthogonal matrix, denoted as $O \in SO(3)$,

$$O^TO = Id\,, \quad \text{so that} \quad O^{-1} = O^T \quad \text{and} \quad \det O = 1\,. \tag{3.19}$$

Geodesic motion on the space of rotations in three dimensions may be represented as a curve $O(t) \in SO(3)$ depending on time t. Its angular velocity is defined as the 3×3 matrix $\widehat{\Omega}$,

$$\widehat{\Omega}(t) = O^{-1}(t)\dot{O}(t)\,, \tag{3.20}$$

which must be skew-symmetric. That is, $\widehat{\Omega}^T = -\,\widehat{\Omega}$, where superscript $(\,\cdot\,)^T$ denotes matrix transpose.

Exercise. Show that the skew symmetry of $\widehat{\Omega}(t)$ follows by taking the time derivative of the defining relation for orthogonal matrices. ★

Answer. The time derivative of $O^T(t)O(t) = Id$ along the curve $O(t)$ yields $\big(O^T(t)O(t)\big)^{\cdot} = 0$, so that

$$0 = \dot{O}^T O + O^T \dot{O} = (O^T \dot{O})^T + O^T \dot{O}\,,$$

and, thus

$$(O^{-1}\dot{O})^T + O^{-1}\dot{O} = \widehat{\Omega}^T + \widehat{\Omega} = 0\,.$$

That is, $\widehat{\Omega}^T = -\,\widehat{\Omega}$. ▲

As for the time derivative, the variational derivative of $O^{-1}O = Id$ yields $\delta(O^{-1}O) = 0$, which leads to *another* skew-symmetric matrix, $\widehat{\Xi}$, defined by

$$\delta O^{-1} = -\,(O^{-1}\delta O)O^{-1}\,,$$

and

$$\widehat{\Xi} := O^{-1}\delta O = -(\delta O^{-1})O = -(O^{-1}\delta O)^T = -\,\widehat{\Xi}^T\,.$$

Lemma

3.35. *The variational derivative of the angular velocity $\widehat{\Omega} = O^{-1}\dot{O}$ satisfies*

$$\delta\widehat{\Omega} = \widehat{\Xi}^{\cdot} + \widehat{\Omega}\widehat{\Xi} - \widehat{\Xi}\widehat{\Omega}\,, \tag{3.21}$$

in which $\widehat{\Xi} = O^{-1}\delta O$.

Proof. The variational formula (3.21) follows by subtracting the time derivative $\widehat{\Xi}^{\cdot} = (O^{-1}\delta O)^{\cdot}$ from the variational derivative $\delta\widehat{\Omega} = \delta(O^{-1}\dot{O})$ in the relations

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

and using equality of cross derivatives $\delta\dot{O} = (\delta O)^{\cdot}$. □

Theorem

3.36 (Geodesic motion on $SO(3)$)**.**
The Euler-Lagrange equation for Hamilton's principle

$$\delta S = 0 \quad \textit{with} \quad S = \int L(\widehat{\Omega})\, dt\,, \tag{3.22}$$

using the quadratic Lagrangian $L : TSO(3) \to \mathbb{R}$,

$$L(\widehat{\Omega}) = -\frac{1}{2}\mathrm{tr}(\widehat{\Omega}\mathbb{A}\widehat{\Omega})\,, \tag{3.23}$$

in which $\mathbb{A}$ *is a symmetric, positive-definite* 3×3 *matrix, takes the matrix commutator form*

$$\frac{d\widehat{\Pi}}{dt} = -\,[\,\widehat{\Omega}\,,\,\widehat{\Pi}\,] \quad \textit{with} \quad \widehat{\Pi} = \mathbb{A}\widehat{\Omega} + \widehat{\Omega}\mathbb{A} = \frac{\delta L}{\delta\widehat{\Omega}} = -\,\widehat{\Pi}^T\,. \tag{3.24}$$

Proof. Taking matrix variations in this Hamilton's principle yields

$$\begin{aligned}
\delta S \;&=:\; \int_a^b \left\langle \frac{\delta L}{\delta\widehat{\Omega}}\,,\,\delta\widehat{\Omega} \right\rangle dt \\
&= -\frac{1}{2}\int_a^b \mathrm{tr}\left(\delta\widehat{\Omega}\,\frac{\delta L}{\delta\widehat{\Omega}} \right) dt \\
&= -\frac{1}{2}\int_a^b \mathrm{tr}\left(\delta\widehat{\Omega}\,(\mathbb{A}\widehat{\Omega} + \widehat{\Omega}\mathbb{A}) \right) dt \\
&= -\frac{1}{2}\int_a^b \mathrm{tr}\left(\delta\widehat{\Omega}\,\widehat{\Pi} \right) dt \\
&= \int_a^b \left\langle\, \widehat{\Pi}\,,\,\delta\widehat{\Omega}\,\right\rangle dt\,.
\end{aligned} \tag{3.25}$$

The first step uses $\delta\widehat{\Omega}^T = -\delta\widehat{\Omega}$ and expresses the pairing in the variational derivative of S for matrices as the ***trace pairing***, e.g.,

$$\left\langle\, M\,,\,N\,\right\rangle =: \frac{1}{2}\mathrm{tr}\Big(M^T\,N\Big) = \frac{1}{2}\mathrm{tr}\Big(N^T\,M\Big)\,. \tag{3.26}$$

The second step applies the variational derivative. After cyclically permuting the order of matrix multiplication under the trace, the fourth step substitutes

$$\widehat{\Pi} = \mathbb{A}\widehat{\Omega} + \widehat{\Omega}\mathbb{A} = \frac{\delta L}{\delta \widehat{\Omega}}\,.$$

Next, substituting formula (3.21) for $\delta\widehat{\Omega}$ into the variation of the action (3.25) leads to

$$\begin{aligned} \delta S &= -\frac{1}{2}\int_a^b \operatorname{tr}\left(\delta\widehat{\Omega}\widehat{\Pi}\right) dt \\ &= -\frac{1}{2}\int_a^b \operatorname{tr}\left(\left(\widehat{\Xi}^{\,\cdot} + \widehat{\Omega}\widehat{\Xi} - \widehat{\Xi}\widehat{\Omega}\right)\widehat{\Pi}\right) dt\,. \end{aligned} \tag{3.27}$$

Permuting cyclically under the trace again yields

$$\operatorname{tr}(\widehat{\Omega}\,\widehat{\Xi}\,\widehat{\Pi}) = \operatorname{tr}(\widehat{\Xi}\,\widehat{\Pi}\,\widehat{\Omega})\,.$$

Integrating by parts (dropping endpoint terms) then yields the equation

$$\delta S = -\frac{1}{2}\int_a^b \operatorname{tr}\left(\widehat{\Xi}\left(-\widehat{\Pi}^{\,\cdot} + \widehat{\Pi}\widehat{\Omega} - \widehat{\Omega}\widehat{\Pi}\right)\right) dt\,. \tag{3.28}$$

Finally, invoking stationarity $\delta S = 0$ for an arbitrary variation

$$\widehat{\Xi} = O^{-1}\delta O\,,$$

yields geodesic dynamics on $SO(3)$ with respect to the metric $\mathbb{A}$ in the commutator form (3.24). □

Remark

3.37 (Interpretation of Theorem 3.36).
Equation (3.24) for the matrix $\widehat{\Pi}$ *describes geodesic motion in the space of* 3×3 *orthogonal matrices with respect to the metric tensor* $\mathbb{A}$. *The matrix* $\widehat{\Pi}$ *is defined as the* ***fibre derivative*** *of the Lagrangian* $L(\widehat{\Omega})$ *with respect to the angular velocity matrix* $\widehat{\Omega}(t) = O^{-1}(t)\dot{O}(t)$. *Thus,* $\widehat{\Pi}$ *is the angular momentum matrix dual to the angular velocity matrix* $\widehat{\Omega}$.

Once the solution for $\widehat{\Omega}(t)$ *is known from the evolution of* $\widehat{\Pi}(t)$, *the orthogonal matrix orientation* $O(t)$ *is determined from one last integration in time, by using the equation*

$$\dot{O}(t) = O(t)\widehat{\Omega}(t)\,. \tag{3.29}$$

This is the ***reconstruction formula****, obtained from the definition (3.20) of the angular velocity matrix. In the classical literature, such an integration is called a* ***quadrature****.*

Corollary

3.38. *Formula (3.24) for the evolution of $\widehat{\Pi}(t)$ is equivalent to the conservation law*

$$\frac{d}{dt}\widehat{\pi}(t) = 0\,, \quad \textit{where} \quad \widehat{\pi}(t) := O(t)\widehat{\Pi}(t)O^{-1}(t)\,. \tag{3.30}$$

Proof. This may be verified by a direct computation that uses the reconstruction formula in (3.29). □

Remark

3.39. *The quantities $\widehat{\pi}$ and $\widehat{\Pi}$ in the rotation of a rigid body are called its spatial and body angular momentum, respectively.*

Exercise. (Noether's theorem)

What does Noether's theorem (Corollary 3.34) imply for geodesic motion on the special orthogonal group $SO(3)$? How does this generalise to $SO(n)$? What about Noether's theorem for geodesic motion on other groups?

Hint: consider the endpoint terms $\mathrm{tr}(\widehat{\Xi}\widehat{\Pi})\big|_a^b$ arising in the variation δS in (3.28) and invoke left-invariance of the Lagrangian (3.23) under $O \to U_\epsilon O$ with $U_\epsilon \in SO(3)$. For this symmetry transformation, $\delta O = \widehat{\Gamma} O$ with $\widehat{\Gamma} = \frac{d}{d\epsilon}\big|_{\epsilon=0} U_\epsilon$, so $\widehat{\Xi} = O^{-1}\widehat{\Gamma} O$.
★

3.4 Euler's equations for the motion of a rigid body

Besides describing geodesic motion in the space of 3×3 orthogonal matrices with respect to the metric tensor $\mathbb{A}$, the dynamics of $\widehat{\Pi}$ in (3.24) turns out to be the matrix version of Euler's equations for rigid body motion.

3.4.1 Physical interpretation of $SO(3)$ matrix dynamics

To see how Euler's equation for a rigid body emerges from geodesic motion in $SO(3)$ with respect to the metric $\mathbb{A}$, we shall use the hat map in equation (2.13) to convert the skew-symmetric matrix dynamics (3.24) into its vector form. Let the principal axes of inertia of the body be the orthonormal eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ of $\mathbb{A}$. Then its principal moments of inertia turn out to be linear combinations of the corresponding (positive) eigenvalues a_1, a_2, a_3. Setting

$$\boldsymbol{\Omega} = \Omega_1 \mathbf{e}_1 + \Omega_2 \mathbf{e}_2 + \Omega_3 \mathbf{e}_3 \,, \tag{3.31}$$

identifies vector components Ω_k, $k = 1, 2, 3$, with the components of the skew-symmetric matrix $\widehat{\Omega}_{ij}$, $i, j = 1, 2, 3$, as

$$\widehat{\Omega}_{ij} = -\, \epsilon_{ijk} \Omega_k \,, \tag{3.32}$$

which takes the skew-symmetric matrix form of (2.15)

$$\widehat{\Omega} = \begin{pmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{pmatrix} . \tag{3.33}$$

This identification yields for $\widehat{\Pi} = \mathbb{A}\widehat{\Omega} + \widehat{\Omega}\mathbb{A}$,

$$\widehat{\Pi} = \begin{pmatrix} 0 & -I_3\Omega_3 & I_2\Omega_2 \\ I_3\Omega_3 & 0 & -I_1\Omega_1 \\ -I_2\Omega_2 & I_1\Omega_1 & 0 \end{pmatrix} , \tag{3.34}$$

with

$$I_1 = a_2 + a_3 \,, \quad I_2 = a_1 + a_3 \,, \quad I_3 = a_1 + a_2 \,. \tag{3.35}$$

These quantities are all positive, because $\mathbb{A}$ is positive definite. Consequently, the skew-symmetric matrix $\widehat{\Pi}$ has principle-axis vector components of

$$\boldsymbol{\Pi} = \Pi_1 \mathbf{e}_1 + \Pi_2 \mathbf{e}_2 + \Pi_3 \mathbf{e}_3 \,, \tag{3.36}$$

and the Lagrangian (3.23) in these vector components is expressed as

$$L = \frac{1}{2}\Big(I_1 \Omega_1^2 + I_2 \Omega_2^2 + I_3 \Omega_1^3 \Big) \,, \tag{3.37}$$

with body angular momentum components,

$$\Pi_i = \frac{\delta L}{\delta \Omega_i} = I_i \Omega_i \,, \quad i = 1, 2, 3, \quad \text{(no sum)} \,. \tag{3.38}$$

In this vector representation, the matrix Euler-Lagrange equation (3.24) becomes Euler's equation for rigid body motion.

In vector form, Euler's equations are,

$$\dot{\mathbf{\Pi}} = -\,\mathbf{\Omega}\times\mathbf{\Pi}\,, \tag{3.39}$$

whose vector components are expressed as

$$\begin{aligned} I_1\dot{\Omega}_1 &= (I_2 - I_3)\Omega_2\Omega_3 = -(a_2 - a_3)\Omega_2\Omega_3\,,\\ I_2\dot{\Omega}_2 &= (I_3 - I_1)\Omega_3\Omega_1 = -(a_3 - a_1)\Omega_3\Omega_1\,,\\ I_3\dot{\Omega}_3 &= (I_1 - I_2)\Omega_1\Omega_2 = -(a_1 - a_2)\Omega_1\Omega_2\,. \end{aligned} \tag{3.40}$$

Corollary. Equation (3.39) implies as in Corollary 3.38 that

$$\dot{\boldsymbol{\pi}}(t) = 0\,, \quad \text{for} \quad \boldsymbol{\pi}(t) = O(t)\mathbf{\Pi}(t)\,, \tag{3.41}$$

on using the hat map $O^{-1}\dot{O}(t) = \widehat{\Omega}(t) = \mathbf{\Omega}(t)\times$, as in (3.32).

Remark

3.40 (Two interpretations of Euler's equations).

1. *Euler's equations describe conservation of spatial angular momentum $\boldsymbol{\pi}(t) = O(t)\mathbf{\Pi}(t)$ under the free rotation around a fixed point of a rigid body with principal moments of inertia (I_1, I_2, I_3) in the moving system of coordinates, whose orthonormal basis*

$$(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$$

comprises the principal axes of the body.

2. *Euler's equations also represent* ***geodesic motion*** *on $SO(3)$ with respect to the metric $\mathbb{A}$ whose orthonormal eigenvectors form the basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ and whose (positive) eigenvalues*

$$(a_1, a_2, a_3)$$

are obtained from linear combinations of the formulas for (I_1, I_2, I_3) in equation (3.35). Thus, a rigid body rotates from one orientation to another along the shortest path in $SO(3)$, as determined by using its principal moments of inertia in a metric.

4 Hamilton


Hamilton

Hamilton's approach to geometric optics led to his formulation of the **canonical equations** of particle motion in mechanics.

Geometric optics may be approached either as a theory of systems of rays constructed by means of the elementary laws of refraction (Ibn Sahl, Snell, Descartes, Fermat, Newton), or as a theory based upon the consideration of systems of surfaces whose orthogonal trajectories are the rays (Huygens, Hamilton).

These two approaches embody the dual pictures of light propagation as either rays or as envelopes of Huygens wavelets. The ray approach to geometric optics via Fermat's principle leads to what may be called Lagrangian optics, in which each ray is characterised by assigning an initial point on it and its direction there, much like specifying the initial position and velocity of Newtonian or Lagrangian particle motion. The Huygens wavelet approach leads to Hamiltonian optics, in which a **characteristic function** measures the time that light takes to travel from one point to another and it depends on the co-ordinates of both the initial and final points.

In a *tour de force* begun in 1823, when he was aged eighteen, Hamilton showed that all significant properties of a geometric optical system may be expressed in terms of this characteristic function and its partial derivatives. In this way, Hamilton completed the wave picture of geometric optics first envisioned by Huygens. Hamilton's work was particularly striking because it encompassed and solved the outstanding problem at the time in optics. Namely, it determined how the bright surfaces called "caustics" are created when light reflects off a curved mirror.

Years after his *tour de force* in optics as a young man, Hamilton realised that the same method applies unchanged to mechanics. One simply replaces the optical axis by the time axis, the light rays by the trajectories of the system of particles, and the optical phase space variables by the mechanical phase space variables. Hamilton's formulation of his canonical equations of particle motion in mechanics was expressed using partial derivatives of a *simplified* form of his characteristic function for optics, now called the **Hamiltonian**.

The connection between the Lagrangian and Hamiltonian approaches to mechanics was made via the **Legendre transform**. Hamilton's methods, as developed by Jacobi, Poincaré and other 19th century scientists became a powerful tool in the analysis and solution of problems in mechanics. Hamilton's analogy between optics and mechanics became a guiding light in the development of the quantum mechanics of atoms and molecules a century later, and his ideas still apply today in scientific research on the quantum interactions of photons, electrons and beyond.

4.1 Legendre transform

One passes from Lagrangian to Hamiltonian dynamics through the Legendre transformation.

Definition

4.1 (Legendre transform and fibre derivative).
The ***Legendre transformation*** *is defined by using the* ***fibre derivative*** *of the Lagrangian,*

$$p = \frac{\partial L}{\partial \dot{q}}\,. \tag{4.1}$$

The name ***fibre derivative*** *refers to Definition 3.19 of the tangent bundle TM of a manifold M in which the velocities $\dot{q} \in T_qM$ at a point $q \in M$ are called its* ***fibres****.*

Remark

4.2. *Since the velocity is in the tangent bundle TM, the fibre derivative of the Lagrangian will be in the cotangent bundle T^*M of manifold M.*

Definition

4.3 (Canonical momentum and Hamiltonian).
The quantity p is also called the ***canonical momentum*** *dual to the configuration variable q. If this relation is invertible for the velocity $\dot{q}(q,p)$, then one may define the* ***Hamiltonian****,*

$$H(p,q) = \langle\, p\,,\, \dot{q}\,\rangle - L(q,\dot{q})\,. \tag{4.2}$$

Remark

4.4. *The Hamiltonian $H(p,q)$ may be obtained from the Legendre transformation $H(p,q) = \langle\, p\,,\, \dot q\,\rangle - L(q,\dot q)$ as a function of the variables (q,p), provided one may solve for $\dot q(q,p)$, which requires the Lagrangian to be* ***non-degenerate****, e.g.,*

$$\det \frac{\partial^2 L}{\partial \dot q \partial \dot q} = \det \frac{\partial p(q,\dot q)}{\partial \dot q} \neq 0 \quad \textit{(suppressing indices)}\,. \tag{4.3}$$

Definition

4.5 (Non-degenerate Lagrangian system)**.**
A Lagrangian system (M,L) is said to be ***non-degenerate*** *if the* ***Hessian matrix***

$$H_L(q,\dot q) = \frac{\partial^2 L}{\partial \dot q \partial \dot q} \quad \textit{(again suppressing indices)} \tag{4.4}$$

is invertible everywhere on the tangent bundle TM. Such Lagrangians are also said to be ***hyperregular*** *[MaRa1994].*

Exercise. The following is an example of a singular Lagrangian

$$L(\mathbf{q},\dot{\mathbf{q}}) = n(\mathbf{q})\,\sqrt{\delta_{ij}\dot q^i \dot q^j}\,,$$

that appears in Fermat's principle for ray paths. This Lagrangian is homogeneous of degree 1 in the tangent vector to the ray path. Such a Lagrangian satisfies

$$\frac{\partial^2 L}{\partial \dot q^i \partial \dot q^j}\dot q^j = 0,$$

so its Hessian matrix with respect to the tangent vectors is singular (has zero determinant). This difficulty is inherent in Finsler geometry. Show that this case may be regularised by transforming to a related Riemannian description, in which the Lagrangian is quadratic in the tangent vector. ★

4.2 Hamilton's canonical equations

Theorem

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

$$[\,L\,]_{q^a} = 0\,,$$

in (3.8) are equivalent to ***Hamilton's canonical equations****,*

$$\dot q = \frac{\partial H}{\partial p}\,, \qquad \dot p = -\,\frac{\partial H}{\partial q}\,, \tag{4.5}$$

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

Proof. The derivatives of the Hamiltonian follow from the differential of its defining equation (4.2) as

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

Consequently,

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

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

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

This proves the equivalence of the Euler-Lagrange equations and Hamilton's canonical equations for non-degenerate, or hyper-regular Lagrangians. □

Remark

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

Definition

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

Remark

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

Remark

4.10 (Momentum vs position in phase space).
As discussed in Definition 9.11, the momenta $p = (p_1, \dots, p_n)$ are coordinates in the cotangent bundle at $q = (q^1, \dots, q^n)$ corresponding to the basis $dq^1, \dots, dq^n$ for T^*_qM. This basis for 1-forms in T^*_qM is dual to the vector basis $\partial/\partial q^1, \dots, \partial/\partial q^n$ for the tangent bundle T_qM at $q = (q^1, \dots, q^n)$.

4.3 Phase space action principle

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

$$S = \int_{t_a}^{t_b} L(q, \dot{q}) + p\Big(\frac{dq}{dt} - \dot{q}\Big)\, dt\,, \tag{4.6}$$

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

$$\begin{aligned} \delta S \;&=\; \int_{t_a}^{t_b} \left(\frac{\partial L}{\partial q} - \frac{dp}{dt}\right)\delta q + \left(\frac{\partial L}{\partial \dot{q}} - p\right)\delta\dot{q} + \Big(\frac{dq}{dt} - \dot{q}\Big)\delta p\, dt \\ &\quad + \Big[p\,\delta q\Big]_{t_a}^{t_b}. \end{aligned} \tag{4.7}$$

The contributions at the endpoints t_a and t_b in time vanish, because the variations δq are assumed to vanish then.

Thus, stationarity of this action under these variations imposes the relations

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

- Combining the first and second of these relations recovers the Euler-Lagrange equations, $[\,L\,]_{q^a} = 0$.
- The third relation constrains the variable $\dot{q}$ to be the time derivative of the trajectory $q(t)$ at any time t.

Substituting the Legendre-transform relation (4.2) into the constrained action (4.6) yields the **phase space action**

$$S = \int_{t_a}^{t_b} \Big(p\frac{dq}{dt} - H(q,p) \Big)\, dt\,. \tag{4.8}$$

Varying the phase space action in (4.8) yields

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

Because the variations δq vanish at the endpoints t_a and t_b in time, the last term vanishes. Thus, stationary variations of the phase-space action in (4.8) recover Hamilton's canonical equations (4.5).

Hamiltonian evolution along a curve $(q(t), p(t)) \in T^*M$ satisfying equations (4.5) induces the evolution of a given function $F(q,p) : T^*M \to \mathbb{R}$ on the phase-space T^*M of a manifold M, as

$$\begin{aligned}
\frac{dF}{dt} &= \frac{\partial F}{\partial q}\frac{dq}{dt} + \frac{\partial H}{\partial q}\frac{dp}{dt} \\
&= \frac{\partial F}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial H}{\partial q}\frac{\partial F}{\partial p} =: \{F\,,\,H\} && (4.9) \\
&= \left(\frac{\partial H}{\partial p}\frac{\partial}{\partial q} - \frac{\partial H}{\partial q}\frac{\partial}{\partial p} \right) F =: X_H F\,. && (4.10)
\end{aligned}$$

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

4.4 Poisson brackets

Definition

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

$$\dot{p} = \{p\,,\,H\}\,,\quad \dot{q} = \{q\,,\,H\}\,. \tag{4.11}$$

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

$$\dot{F}(q,p) = \{F\,,\,H\} = \frac{\partial F}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial H}{\partial q}\frac{\partial F}{\partial p}\,. \tag{4.12}$$

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

Remark

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

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

namely,

$$dF\wedge dH = \det\frac{\partial(F,H)}{\partial(q,p)}dq\wedge dp = \{F,H\}dq\wedge dp\,. \tag{4.13}$$

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

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

Proposition

4.13 (The canonical Poisson bracket).
The definition of the ***canonical Poisson bracket*** *in (4.12) implies the following properties. By direct computation, the bracket is:*

1. *bilinear,*
2. *skew symmetric,* $\{F\,,\,H\} = -\,\{H\,,\,F\}$*,*
3. *satisfies the Leibnitz rule (product rule),*

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

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

4. *satisfies the Jacobi identity*

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

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

4.5 Canonical transformations

Definition

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

Example

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

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

in phase space T^*M*, the Poisson bracket in (4.13) transforms via the Jacobian determinant, as*

$$\begin{aligned} dF\wedge dH &= \{F,H\}dq\wedge dp \\ &= \{F,H\}\det\frac{\partial(q,p)}{\partial(Q,P)}dQ\wedge dP\,. \end{aligned} \tag{4.14}$$

Definition

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

$$\det \frac{\partial(q,p)}{\partial(Q,P)} = 1\,, \quad \textit{so that} \quad dq \wedge dp = dQ \wedge dP\,, \tag{4.15}$$

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

$$\dot{P} = \{P\,,\,H\}\,, \quad \dot{Q} = \{Q\,,\,H\}\,. \tag{4.16}$$

Remark

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

$$\det \frac{\partial(q,p)}{\partial(Q,P)} = \text{constant}\,, \tag{4.17}$$

may still be said to be canonical.

Definition

4.18 (Lie transformation groups)**.**

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

Proposition

4.19. *The canonical transformations form a group.*

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

Remark

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

4.6 Flows of Hamiltonian vector fields

The Leibniz property (product rule) in Proposition 4.13 suggests the canonical Poisson bracket is a type of derivative. This derivation property of the Poisson bracket allows its use in the definition of a **Hamiltonian vector field**.

Definition

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

$$X_H = \{\cdot\,,\,H\} = \frac{\partial H}{\partial p}\frac{\partial}{\partial q} - \frac{\partial H}{\partial q}\frac{\partial}{\partial p}\,, \tag{4.18}$$

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

Proposition

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

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

Definition

4.23 (Hamiltonian flow)**.**

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

$$\frac{d\phi_t}{dt} = X_H\big(\phi_t(q,p)\big) = \{\phi_t,\, H\}\,, \tag{4.19}$$

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

Theorem

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

Proof. According to Definition 4.16, a transformation

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

which depends smoothly on a parameter ϵ is canonical, provided it preserves area in phase space (up to a constant factor that defines the units of area). That is, it is canonical provided it satisfies the condition in equation (4.15), namely

$$dq(\epsilon) \wedge dp(\epsilon) = dq(0) \wedge dp(0)\,. \tag{4.20}$$

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

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

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

$$\begin{aligned}
\frac{d}{d\epsilon}\Big(dq(\epsilon) \wedge dp(\epsilon)\Big) &= d(X_F q) \wedge dp + dq \wedge d(X_F p) \\
&= d(F_{,p}) \wedge dp + dq \wedge d(F_{,q}) \\
&= (F_{,pq}dq + F_{,pp}dp) \wedge dp + dq \wedge (F_{,qq}dq + F_{,qp}dp) \\
&= (F_{,pq} - F_{,qp})dq \wedge dp \\
&= 0\,,
\end{aligned} \tag{4.21}$$

by equality of cross derivatives of F and asymmetry of the wedge product. Therefore, condition (4.20) holds and the transformation is canonical. □

Corollary

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

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

Exercise. (Noether's theorem)

Suppose the phase space action (4.8) is invariant under the infinitesimal transformation $q \to q + \delta q$, with $\delta q = \xi_M(q) \in TM$ for $q \in M$ under the transformations of a Lie group G acting on a manifold M. That is, suppose S in (4.8) satisfies $\delta S = 0$ for $\delta q = \xi_M(q) \in TM$.

What does Noether's theorem imply for this phase space action principle? ★

Answer. Noether's theorem implies conservation of the quantity

$$J^\xi = \big\langle p,\, \xi_M(q) \big\rangle_{T^*M \times TM} \in \mathbb{R}\,, \tag{4.22}$$

arising from integration by parts evaluated at the endpoints. This notation introduces a pairing $\langle \,\cdot\,,\,\cdot\, \rangle_{T^*M \times TM} : T^*M \times TM \to \mathbb{R}$. The conservation of J^ξ is expressed as,

$$\frac{dJ^\xi}{dt} = \{ J^\xi,\, H \} = 0\,. \tag{4.23}$$

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

$$\begin{aligned} 0 = X_{J^\xi} H \;&=\; \frac{\partial J^\xi}{\partial p}\frac{\partial H}{\partial q} - \frac{\partial J^\xi}{\partial q}\frac{\partial H}{\partial p} \\ &=\; \xi_M(q)\partial_q H - p\xi'_M(q)\partial_p H \\ &=\; \frac{d}{d\epsilon}\Big|_{\epsilon=0} H\big(q(\epsilon),\, p(\epsilon)\big)\,. \end{aligned} \tag{4.24}$$

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

Conversely, if the Hamiltonian $H(q,p)$ is invariant under the canonical transformation generated by X_{J^ξ}, then the Noether endpoint quantity J^ξ in (4.22) will be a constant of the canonical motion under H.

▲

Definition

4.26 (Cotangent lift momentum map).
On introducing a pairing $\langle\,\cdot\,,\,\cdot\,\rangle : \mathfrak{g}^*\times\mathfrak{g}\to\mathbb{R}$ one may define a map $J:T^*M\to\mathfrak{g}^*$ in terms of this pairing and the Noether endpoint quantity in (4.22 as

$$J^\xi = \big\langle p,\,\xi_M(q)\big\rangle_{T^*M\times TM} =: \big\langle J(q,p),\,\xi\big\rangle_{\mathfrak{g}^*\times\mathfrak{g}}\,, \tag{4.25}$$

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

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

$$\big\{F\circ J\,,\,H\circ J\big\} = \big\{F\,,\,H\big\}\circ J\,. \tag{4.26}$$

This relation defines a **Lie-Poisson bracket** on $\mathfrak{g}^*$ that inherits the properties in Proposition 4.13 of the canonical Poisson bracket.

★

4.7 Properties of Hamiltonian vector fields

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

Definition

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

$$[X_F\,,\,X_H] = X_F X_H - X_H X_F\,, \tag{4.27}$$

which is again a Hamiltonian vector field.

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

Lemma

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

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

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

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

Summation over cyclic permutations then yields the result. □

Lemma

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

$$\{F,\,\{G,\,H\}\} + \{G,\,\{H,\,F\}\} + \{H,\,\{F,\,G\}\} = 0\,. \tag{4.28}$$

Proof. Formula (4.28) may be proved by direct computation, as in Proposition 4.13. This identity may also be verified formally by the same calculation as in the proof of the previous Lemma, by writing $\{G,\,H\} = G(H) - H(G)$ symbolically. □

Remark

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

Theorem

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

$$X_{\{F\,,\,H\}} = -\,[X_F\,,\,X_H]\,. \tag{4.29}$$

Proof. One computes,

$$\begin{aligned}
[X_G, X_H] &= X_G X_H - X_H X_G \\
&= \{G,\,\cdot\,\}\{H,\,\cdot\,\} - \{H,\,\cdot\,\}\{G,\,\cdot\,\} \\
&= \{G,\,\{H,\,\cdot\,\}\} - \{H,\,\{G,\,\cdot\,\}\} \\
&= \{\{G,\,H\},\,\cdot\,\} = -X_{\{G,H\}}\,.
\end{aligned}$$

The first line is the definition of the commutator of vector fields. The second line is the definition of Hamiltonian vector field in terms of Poisson bracket. The third line is a substitution. The fourth line uses the Jacobi identity (4.28) and skew symmetry. □

5 Rigid-body motion

5.1 Hamiltonian form of rigid body motion

A dynamical system on the tangent space TM of a manifold M

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

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

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

in terms of a Poisson bracket operation,

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

which is bilinear, skew-symmetric, defines a derivative operation satisfying the Leibnitz rule for a product of functions and satisfies the Jacobi identity.

As we shall explain, reduced equations arising from group-invariant Hamilton's principles on Lie groups are naturally Hamiltonian. If we **Legendre transform** the Lagrangian in Hamilton's principle in Theorem 3.36 for geodesic motion on $SO(3)$ – interpreted also as rigid body dynamics – then its simple, beautiful and well-known Hamiltonian formulation emerges.

Definition

5.1. *The Legendre transformation from angular velocity* $\mathbf{\Omega}$ *to angular momentum* $\mathbf{\Pi}$ *is defined by*

$$\frac{\delta L}{\delta \mathbf{\Omega}} = \mathbf{\Pi} \,.$$

That is, the Legendre transformation defines the **body angular momentum vector** by the variations of the rigid body's reduced Lagrangian with respect to the body angular velocity vector. For the Lagrangian in (3.37),

$$L(\mathbf{\Omega}) = \frac{1}{2}\,\mathbf{\Omega}\cdot\mathbb{I}\mathbf{\Omega}\,, \tag{5.1}$$

with moment of inertia tensor $\mathbf{I}$, the body angular momentum,

$$\mathbf{\Pi} = \frac{\delta L}{\delta \mathbf{\Omega}} = \mathbb{I}\mathbf{\Omega}\,, \tag{5.2}$$

has $\mathbb{R}^3$ components,

$$\Pi_i = I_i\Omega_i = \frac{\partial L}{\partial \Omega_i}\,, \quad i=1,2,3, \tag{5.3}$$

in which principal moments of inertia I_i with $i=1,2,3$ are all positive definite. This is also how body angular momentum was defined in Definition 2.33 in the Newtonian setting.

5.2 Lie-Poisson Hamiltonian rigid body dynamics

The Legendre transformation is defined for rigid body dynamics by

$$H(\mathbf{\Pi}) := \mathbf{\Pi}\cdot\mathbf{\Omega} - L(\mathbf{\Omega})\,,$$

in terms of the vector dot product on $\mathbb{R}^3$. From the rigid body Lagrangian in (5.1), one finds the expected expression for the rigid body Hamiltonian,

$$H(\mathbf{\Pi}) = \frac{1}{2}\,\mathbf{\Pi}\cdot\mathbb{I}^{-1}\mathbf{\Pi} := \frac{\Pi_1^2}{2I_1} + \frac{\Pi_2^2}{2I_2} + \frac{\Pi_3^2}{2I_3}\,. \tag{5.4}$$

The Legendre transform for this case is invertible for positive definite I_i, so we may solve for

$$\frac{\partial H}{\partial \mathbf{\Pi}} = \mathbf{\Omega} + \left(\mathbf{\Pi} - \frac{\partial L}{\partial \mathbf{\Omega}}\right)\cdot\frac{\partial \mathbf{\Omega}}{\partial \mathbf{\Pi}} = \mathbf{\Omega}\,.$$

In $\mathbb{R}^3$ coordinates, this relation expresses the body angular velocity as the derivative of the reduced Hamiltonian with respect to the body angular momentum, namely,

$$\boldsymbol{\Omega} = \frac{\partial H}{\partial \boldsymbol{\Pi}}\,.$$

Hence, the reduced Euler-Lagrange equation for L may be expressed equivalently in angular momentum vector components in $\mathbb{R}^3$ and Hamiltonian H as:

$$\frac{d}{dt}(\mathbb{I}\boldsymbol{\Omega}) = \mathbb{I}\boldsymbol{\Omega}\times\boldsymbol{\Omega} \Longleftrightarrow \frac{d\boldsymbol{\Pi}}{dt} = \boldsymbol{\Pi}\times\frac{\partial H}{\partial \boldsymbol{\Pi}} := \{\boldsymbol{\Pi}, H\}\,. \tag{5.5}$$

This expression suggests we introduce the following rigid body Poisson bracket on functions of $\boldsymbol{\Pi}\in\mathbb{R}^3$.

$$\{F, H\}(\boldsymbol{\Pi}) := -\,\boldsymbol{\Pi}\cdot\left(\frac{\partial F}{\partial \boldsymbol{\Pi}}\times\frac{\partial H}{\partial \boldsymbol{\Pi}}\right), \tag{5.6}$$

or, in components,

$$\{\Pi_j, \Pi_k\} = -\,\Pi_i\epsilon_{ijk}\,. \tag{5.7}$$

For the Hamiltonian (5.4), one checks that the Euler equations in terms of the rigid body angular momenta,

$$\begin{aligned}
\frac{d\Pi_1}{dt} &= -\left(\frac{1}{I_2}-\frac{1}{I_3}\right)\Pi_2\Pi_3\,,\\
\frac{d\Pi_2}{dt} &= -\left(\frac{1}{I_3}-\frac{1}{I_1}\right)\Pi_3\Pi_1\,,\\
\frac{d\Pi_3}{dt} &= -\left(\frac{1}{I_1}-\frac{1}{I_2}\right)\Pi_1\Pi_2\,,
\end{aligned} \tag{5.8}$$

that is, the equations in vector form,

$$\frac{d\boldsymbol{\Pi}}{dt} = -\,\boldsymbol{\Omega}\times\boldsymbol{\Pi}\,, \tag{5.9}$$

are equivalent to

$$\frac{dF}{dt} = \{F, H\}\,,\quad \text{with}\quad F = \boldsymbol{\Pi}\,.$$

The Poisson bracket proposed in (5.6) may be rewritten in terms of coordinates $\boldsymbol{\Pi}\in\mathbb{R}^3$ as

$$\{F, H\} = -\,\nabla\frac{|\boldsymbol{\Pi}|^2}{2}\cdot\nabla F\times\nabla H\,, \tag{5.10}$$

where ∇ denotes $\partial/\partial\mathbf{\Pi}$. This is an example of the Nambu $\mathbb{R}^3$ bracket [Na1973], which may be seen to satisfy the defining relations to be a Poisson bracket, by identifying it with the commutator of divergenceless vector fields. In this case, the distinguished function $C(\mathbf{\Pi}) = |\mathbf{\Pi}|^2/2$ and its level sets are the angular momentum spheres. Hence, the Hamiltonian rigid body dynamics (5.9) rewritten as

$$\frac{d\mathbf{\Pi}}{dt} = \Big\{\mathbf{\Pi}, H\Big\} = \nabla\frac{|\mathbf{\Pi}|^2}{2} \times \nabla H \tag{5.11}$$

may be interpreted as a divergenceless flow in $\mathbb{R}^3$ along intersections of level sets of angular momentum spheres $|\mathbf{\Pi}|^2 = const$ with the kinetic energy ellipsoids $H = const$ in equation (5.4).

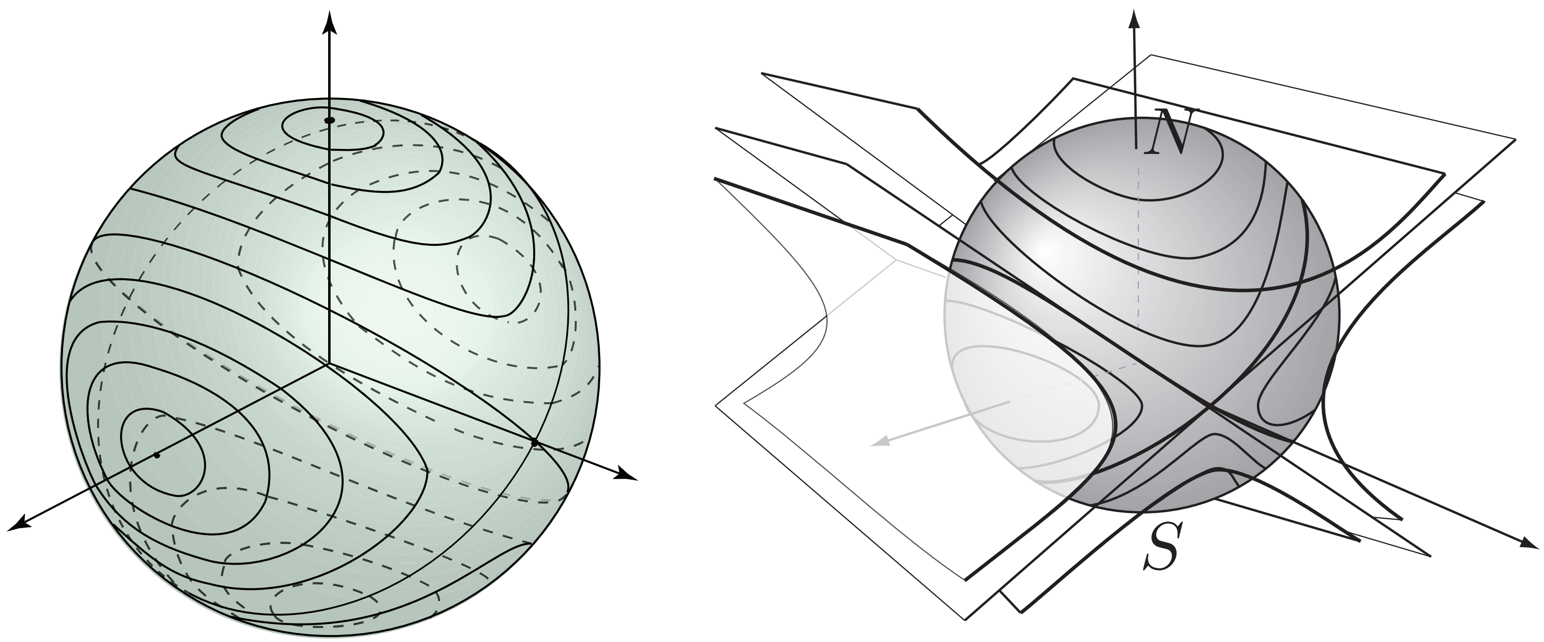


Figure 6: The dynamics of a rotating rigid body may be represented as a divergenceless flow along the intersections in $\mathbb{R}^3$ of the level sets of two conserved quantities: the angular momentum sphere $|\mathbf{\Pi}|^2 = const$ and the hyperbolic cylinders $G = const$ in equation (5.18).

5.3 Geometry of rigid body level sets in $\mathbb{R}^3$

Euler's equations (5.11) are expressible in vector form as

$$\frac{d}{dt}\boldsymbol{\Pi} = \nabla L \times \nabla H\,, \tag{5.12}$$

where H is the rotational kinetic energy

$$H = \frac{\Pi_1^2}{2I_1} + \frac{\Pi_2^2}{2I_2} + \frac{\Pi_3^2}{2I_3}\,, \tag{5.13}$$

with gradient

$$\nabla H = \left(\frac{\partial H}{\partial \Pi_1}, \frac{\partial H}{\partial \Pi_2}, \frac{\partial H}{\partial \Pi_3}\right) = \left(\frac{\Pi_1}{I_1}, \frac{\Pi_2}{I_2}, \frac{\Pi_3}{I_3}\right)\,,$$

and L is half the square of the body angular momentum

$$L = \frac{1}{2}(\Pi_1^2 + \Pi_2^2 + \Pi_3^2)\,, \tag{5.14}$$

with gradient

$$\nabla L = (\Pi_1,\, \Pi_2,\, \Pi_3)\,. \tag{5.15}$$

Since both H and L are conserved, the rigid body motion itself takes place, as we know, along the intersections of the level surfaces of the energy (ellipsoids) and the angular momentum (spheres) in $\mathbb{R}^3$: The centres of the energy ellipsoids and the angular momentum spheres coincide. This, along with the $(\mathbb{Z}_2)^3$ symmetry of the energy ellipsoids, implies that the two sets of level surfaces in $\mathbb{R}^3$ develop collinear gradients (for example, tangencies) at pairs of points which are diametrically opposite on an angular momentum sphere. At these points, collinearity of the gradients of H and L implies stationary rotations, that is, equilibria.

The geometry of the level sets on whose intersections the motion takes place may be recast equivalently by taking linear combinations of H and L. For example, consider the following.

Proposition

5.2. *Euler's equations for the rigid body (5.12) may be written equivalently as*

$$\frac{d}{dt}\boldsymbol{\Pi} = \nabla L \times \nabla G\,, \quad \textit{where} \quad G = H - \frac{L}{I_2}\,, \tag{5.16}$$

or, explicitly, L and G are given by

$$L = \frac{1}{2}(\Pi_1^2 + \Pi_2^2 + \Pi_3^2) \tag{5.17}$$

and

$$G = \Pi_1^2\left(\frac{1}{2I_1} - \frac{1}{2I_2}\right) - \Pi_3^2\left(\frac{1}{2I_2} - \frac{1}{2I_3}\right). \tag{5.18}$$

Proof. The proof is immediate. Since $\nabla L \times \nabla L = 0$,

$$\frac{d}{dt}\mathbf{\Pi} = \nabla L \times \nabla H = \nabla L \times \nabla\left(H - \frac{L}{I_2}\right) = \nabla L \times \nabla G\,. \tag{5.19}$$

□

Remark

5.3. *With the linear combination $G = H - L/I_2$, the solutions of Euler's equations for rigid body dynamics may be realised as flow along the intersections of the spherical level sets of the body angular momentum $L = const$ and a family of* hyperbolic cylinders *$G = const$. These hyperbolic cylinders are translation-invariant along the principal axis of the intermediate moment of inertia and oriented so that the asymptotes of the hyperbolas (at $G = 0$) slice each angular momentum sphere along the four (heteroclinic) orbits that connect the diametrically opposite points on the sphere that lie along the intermediate axis. See Figure 6.*

5.4 Rotor and pendulum

The idea of recasting the geometry of flow lines in $\mathbb{R}^3$ as the intersections of different level sets on which the motion takes place was extended in [?] to reveal a remarkable relationship between the rigid body and the planar pendulum. This relationship was found by further exploiting the symmetry of the triple scalar product appearing in the $\mathbb{R}^3$ bracket (5.10).

Theorem

5.4. *Euler's equations for the rigid body (5.12) may be written equivalently as*

$$\frac{d}{dt}\mathbf{\Pi} = \nabla A \times \nabla B\,, \tag{5.20}$$

where A and B are given by the following linear combinations of H and L,

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{bmatrix} a & b \\ c & e \end{bmatrix}\begin{pmatrix} H \\ L \end{pmatrix}, \tag{5.21}$$

in which the constants a, b, c, e satisfy $ae - bc = 1$ to form an $SL(2,\mathbb{R})$ matrix.

Proof. Recall from equation (5.6) that

$$\begin{aligned}\{F,H\}d^3\Pi &:= dF\wedge dL\wedge dH \\ &= \frac{1}{ae-bc}\,dF\wedge d(aH+bL)\wedge d(cH+eL)\,,\end{aligned}\tag{5.22}$$

for real constants a,b,c,e. Consequently, the rigid body equation will remain invariant under any linear combinations of energy and angular momentum

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{bmatrix} a & b \\ c & e \end{bmatrix}\begin{pmatrix} H \\ L \end{pmatrix},$$

provided the constants a,b,c,e satisfy $ae-bc=1$ to form an SL(2,$\mathbb{R}$) matrix. □

Remark

5.5 (Equilibria)**.**
For a general choice for the linear combination of A and B in (5.21), equilibria occur at points where the cross product of gradients $\nabla A\times\nabla B$ vanishes. This can occur at points where the level sets are tangent (and the gradients are both nonzero), or at points where one of the gradients vanishes.

Corollary

5.6. *Euler's equations for the rigid body (5.12) may be written equivalently as*

$$\frac{d}{dt}\boldsymbol{\Pi} = \nabla N\times\nabla K\,,\tag{5.23}$$

where K and N are

$$K=\frac{\Pi_1^2}{2k_1^2}+\frac{\Pi_2^2}{2k_2^2}\qquad \textit{and}\qquad N=\frac{\Pi_2^2}{2k_3^2}+\frac{1}{2}\Pi_3^2\,,\tag{5.24}$$

for

$$\frac{1}{k_1^2}=\frac{1}{I_1}-\frac{1}{I_3},\qquad \frac{1}{k_2^2}=\frac{1}{I_2}-\frac{1}{I_3},\qquad \frac{1}{k_3^2}=\frac{I_3(I_2-I_1)}{I_2(I_3-I_1)}\,.\tag{5.25}$$

Proof. If $I_1<I_2<I_3$, the choice

$$c=1,\qquad e=\frac{1}{I_3},\qquad a=\frac{I_1I_3}{I_3-I_1}<0,\qquad \text{and}\qquad b=\frac{I_3}{I_3-I_1}<0$$

yields

$$\{F,H\}d^3\Pi := dF \wedge dL \wedge dH = dF \wedge dN \wedge dK\,, \tag{5.26}$$

from which equations (5.23) - (5.25) of the Corollary follow.

Since

$$\begin{pmatrix} H \\ L \end{pmatrix} = \frac{1}{ae-bc}\begin{bmatrix} e & -b \\ -c & a \end{bmatrix}\begin{pmatrix} N \\ K \end{pmatrix}$$

we also have

$$H = eN - bK = \frac{1}{I_3}N + \frac{I_3}{I_3 - I_1}K\,. \tag{5.27}$$

Consequently, we may write

$$dF \wedge dL \wedge dH = dF \wedge dN \wedge dK = -I_3\, dF \wedge dK \wedge dH\,. \tag{5.28}$$

□

With this choice, the orbits for Euler's equations for rigid body dynamics are realised as motion along the intersections of two, orthogonally oriented, *elliptic cylinders*, one elliptic cylinder is a level surface of K, with its translation axis along Π_3 (where $K = 0$), and the other is a level surface of N, with its translation axis along Π_1 (where $N = 0$).

Equilibria occur at points where the cross product of gradients $\nabla K \times \nabla N$ vanishes. In the elliptic cylinder case above, this may occur at points where the elliptic cylinders are tangent, and at points where the axis of one cylinder punctures normally through the surface of the other. The elliptic cylinders are tangent at one $\mathbb{Z}_2$-symmetric pair of points along the Π_2 axis, and the elliptic cylinders have normal axial punctures at two other $\mathbb{Z}_2$-symmetric pairs of points along the Π_1 and Π_3 axes.

5.4.1 Restricting rigid body motion to elliptic cylinders

We pursue the geometry of the elliptic cylinders by restricting the rigid body equations to a level surface of K. On the surface $K = constant$, define new variables θ and p by

$$\Pi_1 = k_1 r\cos\theta, \qquad \Pi_2 = k_2 r\sin\theta, \qquad \Pi_3 = p\,, \quad \text{with} \quad r = \sqrt{2K}\,,$$

so that

$$d^3\Pi := d\Pi_1 \wedge d\Pi_2 \wedge d\Pi_3 = k_1k_2\, dK \wedge d\theta \wedge dp\,.$$

In terms of these variables, the constants of the motion become

$$K = \frac{1}{2}r^2 \qquad \text{and} \qquad N = \frac{1}{2}p^2 + \left(\frac{k_2^2}{2k_3^2}r^2\right)\sin^2\theta.$$

On a constant level surface of K the function $\{F,H\}$ only depends on (θ,p) so the Poisson bracket for rigid body motion on any *particular* elliptic cylinder is given by (5.26) as

$$\begin{aligned}\{F,\,H\}\,d^3\Pi &= -\,dL\wedge dF\wedge dH\\ &= k_1k_2\,dK\wedge\{F,\,H\}_{\mathsf{EllipCyl}}\,d\theta\wedge dp\,.\end{aligned}\tag{5.29}$$

The symplectic structure on the level set $K = constant$ is thus given by the following Poisson bracket on this elliptic cylinder:

$$\{F,H\}_{\mathsf{EllipCyl}} = \frac{1}{k_1k_2}\left(\frac{\partial F}{\partial p}\frac{\partial H}{\partial\theta} - \frac{\partial F}{\partial\theta}\frac{\partial H}{\partial p}\right),$$

which is *symplectic*. In particular, it satisfies

$$\{p,\theta\}_{\mathsf{EllipCyl}} = \frac{1}{k_1k_2}\,.\tag{5.30}$$

The restriction of the Hamiltonian H to the symplectic level set of the elliptic cylinder $K = constant$ is by (5.13)

$$H = \frac{k_1^2K}{I_1} + \frac{1}{I_3}\left[\frac{1}{2}p^2 + \frac{I_3^2(I_2-I_1)}{2(I_3-I_2)(I_3-I_1)}r^2\sin^2\theta\right] = \frac{k_1^2K}{I_1} + \frac{N}{I_3}\,.$$

That is, N/I_3 can be taken as the Hamiltonian on this symplectic level set of K. Note that N/I_3 has the form of kinetic plus potential energy. The equations of motion are thus given by

$$\begin{aligned}\frac{d\theta}{dt} &= \left\{\theta,\frac{N}{I_3}\right\}_{\mathsf{EllipCyl}} = \frac{1}{k_1k_2I_3}\frac{\partial N}{\partial p} = -\,\frac{1}{k_1k_2I_3}p\,,\\ \frac{dp}{dt} &= \left\{p,\frac{N}{I_3}\right\}_{\mathsf{EllipCyl}} = \frac{1}{k_1k_2I_3}\frac{\partial N}{\partial\theta} = \frac{1}{k_1k_2I_3}\frac{k_2^2}{k_3^2}r^2\sin\theta\cos\theta\,.\end{aligned}$$

Combining these equations of motion gives the **pendulum equation**,

$$\frac{d^2\theta}{dt^2} = -\,\frac{r^2}{k_1k_2I_3}\sin 2\theta\,.$$

In terms of the original rigid body parameters, this becomes

$$\frac{d^2\theta}{dt^2} = -\,\frac{K}{I_3^2}\left(\frac{1}{I_1} - \frac{1}{I_2}\right)\sin 2\theta\,.\tag{5.31}$$

Thus, simply by transforming coordinates, we have proved the following result.

Proposition

5.7. *Rigid body motion reduces to pendulum motion on level surfaces of K.*

Corollary

5.8. *The dynamics of a rigid body in three-dimensional body angular momentum space is a union of two-dimensional simple-pendulum phase portraits, as shown in Figure 7.*

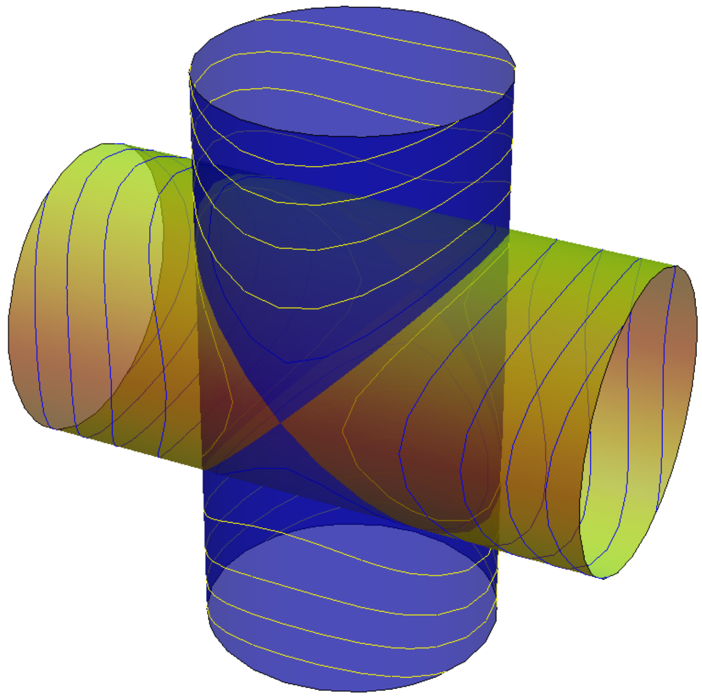


Figure 7: The dynamics of the rigid body in three-dimensional body angular momentum space is recovered by taking the union in $\mathbb{R}^3$ of the intersections of level surfaces of two orthogonal families of concentric cylinders. (Only one member of each family is shown in the figure here, although the curves on each cylinder show other intersections.) On each cylindrical level surface, the dynamics reduces to that of a simple pendulum, as given in equation (5.31).

By restricting to a nonzero level surface of K, the pair of rigid body equilibria along the Π_3 axis are excluded. (This pair of equilibria can be included by permuting the indices of the moments of inertia.) The other two pairs of equilibria, along the Π_1 and Π_2 axes, lie in the

$p = 0$ plane at $\theta = 0$; $\pi/2$, π and $3\pi/2$. Since K is positive, the stability of each equilibrium point is determined by the relative sizes of the principal moments of inertia, which affect the overall sign of the right-hand side of the pendulum equation. The well-known results about stability of equilibrium rotations along the least and greatest principal axes, and instability around the intermediate axis, are immediately recovered from this overall sign, combined with the stability properties of the pendulum equilibria.

For $K > 0$ and $I_1 < I_2 < I_3$; this overall sign is negative, so the equilibria at $\theta = 0$ and π (along the Π_1 axis) are stable, while those at $\theta = \pi/2$ and $3\pi/2$ (along the Π_2 axis) are unstable. The factor of 2 in the argument of the sine in the pendulum equation is explained by the $\mathbb{Z}_2$ symmetry of the level surfaces of K (or, just as well, by their invariance under $\theta \to \theta + \pi$). Under this discrete symmetry operation, the equilibria at $\theta = 0$ and $\pi/2$ exchange with their counterparts at $\theta = \pi$ and $3\pi/2$; respectively, while the elliptical level surface of K is left invariant. By construction, the Hamiltonian N/I_3 in the reduced variables θ and p is also invariant under this discrete symmetry.

6 Spherical pendulum

A spherical pendulum of unit length swings from a fixed point of support under the constant acceleration of gravity g. This motion is equivalent to a particle of unit mass moving on the surface of the unit sphere S^2 under the influence of the gravitational (linear) potential $V(z)$ with $z = \hat{\mathbf{e}}_3 \cdot \mathbf{x}$. The only forces acting on the mass are the reaction from the sphere and gravity. This system may be treated as an enhanced coursework example by using spherical polar coordinates and the traditional methods of Newton, Lagrange and Hamilton. The present section treats this problem more geometrically, inspired by the approach discussed in [CuBa1997, EfMoSa2005].

In this section, the equations of motion for the spherical pendulum will be derived according to the approaches of Lagrange and Hamilton on the tangent bundle TS^2 of $S^2 \in \mathbb{R}^3$:

$$TS^2 = \big\{ (\mathbf{x}, \dot{\mathbf{x}}) \in T\mathbb{R}^3 \simeq \mathbb{R}^6 \;\big|\; 1 - |\mathbf{x}|^2 = 0,\, \mathbf{x} \cdot \dot{\mathbf{x}} = 0 \big\}. \tag{6.1}$$

After the Legendre transformation to the Hamiltonian side, the canonical equations will be transformed to quadratic variables that are invariant under S^1 rotations about the vertical axis. This is the **quotient map** for the spherical pendulum.

Then the **Nambu bracket** in $\mathbb{R}^3$ will be found in these S^1 quadratic invariant variables and the equations will be reduced to the **orbit manifold**, which is the zero level set of a distinguished function called the **Casimir function** for this bracket. On the intersections of the Hamiltonian with the orbit manifold, the reduced equations for the spherical pendulum will simplify to the equations of a quadratically nonlinear oscillator.

The solution for the motion of the spherical pendulum will be finished by finding expressions for its geometrical and dynamical phases.

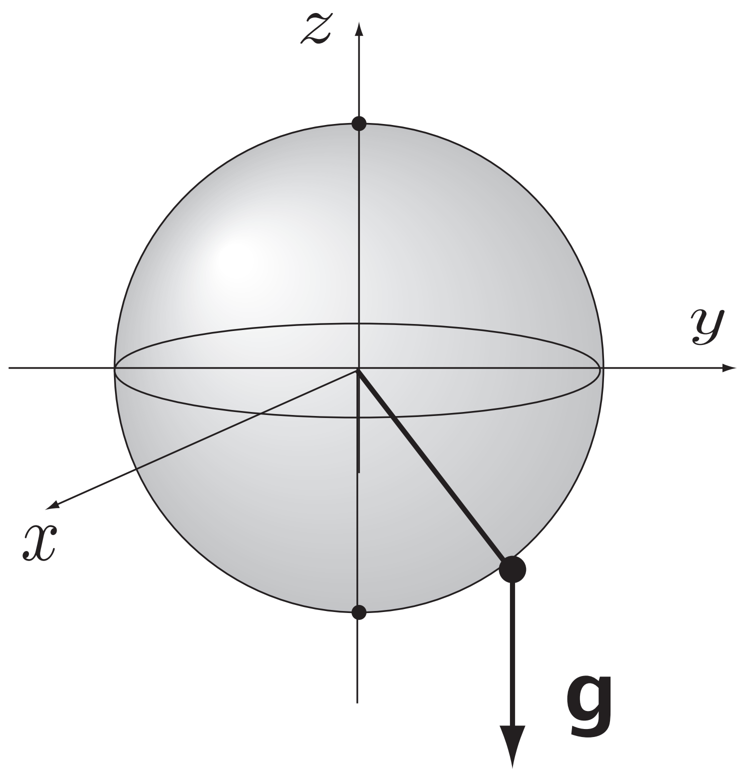


Figure 8: Spherical pendulum moving under gravity on TS^2 in $\mathbb{R}^3$.

The constrained Lagrangian We begin with the Lagrangian $L(\mathbf{x}, \dot{\mathbf{x}}) : T\mathbb{R}^3 \to \mathbb{R}$ given by

$$L(\mathbf{x}, \dot{\mathbf{x}}) = \tfrac{1}{2}|\dot{\mathbf{x}}|^2 - g\hat{\mathbf{e}}_3 \cdot \mathbf{x} - \tfrac{1}{2}\mu(1 - |\mathbf{x}|^2), \tag{6.2}$$

in which the Lagrange multiplier μ constrains the motion to remain on the sphere S^2 by enforcing $(1 - |\mathbf{x}|^2) = 0$ when it is varied in Hamilton's principle. The corresponding Euler-Lagrange equation is

$$\ddot{\mathbf{x}} = -g\hat{\mathbf{e}}_\mathbf{3} + \mu\mathbf{x}\,. \tag{6.3}$$

This equation preserves both of the TS^2 relations $1 - |\mathbf{x}|^2 = 0$ and $\mathbf{x} \cdot \dot{\mathbf{x}} = 0$, provided the Lagrange multiplier is given by

$$\mu = g\hat{\mathbf{e}}_3 \cdot \mathbf{x} - |\dot{\mathbf{x}}|^2\,. \tag{6.4}$$

Remark

6.1. *In Newtonian mechanics, the motion equation obtained by substituting (6.4) into (6.3) may be interpreted as*

$$\ddot{\mathbf{x}} = \mathbf{F} \cdot (\mathrm{Id} - \mathbf{x} \otimes \mathbf{x}) - |\dot{\mathbf{x}}|^2 \mathbf{x} \,,$$

where $\mathbf{F} = -g\hat{\mathbf{e}}_3$ *is the force exerted by gravity on the particle,*

$$\mathbf{T} = \mathbf{F} \cdot (\mathrm{Id} - \mathbf{x} \otimes \mathbf{x})$$

is its component tangential to the sphere and, finally, $-|\dot{\mathbf{x}}|^2 \mathbf{x}$ *is the centripetal force for the motion to remain on the sphere.*

S^1 symmetry and Noether's theorem The Lagrangian in (6.2) is invariant under S^1 rotations about the vertical axis, whose infinitesimal generator is $\delta\mathbf{x} = \hat{\mathbf{e}}_3 \times \mathbf{x}$. Consequently Noether's theorem (Corollary 3.34) that each smooth symmetry of the Lagrangian in an action principle implies a conservation law for its Euler-Lagrange equations, in this case implies that the equations (6.3) conserve

$$J_3(\mathbf{x}, \dot{\mathbf{x}}) = \dot{\mathbf{x}} \cdot \delta\mathbf{x} = \mathbf{x} \times \dot{\mathbf{x}} \cdot \hat{\mathbf{e}}_3 \,, \tag{6.5}$$

which is the angular momentum about the vertical axis.

Legendre transform and canonical equations The fibre derivative of the Lagrangian L in (6.2) is

$$\mathbf{y} = \frac{\partial L}{\partial \dot{\mathbf{x}}} = \dot{\mathbf{x}} \,. \tag{6.6}$$

The variable $\mathbf{y}$ will be the momentum canonically conjugate to the radial position $\mathbf{x}$, after the Legendre transform to the corresponding Hamiltonian,

$$H(\mathbf{x}, \mathbf{y}) = \tfrac{1}{2}|\mathbf{y}|^2 + g\hat{\mathbf{e}}_3 \cdot \mathbf{x} + \tfrac{1}{2}(g\hat{\mathbf{e}}_3 \cdot \mathbf{x} - |\mathbf{y}|^2)(1 - |\mathbf{x}|^2) \,, \tag{6.7}$$

whose canonical equations on $(1 - |\mathbf{x}|^2) = 0$, are

$$\dot{\mathbf{x}} = \mathbf{y} \quad \text{and} \quad \dot{\mathbf{y}} = -g\hat{\mathbf{e}}_3 + (g\hat{\mathbf{e}}_3 \cdot \mathbf{x} - |\mathbf{y}|^2)\mathbf{x} \,. \tag{6.8}$$

This Hamiltonian system on $T^*\mathbb{R}^3$ admits TS^2 as an invariant manifold, provided the initial conditions satisfy the defining relations for TS^2 in (6.1). On TS^2, equations (6.8) conserve the energy

$$E(\mathbf{x}, \mathbf{y}) = \tfrac{1}{2}|\mathbf{y}|^2 + g\hat{\mathbf{e}}_3 \cdot \mathbf{x} \,, \tag{6.9}$$

and the vertical angular momentum

$$J_3(\mathbf{x},\mathbf{y}) = \mathbf{x}\times\mathbf{y}\cdot\hat{\mathbf{e}}_3\,.$$

Under the $(\mathbf{x},\mathbf{y})$ canonical Poisson bracket, the angular momentum component J_3 generates the Hamiltonian vector field

$$\begin{aligned} X_{J_3} = \{\,\cdot\,,\,J_3\} &= \frac{\partial J_3}{\partial\mathbf{y}}\cdot\frac{\partial}{\partial\mathbf{x}} - \frac{\partial J_3}{\partial\mathbf{x}}\cdot\frac{\partial}{\partial\mathbf{y}} \\ &= \hat{\mathbf{e}}_3\times\mathbf{x}\cdot\frac{\partial}{\partial\mathbf{x}} + \hat{\mathbf{e}}_3\times\mathbf{y}\cdot\frac{\partial}{\partial\mathbf{y}}\,, \end{aligned} \tag{6.10}$$

for infinitesimal rotations about the vertical axis $\hat{\mathbf{e}}_3$. Because of the S^1 symmetry of the Hamiltonian in (6.7) under these rotations, we have the conservation law,

$$\dot{J}_3 = \{J_3, H\} = X_{J_3}H = 0\,.$$

6.1 Lie symmetry reduction

Algebra of invariants To take advantage of the S^1 symmetry of the spherical pendulum, we transform to S^1-invariant quantities. A convenient choice of basis for the algebra of polynomials in $(\mathbf{x},\mathbf{y})$ that are S^1-invariant under rotations about the 3-axis is given by [EfMoSa2005]

$$\begin{array}{llllll} \sigma_1 &= x_3 & \sigma_3 &= y_1^2+y_2^2+y_3^2 & \sigma_5 &= x_1y_1+x_2y_2 \\ \sigma_2 &= y_3 & \sigma_4 &= x_1^2+x_2^2, & \sigma_6 &= x_1y_2-x_2y_1 \end{array}$$

Quotient map The transformation defined by

$$\pi:\ (\mathbf{x},\mathbf{y})\to\{\sigma_j(\mathbf{x},\mathbf{y}),\ j=1,\dots,6\} \tag{6.11}$$

is the **quotient map** $T\mathbb{R}^3\to\mathbb{R}^6$ for the spherical pendulum. Each of the fibres of the quotient map π is an S^1 orbit generated by the Hamiltonian vector field X_{J_3} in (6.10).

The six S^1-invariants that define the quotient map in (6.11) for the spherical pendulum satisfy the cubic algebraic relation

$$\sigma_5^2+\sigma_6^2 = \sigma_4(\sigma_3-\sigma_2^2)\,. \tag{6.12}$$

They also satisfy the positivity conditions

$$\sigma_4\ge 0,\quad \sigma_3\ge\sigma_2^2. \tag{6.13}$$

In these variables, the defining relations (6.1) for TS^2 become

$$\sigma_4+\sigma_1^2 = 1 \quad\text{and}\quad \sigma_5+\sigma_1\sigma_2 = 0\,. \tag{6.14}$$

Perhaps not unexpectedly, since TS^2 is invariant under the S^1 rotations, it is also expressible in terms of S^1-invariants. The three relations in equations (6.12) – (6.14) will define the **orbit manifold** for the spherical pendulum in $\mathbb{R}^6$.

Reduced space and orbit manifold in $\mathbb{R}^3$ On TS^2, the variables $\sigma_j(\mathbf{x},\mathbf{y})$, $j=1,\dots,6$, satisfying (6.14) allow the elimination of σ_4 and σ_5 to satisfy the algebraic relation

$$\sigma_1^2\sigma_2^2+\sigma_6^2=(\sigma_3-\sigma_2^2)(1-\sigma_1^2)\,,$$

which on expansion simplifies to

$$\sigma_2^2+\sigma_6^2=\sigma_3(1-\sigma_1^2)\,, \tag{6.15}$$

where $\sigma_3\ge 0$ and $(1-\sigma_1^2)\ge 0$. Restoring $\sigma_6=J_3$, we may write the previous equation as

$$C(\sigma_1,\sigma_2,\sigma_3;J_3^2)=\sigma_3(1-\sigma_1^2)-\sigma_2^2-J_3^2=0\,. \tag{6.16}$$

This is the **orbit manifold** for the spherical pendulum in $\mathbb{R}^3$. The motion takes place on the following family of surfaces depending on $(\sigma_1,\sigma_2,\sigma_3)\in\mathbb{R}^3$ and parameterised by the conserved value of J_3^2,

$$\sigma_3=\frac{\sigma_2^2+J_3^2}{1-\sigma_1^2}\,. \tag{6.17}$$

The orbit manifold for the spherical pendulum is a graph of σ_3 over $(\sigma_1,\sigma_2)\in\mathbb{R}^2$, provided $1-\sigma_1^2\neq 0$. The two solutions of $1-\sigma_1^2=0$ correspond to the North and South poles of the sphere. In the case $J_3^2=0$, the spherical pendulum restricts to the *planar* pendulum.

Reduced Poisson bracket in $\mathbb{R}^3$ When evaluated on TS^2, the Hamiltonian for the spherical pendulum is expressed in these S^1-invariant variables by the linear relation

$$H=\tfrac{1}{2}\sigma_3+g\sigma_1\,, \tag{6.18}$$

whose level surfaces are planes in $\mathbb{R}^3$. The motion in $\mathbb{R}^3$ takes place on the intersections of these Hamiltonian planes with the level sets of J_3^2 given by $C=0$ in equation (6.16). Consequently, in $\mathbb{R}^3$-vector form, the motion is governed by the cross-product formula

$$\dot{\boldsymbol{\sigma}}=\frac{\partial C}{\partial\boldsymbol{\sigma}}\times\frac{\partial H}{\partial\boldsymbol{\sigma}}\,. \tag{6.19}$$

In components, this evolution is expressed as

$$\dot\sigma_i=\{\sigma_i,H\}=\epsilon_{ijk}\frac{\partial C}{\partial\sigma_j}\frac{\partial H}{\partial\sigma_k}\quad\text{with}\quad i,j,k=1,2,3. \tag{6.20}$$

The motion may be expressed in Hamiltonian form by introducing the following bracket operation, defined for a function F of the S^1-invariant vector $\boldsymbol{\sigma}=(\sigma_1,\sigma_2,\sigma_3)\in\mathbb{R}^3$ by

$$\{F,H\}=-\,\frac{\partial C}{\partial\boldsymbol{\sigma}}\cdot\frac{\partial F}{\partial\boldsymbol{\sigma}}\times\frac{\partial H}{\partial\boldsymbol{\sigma}}=-\,\epsilon_{ijk}\frac{\partial C}{\partial\sigma_i}\frac{\partial F}{\partial\sigma_j}\frac{\partial H}{\partial\sigma_k}\,. \tag{6.21}$$

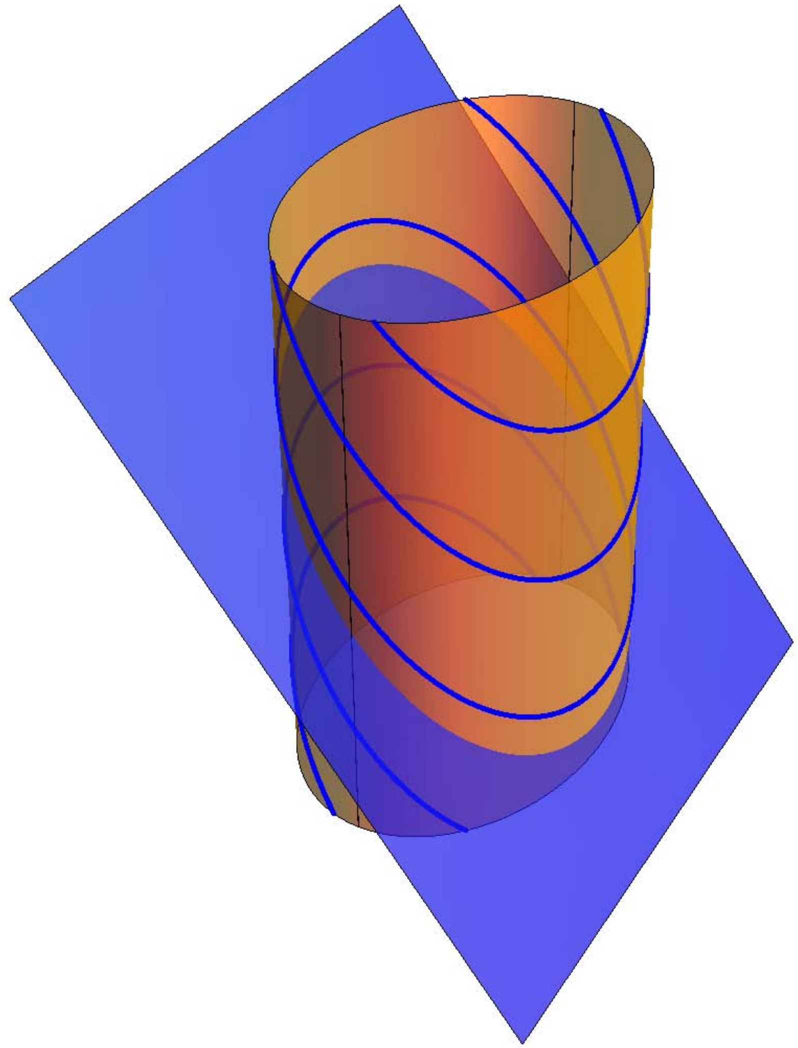


Figure 9: The dynamics of the spherical pendulum in the space of S^1 invariants $(\sigma_1, \sigma_2, \sigma_3)$ is recovered by taking the union in $\mathbb{R}^3$ of the intersections of level sets of two families of surfaces. These surfaces are the roughly cylindrical level sets of angular momentum about the vertical axis given in (6.17) and the (planar) level sets of the Hamiltonian in (6.18). (Only one member of each family is shown in the figure here, although the curves show a few of the other intersections.) On each *planar* level set of the Hamiltonian, the dynamics reduces to that of a quadratically nonlinear oscillator for the verical coordinate (σ_1) given in equation (6.24).

The bracket in (6.21) is another example of the Nambu $\mathbb{R}^3$ bracket introduced in [Na1973], which satisfies the defining relations to be a Poisson bracket. In this case, the distinguished function $C(\sigma_1, \sigma_2, \sigma_3; J_3^2)$ in (6.16) defines a level set of the squared vertical angular momentum J_3^2 in $\mathbb{R}^3$ given by $C = 0$. The distinguished function C is a **Casimir function** for the Nambu bracket in $\mathbb{R}^3$. That is, the Nambu bracket in (6.21) with C obeys $\{C, H\} = 0$ for *any* Hamiltonian $H(\sigma_1, \sigma_2, \sigma_3) : \mathbb{R}^3 \to \mathbb{R}$. Consequently, the motion governed by this $\mathbb{R}^3$ bracket takes place on level sets of J_3^2 given by $C = 0$.

Poisson map Introducing the Nambu bracket in (6.21) ensures that the orbit map for the spherical pendulum $\pi : T\mathbb{R}^3 \to \mathbb{R}^6$ is a **Poisson map**. That is, the subspace obtained by using the relations (6.14) to restrict to the invariant manifold TS^2 produces a set of Poisson brackets $\{\sigma_i, \sigma_j\}$ for $i, j = 1, 2, 3,$ that close amongst themselves. Namely,

$$\{\sigma_i, \sigma_j\} = \epsilon_{ijk} \frac{\partial C}{\partial \sigma_k}\,, \tag{6.22}$$

with C given in (6.16). These brackets may be expressed in tabular form, as

$\{\,\cdot\,,\,\cdot\,\}$	σ_1	σ_2	σ_3
σ_1	0	$1 - \sigma_1^2$	$2\sigma_2$
σ_2	$-1 + \sigma_1^2$	0	$-\,2\sigma_1\sigma_3$
σ_3	$-\,2\sigma_2$	$2\sigma_1\sigma_3$	0

In addition, $\{\sigma_i, \sigma_6\} = 0$ for $i = 1, 2, 3,$ since $\sigma_6 = J_3$ and the $\{\sigma_i \big| \, i = 1, 2, 3\}$ are all S^1-invariant under X_{J_3} in (6.10).

Exercise. Prove that the Nambu bracket in (6.21) satisfies the defining properties in Proposition 4.13 that are required for it to be a genuine Poisson bracket. ★

Reduced motion: Restriction in $\mathbb{R}^3$ to Hamiltonian planes The individual components of the equations of motion may be obtained from (6.20) as

$$\dot{\sigma}_1 = -\sigma_2\,, \qquad \dot{\sigma}_2 = \sigma_1\sigma_3 + g(1 - \sigma_1^2)\,, \qquad \dot{\sigma}_3 = 2g\sigma_2\,. \tag{6.23}$$

Substituting $\sigma_3 = 2(H - g\sigma_1)$ from equation (6.18) and setting the acceleration of gravity to be unity $g = 1$ yields

$$\ddot{\sigma}_1 = 3\sigma_1^2 - 2H\sigma_1 - 1 \tag{6.24}$$

which has equilibria at $\sigma_1^{\pm} = \frac{1}{3}(H \pm \sqrt{H^2+3})$ and conserves the energy integral

$$\tfrac{1}{2}\dot{\sigma}_1^2 + V(\sigma_1) = E \tag{6.25}$$

with the potential $V(\sigma_1)$ parameterised by H in (6.18) and given by

$$V(\sigma_1) = -\,\sigma_1^3 + H\sigma_1^2 + \sigma_1 \tag{6.26}$$

Equation (6.25) is an energy equation for a particle of unit mass, with position σ_1 and energy E, moving in a cubic potential field $V(\sigma_1)$. For $H = 0$, its equilibria in the $(\sigma_1, \dot{\sigma}_1)$ phase plane are at $(\sigma_1, \dot{\sigma}_1) = (\pm\sqrt{3}/3, 0)$, as sketched in Figure 10.

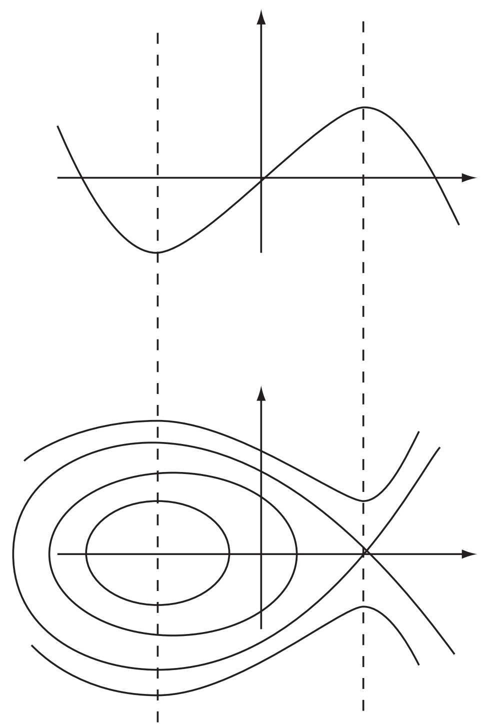


Figure 10: The upper panel shows a sketch of the cubic potential $V(\sigma_1)$ in equation (6.26) for the case $H = 0$. For $H = 0$, the potential has three zeros located at $\sigma_1 = 0, \pm 1$ and two critical points (relative equilibria) at $\sigma_1 = -\sqrt{3}/3$ (centre) and $\sigma_1 = +\sqrt{3}/3$ (saddle). The lower panel shows a sketch of its fish-shaped saddle-centre configuration in the $(\sigma_1, \dot{\sigma}_1)$ phase plane, comprising several level sets of $E(\sigma_1, \dot{\sigma}_1)$ from equation (6.25) for $H = 0$.

Each curve in the lower panel of Figure 10 represents the intersection in the reduced phase-space with S^1-invariant coordinates $(\sigma_1, \sigma_2, \sigma_3) \in \mathbb{R}^3$ of one of the Hamiltonian planes (6.18) with a level set of J_3^2 given by $C = 0$ in equation (6.16). The critical points of the potential are relative equilibria, corresponding to S^1-periodic solutions. The case $H = 0$ includes the homoclinic trajectory, for which the level set $E = 0$ in (6.25) starts and ends with zero velocity at the North pole of the unit sphere. Refer to Section ?? for a discussion of the properties of motion in a cubic potential and the details of how to compute its homoclinic trajectory.

6.2 Geometric phase for the spherical pendulum

We write the Nambu bracket (6.21) for the spherical pendulum as a differential form in $\mathbb{R}^3$

$$\{F, H\}\, d^3\sigma = dC \wedge dF \wedge dH\,, \tag{6.27}$$

with oriented volume element $d^3\sigma = d\sigma_1 \wedge d\sigma_2 \wedge d\sigma_3$. Hence, on a level set of H we have the canonical Poisson bracket

$$\{f, h\} d\sigma_1 \wedge d\sigma_2 = df \wedge dh = \left(\frac{\partial f}{\partial \sigma_1}\frac{\partial h}{\partial \sigma_2} - \frac{\partial f}{\partial \sigma_2}\frac{\partial h}{\partial \sigma_1} \right) d\sigma_1 \wedge d\sigma_2 \tag{6.28}$$

and we recover equation (6.24) in canonical form with Hamiltonian

$$h(\sigma_1, \sigma_2) = -\Big(\tfrac{1}{2}\sigma_2^2 - \,\sigma_1^3 + H\sigma_1^2 + \sigma_1 \Big) = -\Big(\tfrac{1}{2}\sigma_2^2 + V(\sigma_1) \Big), \tag{6.29}$$

which, not unexpectedly, is also the conserved energy integral in (6.25) for motion on level sets of H.

For the S^1 reduction considered in the present case, the canonical 1-form is

$$p_i dq_i = \sigma_2\, d\sigma_1 + H d\psi\,, \tag{6.30}$$

where σ_1 and σ_2 are the symplectic coordinates for the level surface of H on which the reduced motion takes place and $\psi \in S^1$ is canonically conjugate to H.

Our goal is to finish the solution for the spherical pendulum motion by reconstructing the phase $\psi \in S^1$ from the symmetry-reduced motion in $(\sigma_1, \sigma_2, \sigma_3) \in \mathbb{R}^3$ on a level set of H. Rearranging equation (6.30) gives

$$H d\psi = -\,\sigma_2\, d\sigma_1 + p_i dq_i\,. \tag{6.31}$$

Thus, the phase change around a closed periodic orbit on a level set of H in the $(\sigma_1, \sigma_2, \psi, H)$ phase space decomposes into the sum of the following two parts:

$$\oint H\, d\psi = H\, \Delta\psi = \underbrace{-\oint \sigma_2\, d\sigma_1}_{\text{geometric}} + \underbrace{\oint p_i dq_i}_{\text{dynamic}}\,. \tag{6.32}$$

On writing this decomposition of the phase as

$$\Delta\psi = \Delta\psi_{geom} + \Delta\psi_{dyn}\,, \tag{6.33}$$

one sees from (6.23) that

$$H\Delta\psi_{geom} = \oint \sigma_2^2\,dt = \iint d\sigma_1 \wedge d\sigma_2 \tag{6.34}$$

is the area enclosed by the periodic orbit on a level set of H. Thus, the name: **geometric phase** for $\Delta\psi_{geom}$, because this part of the phase equals the geometric area of the periodic orbit. The rest of the phase is given by

$$H\Delta\psi_{dyn} = \oint p_i\,dq_i = \int_0^T (-\sigma_2\dot\sigma_1 + H\dot\psi)\,dt\,. \tag{6.35}$$

Hence, from the canonical equations $\dot\sigma_1 = \partial h/\partial\sigma_2$ and $\dot\psi = \partial h/\partial H$ with Hamiltonian h in (6.29), we have

$$\begin{aligned}\Delta\psi_{dyn} &= \frac{1}{H}\int_0^T \left(\sigma_2\frac{\partial h}{\partial\sigma_2} + H\frac{\partial h}{\partial H}\right)dt\\ &= \frac{2T}{H}\Big(h + \big\langle V(\sigma_1)\big\rangle - \tfrac{1}{2}H\big\langle\sigma_1^2\big\rangle\Big)\\ &= \frac{2T}{H}\Big(h + \big\langle V(\sigma_1)\big\rangle\Big) - T\big\langle\sigma_1^2\big\rangle\,,\end{aligned} \tag{6.36}$$

where T is the period of the orbit around which the integration is performed and the angle brackets $\langle\,\cdot\,\rangle$ denote time average.

The second summand $\Delta\psi_{dyn}$ in (6.33) depends on the Hamiltonian $h = E$, the orbital period T, the value of the level set H and the time averages of the potential energy and σ_1^2 over the orbit. Thus, $\Delta\psi_{dyn}$ deserves the name **dynamic phase**, since it depends on the several aspects of the dynamics along the orbit, not just its area.

This finishes the solution for the periodic motion of the spherical pendulum up to quadratures for the phase.

7 Poincaré and symplectic manifolds

Henri Poincaré

The geometry of Hamiltonian mechanics is best expressed by using exterior calculus on symplectic manifolds. Exterior calculus began with H. Poincaré and was eventually perfected by E. Cartan using methods of S. Lie. This chapter introduces key definitions and develops the necessary ingredients of exterior calculus.

This chapter casts the ideas underlying the examples we have been studying heuristically in the previous chapters into the language of differential forms.

The goals of the next few sections are, as follows.

1. Define differential forms using exterior product (wedge product) in a local basis.
2. Define the push-forward and pull-back of a differential form under a smooth invertible map.
3. Define the operation of contraction, or substitution of a vector field into a differential form.
4. Define the exterior derivative of a differential form.
5. Define Lie derivative in two equivalent ways, either dynamically as the tangent to the flow of a smooth invertible map acting by push-forward on a differential form, or algebraically by using Cartan's formula.
6. Derive the various identities for Lie derivatives acting on differential forms and illustrate them using steady incompressible fluid flows as an example.
7. Explain Nambu's bracket for divergenceless vector fields in $\mathbb{R}^3$ in the language of differential forms.
8. Define the Hodge star operation and illustrate its application in Maxwell's equations.
9. Explain Poincaré's Lemma for closed, exact and co-exact forms.

We begin by recalling Hamilton's canonical equations and using them to demonstrate Poincaré's theorem for Hamiltonian flows heuristically, by a simple direct calculation. This will serve to motivate further discussion of manifolds, tangent bundles, cotangent bundles, vector fields and differential forms in the remainder of this Chapter.

Definition

7.1 (Hamilton's canonical equations)**.**
Hamilton's canonical equations *are written on* ***phase space****, a locally Euclidean space with pairs of coordinates denoted (q,p). Namely,*

$$\frac{dq}{dt} = \frac{\partial H}{\partial p}\,, \qquad \frac{dp}{dt} = -\,\frac{\partial H}{\partial q}\,, \tag{7.1}$$

where $\partial H/\partial q$ and $\partial H/\partial p$ are the gradients of a smooth function on phase space $H(q,p)$ called the ***Hamiltonian****.*

The set of curves in phase space $(q(t),p(t))$ satisfying Hamilton's canonical equations (7.1) is called a ***Hamiltonian flow****.*

Definition

7.2 (Symplectic 2-form)**.**
The oriented area in phase space

$$\omega = dq \wedge dp = -\,dp \wedge dq$$

is called the ***symplectic 2-form****.*

Definition

7.3 (Sym·plec·tic)**.**
From the Greek for plaiting, braiding or joining together.

Definition

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

Theorem

7.5 (Poincaré's theorem). *Hamiltonian flows are symplectic. That is, they preserve the oriented phase space area $\omega = dq \wedge dp$.*

Proof. Preservation of ω may first be verified via the same formal calculation used to prove its preservation (4.20) in Theorem 4.24. Namely, along the characteristic equations of the Hamiltonian vector field $(dq/dt, dp/dt) = (\dot{q}(t), \dot{p}(t)) = (H_p, -H_q)$, for a solution of Hamilton's equations for a smooth Hamiltonian function $H(q,p)$, the flow of the symplectic two-form ω is governed by

$$\begin{aligned}\frac{d\omega}{dt} &= d\dot{q} \wedge dp + dq \wedge d\dot{p} = dH_p \wedge dp - dq \wedge dH_q \\ &= (H_{pq}dq + H_{pp}dp) \wedge dp - dq \wedge (H_{qq}dq + H_{qp}dp) \\ &= H_{pq}\, dq \wedge dp - H_{qp}\, dq \wedge dp = (H_{pq} - H_{qp})\, dq \wedge dp = 0\,.\end{aligned}$$

The first step uses the ***product rule*** for differential forms, the second uses antisymmetry of the wedge product $(dq \wedge dp = -\, dp \wedge dq)$ and last step uses equality of cross derivatives $H_{pq} = H_{qp}$ for a smooth Hamiltonian function H. □

8 Preliminaries for exterior calculus

8.1 Manifolds and bundles

Let us review some of the fundamental concepts that have already begun to emerge in the previous chapter and cast them into the language of exterior calculus.

Definition

8.1 (Smooth submanifold of $\mathbb{R}^{3N}$).
A smooth K-dimensional submanifold M of the Euclidean space $\mathbb{R}^{3N}$ is any subset which in a neighbourhood of every point on it is a graph of a smooth mapping of $\mathbb{R}^K$ into $\mathbb{R}^{(3N-K)}$ (where $\mathbb{R}^K$ and $\mathbb{R}^{(3N-K)}$ are coordinate subspaces of $\mathbb{R}^{3N} \simeq \mathbb{R}^K \times \mathbb{R}^{(3N-K)}$).

This means that every point in M has an open neighbourhood U such that the intersection $M \cap U$ is the graph of some smooth function expressing $(3N-K)$ of the standard coordinates of $\mathbb{R}^{3N}$ in terms of the other K coordinates, e.g., $(x,y,z) = (x, f(x,z), z)$ in $\mathbb{R}^3$. This is also called an **embedded submanifold**.

Definition

8.2 (Tangent vectors and tangent bundle)**.**
The solution $q(t) \in M$ is a curve (or trajectory) in manifold M parameterised by time in some interval $t \in (t_1\,,\,t_2)$. The ***tangent vector*** *of the curve $q(t)$ is the velocity $\dot{q}(t)$ along the trajectory that passes though the point $q \in M$ at time t. This is written $\dot{q} \in T_qM$, where T_qM is the tangent space at position q on the manifold M. Taking the union of the tangent spaces T_qM over the entire configuration manifold defines the* ***tangent bundle*** *$(q, \dot{q}) \in TM$.*

Remark

8.3 (Tangent and cotangent bundles)**.**
The ***configuration space*** *M has coordinates $q \in M$. The union of positions on M and tangent vectors (velocities) at each position comprises the tangent bundle TM. Its positions and momenta have* ***phase space*** *coordinates expressed as $(q, p) \in T^*M$, where T^*M is the* ***cotangent bundle*** *of the configuration space.*

The terms ***tangent bundle*** *and* ***cotangent bundle*** *introduced earlier are properly defined in the context of manifolds. See especially Definition 9.14 in the next section for a precise definition of the cotangent bundle of a manifold. Until now, we have gained intuition about geometric mechanics in the context of examples, by thinking of the tangent bundle as simply the space of positions and velocities. Likewise, we have regarded the cotangent bundle simply as a pair of vectors on an optical screen, or as the space of positions and canonical momenta for a system of particles. In this chapter, these intuitive definitions will be formalised and made precise by using the language of differential forms.*

8.2 Contraction

Definition

8.4 (Contraction)**.**
In exterior calculus, the operation of ***contraction*** *denoted as $\lrcorner$ introduces a pairing between vector fields and differential forms. Contraction is also called* ***substitution*** *of a vector field into a differential form. For basis elements in phase space, contraction defines* ***duality relations****,*

$$\partial_q \lrcorner\, dq = 1 = \partial_p \lrcorner\, dp\,, \quad \textit{and} \quad \partial_q \lrcorner\, dp = 0 = \partial_p \lrcorner\, dq\,, \tag{8.1}$$

so that differential forms are linear functions of vector fields. A ***Hamiltonian vector field****:*

$$X_H = \dot{q}\frac{\partial}{\partial q} + \dot{p}\frac{\partial}{\partial p} = H_p\partial_q - H_q\partial_p = \{\,\cdot\,,\,H\,\}\,, \tag{8.2}$$

satisfies the intriguing linear functional relations with the basis elements in phase space,

$$X_H \lrcorner\, dq = H_p \quad and \quad X_H \lrcorner\, dp = -\, H_q\,. \tag{8.3}$$

Definition

8.5 (Contraction rules with higher forms)**.**
*The rule for contraction or substitution of a vector field into a differential form is to sum the substitutions of X_H over the permutations of the factors in the differential form that bring the corresponding dual basis element into its **leftmost position**. For example, substitution of the Hamiltonian vector field X_H into the symplectic form $\omega = dq \wedge dp$ yields*

$$X_H \lrcorner\, \omega = X_H \lrcorner\, (dq \wedge dp) = (X_H \lrcorner\, dq)\, dp - (X_H \lrcorner\, dp)\, dq\,.$$

In this example, $X_H \lrcorner\, dq = H_p$ and $X_H \lrcorner\, dp = -\, H_q$, so

$$X_H \lrcorner\, \omega = H_p dp + H_q dq = dH\,,$$

which follows from the duality relations (8.1).

This calculation proves the following.

Theorem

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

$$X_H \lrcorner\, \omega \quad = \quad dH \quad with \quad \omega = dq \wedge dp\,. \tag{8.4}$$

Remark

8.7. *The purely geometric nature of relation (8.4) argues for it to be taken as the **definition** of a Hamiltonian vector field.*

Lemma

8.8 ($d^2 = 0$ for smooth phase space functions)**.**

Proof. For any smooth phase space function $H(q,p)$, one computes

$$dH = H_q dq + H_p dp$$

and taking the second exterior derivative yields

$$\begin{aligned} d^2 H &= H_{qp}\, dp \wedge dq + H_{pq}\, dq \wedge dp \\ &= (H_{pq} - H_{qp})\, dq \wedge dp = 0\,. \end{aligned}$$

□

Relation (8.4) also implies the following.

Corollary

8.9. *The flow of X_H preserves the exact 2-form ω for any Hamiltonian H.*

Proof. Preservation of ω may be verified first by a formal calculation using (8.4). Along $(dq/dt, dp/dt) = (\dot q, \dot p) = (H_p, -H_q)$, for a solution of Hamilton's equations, we have

$$\begin{aligned} \frac{d\omega}{dt} &= d\dot q \wedge dp + dq \wedge d\dot p = dH_p \wedge dp - dq \wedge dH_q \\ &= d(H_p\, dp + H_q\, dq) = d(X_H \,\lrcorner\, \omega) = d(dH) = 0\,. \end{aligned}$$

The first step uses the ***product rule*** for differential forms and the third and last steps use the property of the exterior derivative d that $d^2 = 0$ for continuous forms. The latter is due to equality of cross derivatives $H_{pq} = H_{qp}$ and antisymmetry of the wedge product: $dq \wedge dp = -\, dp \wedge dq$. □

Definition

8.10 (Symplectic flow)**.**
A flow is ***symplectic****, if it preserves the phase space area, or symplectic two-form, $\omega = dq \wedge dp$.*

According to this definition, Corollary 8.9 may be simply re-stated as

Corollary

8.11 (Poincaré's theorem)**.**
The flow of a Hamiltonian vector field is symplectic.

Definition

8.12 (Canonical transformations)**.**
A smooth invertible map g of the phase space T^*M is called a ***canonical transformation****, if it preserves the canonical symplectic form ω on T^*M, i.e., $g^*\omega = \omega$, where $g^*\omega$ denotes the transformation of ω under the map g.*

Remark

8.13. *The usage of the notation $g^*\omega$ as the transformation of ω under the map g foreshadows the idea of* ***pull-back****, made more precise in Definition 9.18.*

Remark

8.14 (Criterion for a canonical transformation)**.** *Suppose in the original coordinates (p, q) the symplectic form is expressed as $\omega = dq \wedge dp$. A transformation $g : T^*M \mapsto T^*M$ written as $(Q, P) = (Q(p, q), P(p, q)$ is canonical if the direct computation shows that $dQ \wedge dP = c\, dq \wedge dp$, up to a constant factor c. (Such a constant factor c is unimportant, since it may be absorbed into the units of time in Hamilton's canonical equations.)*

Remark

8.15. *By Corollary 8.11 of Poincaré's Theorem 7.5, the Hamiltonian phase flow g_t is a one-parameter group of canonical transformations.*

Theorem

8.16 (Preservation of Hamiltonian form)**.**
Canonical transformations preserve Hamiltonian form.

Proof. The coordinate-free relation (8.4) keeps its form if

$$dQ \wedge dP = c\, dq \wedge dp \,,$$

up to the constant factor c. Hence, Hamilton's equations re-emerge in canonical form in the new coordinates, up to a rescaling by c which may be absorbed into the units of time. □

Remark

8.17 (Lagrange-Poincaré theorem)**.**
Lagrange's equations

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

imply an evolution equation for the differential one-form

$$\begin{aligned} \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot q}\, dq\Big) &= \Big(\frac{d}{dt}\frac{\partial L}{\partial \dot q}\Big)\, dq + \frac{\partial L}{\partial \dot q}\, d\dot q \\ &= dL\,. \end{aligned}$$

Applying the exterior derivative, commuting it with the time derivative and using $d^2 = 0$ *yields*

$$\frac{d}{dt}\Big(d\,\frac{\partial L}{\partial \dot q} \wedge dq\Big) = 0\,, \tag{8.5}$$

whose preservation is Lagrange's counterpart of Poincaré's theorem on the symplectic property of Hamiltonian flows, found and used in ray optics almost a century before Poincaré!

The components $\partial L/\partial \dot q^a$ of the differential one-form,

$$\theta_L = \frac{\partial L}{\partial \dot q^a}\, dq^a\,,$$

transform under a change of coordinates on M as a covariant vector. That is, under a change of coordinates $Q^i = Q^i(q)$, we find

$$\theta_L = \frac{\partial L}{\partial \dot q^a}\, dq^a = \frac{\partial L}{\partial \dot q^a}\frac{\partial q^a}{\partial Q^b}\, dQ^b\,.$$

Proposition

8.18. *A Lagrangian system (M, L) is* ***non-degenerate*** *(hyperregular) if and only if the two-form $d\theta_L$ on TM is non-degenerate.*

Proof. In coordinates on TM with indices $a, b = 1, \ldots, K$,

$$d\theta_L = \frac{\partial^2 L}{\partial \dot{q}^a \partial \dot{q}^b} d\dot{q}^b \wedge dq^a + \frac{\partial^2 L}{\partial \dot{q}^a \partial q^b} dq^b \wedge dq^a \,, \tag{8.6}$$

so that the $2K \times 2K$ matrix corresponding to the two-form $d\theta_L$ is non-degenerate if and only if the $K \times K$ matrix H_L in (4.3) is non-degenerate. □

Definition

8.19 (Canonical, or Liouville one-form)**.**
The one-form θ on T^*M, defined in phase space coordinates by

$$\theta = p_a dq^a = p \cdot dq \,,$$

is called the ***canonical, or Liouville one-form****. Its exterior derivative yields (minus) the symplectic two-form,*

$$d\theta = -\,\omega = dp_a \wedge dq^a \,.$$

Definition

8.20 (Cotangent lift)**.**
A change of base coordinates $Q^b = Q^b(q)$ in the cotangent bundle T^*M of a manifold M induces a change in its fibre coordinates

$$p_a = P_b \frac{\partial Q^b}{\partial q^a} \quad \textit{such that} \quad p_a dq^a = P_b dQ^b \,,$$

so (Q^b, P_b) are also canonical coordinates. This transformation of the fibre coordinates (canonical momenta) is called the ***cotangent lift*** *of the base transformation.*

8.3 Hamilton-Jacobi equation

Definition

8.21 (Steady generating functions)**.**
A sufficient condition for a transformation $(Q,P)=(Q(p,q),P(p,q))$ to be canonical is that

$$P\cdot dQ - p\cdot dq = dF\,. \tag{8.7}$$

Following Hamilton's approach to geometric optics, this relation defines a ***generating function*** *F, which may be chosen to depend on one of the old phase space variables (p,q) and one of the new phase space variables (Q,P).*

Remark

8.22 (Time-dependent generating functions)**.**

Generating functions based on the phase space action in (4.8) lead to the ***Hamilton-Jacobi equation***. For this, one considers a time-dependent transformation $(Q,P)=(Q(p,q,t),P(p,q,t))$, under which the integrand of the phase space action in (4.8) transforms as

$$p\cdot dq - H(q,p)dt = P\cdot dQ - K(Q,P)dt + dS\,, \tag{8.8}$$

in which we require the transformed Hamiltonian to vanish identically, that is

$$K(Q,P)\equiv 0\,.$$

Hence, all its derivatives are also zero, and Hamilton's equations become trivial:

$$\frac{dP}{dt} = 0 = \frac{dQ}{dt}\,.$$

That is, the new generalised coordinates Q and momenta P are constants of motion. Under this condition, one may rearrange equation (8.8), so that

$$\begin{aligned} dS &= \frac{\partial S}{\partial q}\cdot dq + \frac{\partial S}{\partial t}\,dt + \frac{\partial S}{\partial Q}\cdot dQ \\ &= p\cdot dq - H(q,p)dt - P\cdot dQ\,. \end{aligned} \tag{8.9}$$

Consequently, the generating function $S(q,t,Q)$ satisfies, term by term,

$$\frac{\partial S}{\partial q} = p\,, \quad \frac{\partial S}{\partial Q} = -P\,, \quad \frac{\partial S}{\partial t} + H(q,p) = 0\,. \tag{8.10}$$

Combining these equations results in the ***Hamilton-Jacobi equation***, written in the form,

$$\frac{\partial S}{\partial t}(q,t,Q) + H\Big(q, \frac{\partial S}{\partial q}\Big) = 0\,. \tag{8.11}$$

Thus, the Hamilton-Jacobi equation (8.11) is a single, first-order nonlinear partial differential equation for the function S of the N generalised coordinates $q = (q_1, \dots, q_N)$ and the time t. The generalised momenta do not appear, except as derivatives of S. Remarkably, when the $2N$ constant parameters Q and P are identified with the initial values $Q = q(t_a)$, $P = p(t_a)$, the function S is equal to the classical action,

$$S(q,t,Q) = \int_{t_a}^{t} dS = \int_{t_a}^{t} p \cdot dq - H(q,p)\,dt\,. \tag{8.12}$$

In geometrical optics, the solution S of the Hamilton-Jacobi equation (8.11) is called ***Hamilton's characteristic function***.

Remark

8.23 (Hamilton's characteristic function in optics)**.**

Hamilton's characteristic function S in (8.12) has an interesting interpretation in terms of geometric optics. As we saw in Chapter **??**, the tangents to Fermat's light rays in an isotropic medium are normal to Huygens wave fronts. The phase of such a wave front is given by [BoWo1965]

$$\phi = \int \mathbf{k} \cdot d\mathbf{r} - \omega(\mathbf{k}, \mathbf{r})\,dt\,. \tag{8.13}$$

The Huygens wave front is a travelling wave, for which the phase ϕ is constant. For such a wave, the phase shift $\int \mathbf{k} \cdot d\mathbf{r}$ along a ray trajectory such as $\mathbf{r}(t)$ in Figure **??** is given by $\int \omega dt$.

On comparing the phase relation in (8.13) to the Hamilton-Jacobi solution in equation (8.12), one sees that Hamilton's characteristic function S plays the role of the phase ϕ of the wave front. The frequency ω of the travelling wave plays the role of the Hamiltonian and the wavevector $\mathbf{k}$ corresponds to the canonical momentum. Physically, the index of refraction $n(\mathbf{r})$ of the medium at position $\mathbf{r}$ enters the travelling wave phase speed ω/k as

$$\frac{\omega}{k} = \frac{c}{n(\mathbf{r})}\,, \qquad k = |\mathbf{k}|\,,$$

where c is the speed of light in a vacuum. Consequently, we may write Hamilton's canonical equations for a wave front as

$$\frac{d\mathbf{r}}{dt} = \frac{\partial\omega}{\partial\mathbf{k}} = \frac{c}{n}\frac{\mathbf{k}}{k} = \frac{c^2}{n^2\omega}\mathbf{k}\,, \tag{8.14}$$

$$\frac{d\mathbf{k}}{dt} = -\frac{\partial\omega}{\partial\mathbf{r}} = \frac{ck}{2n^3}\frac{\partial n^2}{\partial\mathbf{r}} = \frac{\omega}{n}\frac{\partial n}{\partial\mathbf{r}}\,. \tag{8.15}$$

After a short manipulation, these canonical equations combine into

$$\frac{n^2}{c}\frac{d}{dt}\left(\frac{n^2}{c}\frac{d\mathbf{r}}{dt}\right) = \frac{1}{2}\frac{\partial n^2}{\partial\mathbf{r}}\,. \tag{8.16}$$

In terms of a different variable time increment $cdt = n^2 d\tau$, equation (8.16) may also be expressed in the form of

$$\frac{d^2\mathbf{r}}{d\tau^2} = \frac{1}{2}\frac{\partial n^2}{\partial\mathbf{r}} \qquad \textit{(Newton's 2nd Law)} \tag{8.17}$$

If instead of τ we define the variable time increment $cdt = nd\sigma$, then equation (8.16) takes the form of the ***eikonal equation*** (**??**) for the paths of light rays in geometric optics, $\mathbf{r}(\sigma) \in \mathbb{R}^3$ as

$$\frac{d}{d\sigma}\left(n(\mathbf{r})\frac{d\mathbf{r}}{d\sigma}\right) = \frac{\partial n}{\partial\mathbf{r}} \qquad \textit{(Eikonal equation)} \tag{8.18}$$

As discussed in Chapter **??**, this equation follows from Fermat's principle of stationarity of the optical length under variations of the ray paths,

$$\delta\int_A^B n(\mathbf{r}(\sigma))\,d\sigma = 0 \qquad \textit{(Fermat's principle)} \tag{8.19}$$

with arclength parameter σ, satisfying $d\sigma^2 = d\mathbf{r}(\sigma)\cdot d\mathbf{r}(\sigma)$ and, hence, $|d\mathbf{r}/d\sigma| = 1$.

From this vantage point, one sees that replacing $\mathbf{k} \to \frac{\omega}{c}\nabla S$ in Hamilton's first equation (8.14) yields

$$n(\mathbf{r})\frac{d\mathbf{r}}{d\sigma} = \nabla S(\mathbf{r}) \qquad \textit{(Huygens equation)} \tag{8.20}$$

from which the eikonal equation (8.18) may be recovered by differentiating, using $d/d\sigma = n^{-1}\nabla S\cdot\nabla$ and $|\nabla S|^2 = n^2$.

Thus, the Hamilton-Jacobi equation (8.11) includes and unifies the ideas that originated with Fermat, Huygens and Newton.

Remark

8.24 (The threshold of quantum mechanics).

In a paper based on his PhD thesis, Feynman [Fe1948] derived a new formulation of quantum mechanics based on summing the complex amplitude for each path $\exp(iS/\hbar)$, with action S given by the Hamilton-Jacobi solution in equation (8.12), over all possible paths between the initial and final points. Earlier Dirac [Di1933] had considered a similar idea, but Dirac had considered only the classical path. Feynman showed that quantum mechanics emerges when the amplitudes $\exp(iS/\hbar)$ for *all* paths are summed. That is, the amplitudes for all paths are added together, then their modulus-squared is taken according to the quantum mechanical rule for obtaining a probability density. Perhaps not unexpectedly, Feynman's original paper [Fe1948] which laid the foundations of a new formulation of quantum mechanics was *rejected* by the mainstream scientific journal then, Physical Review!

Feynman's formulation of quantum mechanics provides an extremely elegant view of classical mechanics as being the $\hbar \to 0$ limit of quantum mechanics, through the principle of stationary phase. In this limit, only the path for which S is stationary (i.e., satisfies Hamilton's principle) contributes to the sum over all paths, and the particle traverses a single trajectory, rather than many. For more information, see [Fe1948, FeHi1965, Di1981].

9 Differential forms and Lie derivatives

9.1 Exterior calculus with differential forms

Various concepts involving differential forms have already emerged heuristically in our earlier discussions of the relations among Lagrangian and Hamiltonian formulations of mechanics. In this chapter, we shall reprise the relationships among these concepts and set up a framework for using differential forms that generalises the theorems of vector calculus involving grad, div and curl, and the integral theorems of Green, Gauss and Stokes so that they apply to manifolds of arbitrary dimension.

Definition

9.1 (Velocity vectors of smooth curves).

Consider an arbitrary curve $c(t)$ that maps an open interval $t \in (-\epsilon, \epsilon) \subset \mathbb{R}$ around the point $t = 0$ to the manifold M:

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

*with $c(0) = x$. Its **velocity vector** at x is defined by $c'(0) := \frac{dc}{dt}\big|_{t=0} = v$.*

Definition

9.2 (Tangent space to a smooth manifold)**.**
The space of velocities v tangent to the manifold at a point $x \in M$ forms a vector space called the ***tangent space*** *to M at $x \in M$. This vector space is denoted as T_xM.*

Definition

9.3 (Tangent bundle over a smooth manifold)**.**
The disjoint union of tangent spaces to M at the points $x \in M$ given by

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

is a vector space called the ***tangent bundle*** *to M and is denoted as TM.*

Definition

9.4. *[Differential of a smooth function]*
Let $f : M \mapsto \mathbb{R}$ be a smooth, real-valued function on an n-dimensional manifold M. The ***differential*** *of f at a point $x \in M$ is a linear map $df(x) : T_xM \mapsto \mathbb{R}$, from the tangent space T_xM of M at x to the real numbers.*

Definition

9.5 (Differentiable map)**.**
A map $f : M \to N$ from manifold M to manifold N is said to be differentiable *(resp. C^k) if it is represented in local coordinates on M and N by differentiable (resp. C^k) functions.*

Definition

9.6 (Derivative of a differentiable map)**.**
The ***derivative*** *of a differentiable map*

$$f : M \to N$$

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

$$T_x f : T_x M \to T_x N,$$

constructed for $v \in T_x M$ by using the chain rule to compute,

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

Thus $T_x f \cdot v$ is the ***velocity vector*** *at $t = 0$ of the curve $f \circ c : \mathbb{R} \to N$ at the point $x \in M$.*

Remark

9.7. *The tangent vectors of the map $f : M \to N$ define a space of linear operators at each point x in M, satisfying*
(i) $T_x(f + g) = T_x f + T_x g$ (linearity), and
(ii) $T_x(fg) = (T_x f)g + f(T_x g)$ (the Leibniz rule).

Definition

9.8 (Tangent lift)**.**
The union $Tf = \bigcup_x T_x f$ of the derivatives $T_x f : T_x M \to T_x N$ over points $x \in M$ is called the ***tangent lift*** *of the map $f : M \to N$.*

Remark

9.9. *The chain-rule definition of the derivative $T_x f$ of a differentiable map at a point x depends on the function f and the vector v. Other degrees of differentiability are possible. For example, if M and N are manifolds and $f : M \to N$ is of class C^{k+1}, then the tangent lift (Jacobian) $T_x f : T_x M \to T_x N$ is C^k.*

Definition

9.10 (Vector field)**.**
A ***vector field*** *X on a manifold M is a map : $M \to TM$ that assigns a vector $X(x)$ at any point $x \in M$. The real vector space of vector fields on M is denoted $\mathfrak{X}(M)$.*

Definition

9.11 (Local basis of a vector field)**.**
*A **basis** of the vector space T_xM may be obtained by using the gradient operator, written as $\nabla = (\partial/\partial x^1, \partial/\partial x^2, \dots, \partial/\partial x^n)$ in local coordinates. In these local coordinates a vector field X has **components** X^j given by*

$$X = X^j \frac{\partial}{\partial x^j} =: X^j \partial_j \,,$$

where repeated indices are summed over their range. (In this case $j = 1, 2, \dots, n$.)

Definition

9.12 (Dual basis)**.**
*As in Definition 9.11, relative to the local coordinate basis $\partial_j = \partial/\partial x^j$, $j = 1, 2, \dots, n$ of the tangent space T_xM, one may write the **dual basis** as dx^k, $k = 1, 2, \dots, n$, so that, in familiar notation, the differential of a function f is given by*

$$df = \frac{\partial f}{\partial x^k}\, dx^k \,,$$

and again one sums on repeated indices.

Definition

9.13 (Subscript-comma notation)**.**
***Subscript-comma notation** abbreviates partial derivatives as*

$$f_{,k} := \frac{\partial f}{\partial x^k} \,, \quad \text{so that} \quad df = f_{,k}\, dx^k \,.$$

Definition

9.14 (Cotangent space of M at x)**.**
*Being a linear map from the tangent space T_xM of M at x to the reals, the differential defines the space T_x^*M dual to T_xM. The dual space T_x^*M is called the **cotangent space** of M at x.*

Definition

9.15 (Tangent and cotangent bundles)**.** *The union of tangent spaces T_xM over all $x \in M$ is the **tangent bundle** TM of the manifold M. Its dual is the **cotangent bundle**, denoted T^*M.*

9.2 Pull-back and push-forward notation: coordinate-free representation

We introduce the pull-back and push-forward notation for changes of basis by variable transformations in functions, vector fields and differentials. Let $\phi : M \to N$ be a smooth invertible map from the manifold M to the manifold N.

$\phi^* f$ **pull-back of a function**: $\phi^* f = f \circ \phi$.

$\phi_* g$ **push-forward of a function**: $\phi_* g = g \circ \phi^{-1}$.

$\phi_* X$ **push-forward of a vector field X by ϕ**:

$$(\phi_* X)\big(\phi(z)\big) = T_z\phi \cdot X(z) \, . \tag{9.1}$$

The push-forward of a vector field X by ϕ has components,

$$(\phi_* X)^l \frac{\partial}{\partial \phi^l(z)} = X^J(z) \frac{\partial}{\partial z^J} \, ,$$

so that

$$(\phi_* X)^l = \frac{\partial \phi^l(z)}{\partial z^J} X^J(z) =: (T_z\phi \cdot X(z))^l \, . \tag{9.2}$$

This formula defines the notation $T_z\phi \cdot X(z)$.

$\phi^* Y$ **pull-back of a vector field Y by ϕ**:

$$\phi^* Y = (\phi^{-1})_* Y \, .$$

$\phi^* df$ **pull-back of differential df of function f by ϕ**:

$$\phi^* df = d(f \circ \phi) = d(\phi^* f) \, . \tag{9.3}$$

In components, this is

$$\phi^* df = df(\phi(z)) = \frac{\partial f}{\partial \phi^l(z)}(T_z\phi \cdot dz)^l = \frac{\partial f}{\partial z^J} dz^J ,$$

in which

$$(T_z\phi \cdot dz)^l = \frac{\partial \phi^l(z)}{\partial z^J}\, dz^J . \tag{9.4}$$

9.3 Wedge product of differential forms

Differential forms of higher degree may be constructed locally from the one-form basis dx^j, $j = 1, 2, \ldots, n$, by composition with the **wedge product**, or **exterior product**, denoted by the symbol $\wedge$. The geometric construction of higher-degree forms is intuitive and the wedge product is natural, if one imagines first composing the one-form basis as a set of line elements in space to construct oriented surface elements as two-forms $dx^j \wedge dx^k$, then volume elements as three-forms $dx^j \wedge dx^k \wedge dx^l$, etc. For these surface and volume elements to be **oriented**, the wedge product must be antisymmetric. That is, $dx^j \wedge dx^k = -\, dx^k \wedge dx^j$ under exchange of the order in a wedge product. By using this construction, any k-form $\alpha \in \Lambda^k$ on M may be written locally at a point $m \in M$ in the dual basis dx^j as

$$\alpha_m = \alpha_{i_1 \ldots i_k}(m) dx^{i_1} \wedge \cdots \wedge dx^{i_k} \in \Lambda^k ,\ i_1 < i_2 < \cdots < i_k , \tag{9.5}$$

where the sum over repeated indices is **ordered**, so that it must be taken over all i_j satisfying $i_1 < i_2 < \cdots < i_k$.

The rules for composition with the wedge product in the construction of k-forms Λ^k with $k \le n$ on an n-dimensional manifold are summarised in the following proposition.

Proposition

9.16 (Wedge product rules)**.**
The properties of the wedge product among differential forms in n dimensions are:

(i) *$\alpha \wedge \beta$ is **associative**: $\alpha \wedge (\beta \wedge \gamma) = (\alpha \wedge \beta) \wedge \gamma$.*

(ii) *$\alpha \wedge \beta$ is* ***bilinear*** *in α and β:*

$$\begin{aligned}(a\alpha_1 + b\alpha_2) \wedge \beta &= a\alpha_1 \wedge \beta + b\alpha_2 \wedge \beta \\ \alpha \wedge (c\beta_1 + e\beta_2) &= c\alpha \wedge \beta_1 + e\alpha \wedge \beta_2 \,,\end{aligned}$$

for $a, b, c, e \in \mathbb{R}$.

(iii) *$\alpha \wedge \beta$ is* ***anticommutative****: $\alpha \wedge \beta = (-1)^{kl} \beta \wedge \alpha$, where α is a k-form and β is an l-form. The prefactor $(-1)^{kl}$ counts the signature of the switches in sign required in reordering the wedge product so that its basis indices are strictly increasing, that is, they satisfy $i_1 < i_2 < \cdots < i_{k+l}$.*

9.4 Pull-back & push-forward of differential forms

Smooth invertible maps act on differential forms by the operations of pull-back and push-forward.

Definition

9.17 (Diffeomorphism).
A smooth invertible map whose inverse is also smooth is said to be a ***diffeomorphism****.*

Definition

9.18 (Pull-back and push-forward).
Let $\phi : M \to N$ be a smooth invertible map from the manifold M to the manifold N and let α be a k-form on N. The ***pull-back*** *$\phi^*\alpha$ of α by ϕ is defined as the k-form on M given by*

$$\phi^*\alpha_m = \alpha_{i_1 \dots i_k}(\phi(m))(T_m\phi \cdot dx)^{i_1} \wedge \cdots \wedge (T_m\phi \cdot dx)^{i_k} \,, \tag{9.6}$$

with $i_1 < i_2 < \cdots < i_k$. If the map ϕ is a diffeomorphism, the ***push-forward*** *$\phi_*\alpha$ of a k-form α by the map ϕ is defined by $\phi_*\alpha = (\phi^*)^{-1}\alpha$. That is, for diffeomorphisms, pull-back of a differential form is the inverse of push-forward.*

Example

9.19. *In the definition (9.6) of the pull-back of the k-form α, the additional notation $T_m\phi$ expresses the chain rule for change of variables in local coordinates. For example,*

$$(T_m\phi\cdot dx)^{i_1} = \frac{\partial\phi^{i_1}(m)}{\partial x^{i_A}}\,dx^{i_A}\,.$$

Thus, the pull-back of a one-form is given as in (9.2) and (9.4),

$$\begin{aligned}\phi^*\big(\mathbf{v}(\mathbf{x})\cdot d\mathbf{x}\big) &= \mathbf{v}\big(\phi(\mathbf{x})\big)\cdot d\phi(\mathbf{x})\\ &= v_{i_1}\big(\phi(\mathbf{x})\big)\left(\frac{\partial\phi^{i_1}(\mathbf{x})}{\partial x^{i_A}}\,dx^{i_A}\right)\\ &= \mathbf{v}\big(\phi(\mathbf{x})\big)\cdot(T_{\mathbf{x}}\phi\cdot d\mathbf{x})\,.\end{aligned}$$

Pull-backs of other differential forms may be built up from their basis elements, by the following.

Proposition

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

$$\phi^*(\alpha\wedge\beta) = \phi^*\alpha\wedge\phi^*\beta\,. \tag{9.7}$$

Remark

9.21. *The Definition 8.12 of a canonical transformation may now be rephrased using the pull-back operation, as follows. A smooth invertible transformation ϕ is canonical, if*

$$\phi^*\omega = c\,\omega\,,$$

for some constant $c\in\mathbb{R}$.

Likewise, Poincaré's Theorem 7.5 of invariance of the symplectic 2-form under a Hamiltonian flow ϕ_t depending on a real parameter t may be expressed in terms of the pull-back operation as

$$\phi_t^*(dq\wedge dp) = dq\wedge dp\,.$$

9.5 Summary of differential-form operations

Besides the wedge product, three basic operations are commonly applied to differential forms. These are contraction, exterior derivative and Lie derivative.

Contraction $\lrcorner$ with a vector field X lowers the degree:

$$X \lrcorner \Lambda^k \mapsto \Lambda^{k-1}\,.$$

Exterior derivative d raises the degree:

$$d\Lambda^k \mapsto \Lambda^{k+1}\,.$$

Lie derivative $\pounds_X$ by vector field X preserves the degree:

$$\pounds_X \Lambda^k \mapsto \Lambda^k\,, \quad \text{where} \quad \pounds_X \Lambda^k = \frac{d}{dt}\bigg|_{t=0} \phi_t^* \Lambda^k\,,$$

in which ϕ_t is the flow of the vector field X.

Lie derivative $\pounds_X$ satisfies Cartan's formula:

$$\pounds_X \alpha = X \lrcorner d\alpha + d(X \lrcorner \alpha) \quad \text{for} \quad \alpha \in \Lambda^k\,.$$

Remark

9.22.
Note that Lie derivative commutes with exterior derivative. That is,

$$d(\pounds_X \alpha) = \pounds_X d\alpha\,, \quad \textit{for} \quad \alpha \in \Lambda^k(M) \quad \textit{and} \quad X \in \mathfrak{X}(M)\,.$$

9.6 Contraction, or interior product

Definition

9.23 (Contraction, or interior product)**.**
Let $\alpha \in \Lambda^k$ be a k-form on a manifold M

$$\alpha = \alpha_{i_1 \dots i_k} dx^{i_1} \wedge \dots \wedge dx^{i_k} \in \Lambda^k \,, \quad \text{with} \quad i_1 < i_2 < \dots < i_k \,,$$

and let $X = X^j \partial_j$ be a vector field. The contraction, or interior product $X \,\lrcorner\, \alpha$ of a vector field X with a k-form α is defined by

$$X \,\lrcorner\, \alpha = X^j \alpha_{j i_2 \dots i_k} dx^{i_2} \wedge \dots \wedge dx^{i_k} \,. \tag{9.8}$$

Note that

$$\begin{aligned} X \,\lrcorner\, (Y \,\lrcorner\, \alpha) &= X^l Y^m \alpha_{m l i_3 \dots i_k} dx^{i_3} \wedge \dots \wedge dx^{i_k} \\ &= -Y \,\lrcorner\, (X \,\lrcorner\, \alpha) \,, \end{aligned}$$

by antisymmetry of $\alpha_{m l i_3 \dots i_k}$, particularly in its first two indices.

Remark

9.24 (Examples of contraction)**.**

(1) *A mnemonic device for keeping track of signs in contraction or substitution of a vector field into a differential form is to sum the substitutions of $X = X^j \partial_j$ over the permutations that bring the corresponding dual basis element into the leftmost position in the k-form α. For example, in two dimensions, contraction of the vector field $X = X^j \partial_j = X^1 \partial_1 + X^2 \partial_2$ into the two-form $\alpha = \alpha_{jk} dx^j \wedge dx^k$ with $\alpha_{21} = -\alpha_{12}$, yields*

$$X \,\lrcorner\, \alpha = X^j \alpha_{j i_2} dx^{i_2} = X^1 \alpha_{12} dx^2 + X^2 \alpha_{21} dx^1 \,.$$

Likewise, in three dimensions, contraction of the vector field $X = X^1 \partial_1 + X^2 \partial_2 + X^3 \partial_3$ into the three-form $\alpha = \alpha_{123} dx^1 \wedge dx^2 \wedge dx^3$ with $\alpha_{213} = -\alpha_{123}$, etc. yields

$$\begin{aligned} X \,\lrcorner\, \alpha &= X^1 \alpha_{123} dx^2 \wedge dx^3 + \textit{cyclic permutations}, \\ &= X^j \alpha_{j i_2 i_3} dx^{i_2} \wedge dx^{i_3} \quad \text{with} \quad i_2 < i_3 \,. \end{aligned}$$

(2) *The rule for contraction of a vector field with a differential form develops from the relation*

$$\partial_j \,\lrcorner\, dx^k = \delta^k_j \,,$$

in the coordinate basis $e_j = \partial_j := \partial/\partial x^j$ *and its dual basis* $e^k = dx^k$. *Contraction of a vector field with a one-form yields the dot product, or inner product, between a covariant vector and a contravariant vector is given by*

$$X^j\partial_j \,\lrcorner\, v_k dx^k = v_k\delta^k_j X^j = v_j X^j \,,$$

or, in vector notation,

$$X \,\lrcorner\, \mathbf{v}\cdot d\mathbf{x} = \mathbf{v}\cdot\mathbf{X} \,.$$

This is the ***dot product of vectors*** $\mathbf{v}$ *and* $\mathbf{X}$.

Our previous calculations for 2-forms and 3-forms provide the following additional expressions for contraction of a vector field with a differential form,

$$\begin{aligned} X \,\lrcorner\, \mathbf{B}\cdot d\mathbf{S} &= -\mathbf{X}\times\mathbf{B}\cdot d\mathbf{x} \,, \\ X \,\lrcorner\, d^{\,3}x &= \mathbf{X}\cdot d\mathbf{S} \,, \\ d(X \,\lrcorner\, d^{\,3}x) &= d(\mathbf{X}\cdot d\mathbf{S}) = (\mathrm{div}\mathbf{X})\, d^{\,3}x \,. \end{aligned}$$

Remark

9.25 (Physical examples of contraction).
The first of these contraction relations represents the Lorentz, or Coriolis force, when $\mathbf{X}$ *is particle velocity and* $\mathbf{B}$ *is either magnetic field, or rotation rate, respectively. The second contraction relation is the flux of the vector* $\mathbf{X}$ *through a surface element. The third is the exterior derivative of the second, thereby yielding the divergence of the vector* $\mathbf{X}$.

Exercise. *Show that*

$$X \,\lrcorner\, (X \,\lrcorner\, \mathbf{B}\cdot d\mathbf{S}) = 0 \,,$$

and

$$(X \,\lrcorner\, \mathbf{B}\cdot d\mathbf{S}) \wedge \mathbf{B}\cdot d\mathbf{S} = 0 \,,$$

for any vector field X *and 2-form* $\mathbf{B}\cdot d\mathbf{S}$. ★

(3) *By the linearity of its definition (9.8), contraction of a vector field X with a differential k-form α satisfies*

$$(hX) \lrcorner \, \alpha = h(X \lrcorner \, \alpha) = X \lrcorner \, h\alpha \,.$$

Proposition

9.26 (Contracting through wedge product)**.**
Let α be a k-form and β be a one-form on a manifold M and let $X = X^j \partial_j$ be a vector field. Then the contraction of X through the wedge product $\alpha \wedge \beta$ satisfies

$$X \lrcorner \, (\alpha \wedge \beta) = (X \lrcorner \, \alpha) \wedge \beta + (-1)^k \alpha \wedge (X \lrcorner \, \beta) \,. \tag{9.9}$$

Proof. The proof is a straightforward calculation using the definition of contraction. The exponent k in the factor $(-1)^k$ counts the number of exchanges needed to get the one-form β to the left-most position through the k-form α. □

Proposition

9.27 (Contraction commutes with pull-back)**.**
That is,

$$\phi^*(X(m) \lrcorner \, \alpha) = X(\phi(m)) \lrcorner \, \phi^* \alpha \,. \tag{9.10}$$

Proof. Direct verification using the relation between pull-back of forms and push-forward of vector fields. □

Definition

9.28 (Alternative notations for contraction)**.**
Besides the hook notation with $\lrcorner$, one also finds in the literature the following two alternative notations for contraction of a vector field X with k-form $\alpha \in \Lambda^k$ on a manifold M.

$$X \lrcorner \, \alpha = i_X \alpha = \alpha(X, \underbrace{\cdot, \cdot, \dots, \cdot}_{k-1 \text{ slots}}) \in \Lambda^{k-1} \,. \tag{9.11}$$

In the last alternative, one leaves a dot $(\,\cdot\,)$ in each remaining slot of the form that results after contraction. For example, contraction of the Hamiltonian vector field $X_H = \{\,\cdot\,,\, H\}$ with the symplectic 2-form $\omega \in \Lambda^2$ produces the 1-form,

$$X_H \lrcorner \, \omega = \omega(X_H, \,\cdot\,) = -\,\omega(\,\cdot\,, X_H) = dH \,.$$

Proposition

9.29 (Hamiltonian vector field definitions).
The two definitions of Hamiltonian vector field X_H

$$dH = X_H \,\lrcorner\, \omega \quad \textit{and} \quad X_H = \{\,\cdot\,,\,H\}\,,$$

are equivalent.

Proof. The symplectic Poisson bracket satisfies $\{F, H\} = \omega(X_F, X_H)$, because

$$\omega(X_F, X_H) := X_H \,\lrcorner\, X_F \,\lrcorner\, \omega = X_H \,\lrcorner\, dF = -X_F \,\lrcorner\, dH = \{F,\,H\}\,.$$

□

Remark

9.30. *The relation $\{F,\,H\} = \omega(X_F, X_H)$ means that the Hamiltonian vector field defined via the symplectic form coincides exactly with the Hamiltonian vector field defined using the Poisson bracket.*

9.7 Exterior derivative

Definition

9.31 (Exterior derivative of a k-form).
The exterior derivative of the k-form α written locally as

$$\alpha = \alpha_{i_1\dots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}\,,$$

in which one sums on all i_j satisfying $i_1 < i_2 < \cdots < i_k$), is the $(k+1)$-form $d\alpha$ written in coordinates as

$$d\alpha = d\alpha_{i_1\dots i_k} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}\,, \qquad \textit{with } i_1 < i_2 < \cdots < i_k\,,$$

where $d\alpha_{i_1\dots i_k} = (\partial\alpha_{i_1\dots i_k}/\partial x^j)\,dx^j$ summed on all j.

With this local definition of $d\alpha$ in coordinates, one may verify the following properties.

Proposition

9.32 (Properties of the exterior derivative)**.**

(i) *If α is a zero-form ($k=0$), that is $\alpha = f \in C^\infty(M)$, then df is the one-form given by the differential of f.*

(ii) *$d\alpha$ is **linear** in α, that is*

$$d(c_1\alpha_1 + c_2\alpha_2) = c_1 d\alpha_1 + c_2 d\alpha_2 \quad \textit{for constants} \quad c_1,\, c_2 \in \mathbb{R}\,.$$

(iii) *$d\alpha$ satisfies the **product rule** ,*

$$d(\alpha\wedge\beta) = (d\alpha)\wedge\beta + (-1)^k\alpha\wedge d\beta\,, \tag{9.12}$$

where α is a k-form and β is a one-form.

(iv) *$d^2 = 0$, that is, $d(d\alpha) = 0$ for any k-form α.*

(v) *d is a **local operator**, that is, $d\alpha$ depends only on local properties of α restricted to any open neighbourhood of x.*

9.8 Exercises in exterior calculus operations

Vector notation for differential basis elements
One denotes differential basis elements dx^i and $dS_i = \frac{1}{2}\epsilon_{ijk}dx^j \wedge dx^k$, for $i, j, k = 1, 2, 3$, in vector notation as

$$
\begin{aligned}
d\mathbf{x} &:= (dx^1, dx^2, dx^3)\,,\\
d\mathbf{S} &= (dS_1, dS_2, dS_3)\\
&:= (dx^2 \wedge dx^3,\, dx^3 \wedge dx^1,\, dx^1 \wedge dx^2)\,,\\
dS_i &:= \frac{1}{2}\epsilon_{ijk}dx^j \wedge dx^k\,,\\
d^{\,3}x &= d\mathrm{Vol} := dx^1 \wedge dx^2 \wedge dx^3\\
&= \frac{1}{6}\epsilon_{ijk}dx^i \wedge dx^j \wedge dx^k\,.
\end{aligned}
$$

Exercise. Vector calculus operations
Show that contraction of the vector field $X = X^j\partial_j =: \mathbf{X}\cdot\nabla$ with the differential basis elements recovers the following familiar operations among vectors:

$$
\begin{aligned}
X \,\lrcorner\, d\mathbf{x} &= \mathbf{X}\,,\\
X \,\lrcorner\, d\mathbf{S} &= \mathbf{X} \times d\mathbf{x}\,,\\
(\text{or,}\quad X \,\lrcorner\, dS_i &= \epsilon_{ijk}X^j dx^k)\\
Y \,\lrcorner\, X \,\lrcorner\, d\mathbf{S} &= \mathbf{X} \times \mathbf{Y}\,,\\
X \,\lrcorner\, d^{\,3}x &= \mathbf{X}\cdot d\mathbf{S} = X^k dS_k\,,\\
Y \,\lrcorner\, X \,\lrcorner\, d^{\,3}x &= \mathbf{X} \times \mathbf{Y}\cdot d\mathbf{x} = \epsilon_{ijk}X^iY^j dx^k\,,\\
Z \,\lrcorner\, Y \,\lrcorner\, X \,\lrcorner\, d^{\,3}x &= \mathbf{X} \times \mathbf{Y}\cdot \mathbf{Z}\,.
\end{aligned}
$$

★

Exercise. Exterior derivatives in vector notation
Show that the exterior derivative and wedge product satisfy the following relations in components and in three-dimensional vector

notation:

$$
\begin{aligned}
df &= f_{,j}\, dx^j =: \nabla f \cdot d\mathbf{x}\,,\\
0 = d^2 f &= f_{,jk}\, dx^k \wedge dx^j\,,\\
df \wedge dg &= f_{,j}\, dx^j \wedge g_{,k}\, dx^k\\
&=: (\nabla f \times \nabla g) \cdot d\mathbf{S}\,,\\
df \wedge dg \wedge dh &= f_{,j}\, dx^j \wedge g_{,k}\, dx^k \wedge h_{,l}\, dx^l\\
&=: (\nabla f \cdot \nabla g \times \nabla h)\, d^{\,3}x\,.
\end{aligned}
$$

★

Exercise. Vector calculus formulas
Show that the exterior derivative yields the following vector calculus formulas:

$$
\begin{aligned}
df &= \nabla f \cdot d\mathbf{x}\,,\\
d(\mathbf{v} \cdot d\mathbf{x}) &= (\operatorname{curl} \mathbf{v}) \cdot d\mathbf{S}\,,\\
d(\mathbf{A} \cdot d\mathbf{S}) &= (\operatorname{div} \mathbf{A})\, d^{\,3}x\,.
\end{aligned}
$$

The compatibility condition $d^2 = 0$ is written for these forms as

$$
\begin{aligned}
0 = d^2 f = d(\nabla f \cdot d\mathbf{x}) &= (\operatorname{curl} \operatorname{grad} f) \cdot d\mathbf{S}\,,\\
0 = d^2(\mathbf{v} \cdot d\mathbf{x}) = d\big((\operatorname{curl} \mathbf{v}) \cdot d\mathbf{S}\big) &= (\operatorname{div} \operatorname{curl} \mathbf{v})\, d^{\,3}x\,.
\end{aligned}
$$

The product rule (9.12) is written for these forms as

$$
\begin{aligned}
d\big(f(\mathbf{A} \cdot d\mathbf{x})\big) &= df \wedge \mathbf{A} \cdot d\mathbf{x} + f \operatorname{curl}\mathbf{A} \cdot d\mathbf{S}\\
&= \big(\nabla f \times \mathbf{A} + f \operatorname{curl}\mathbf{A}\big) \cdot d\mathbf{S}\\
&= \operatorname{curl}(f\mathbf{A}) \cdot d\mathbf{S}\,,
\end{aligned}
$$

and

$$
\begin{aligned}
d\big((\mathbf{A}\cdot d\mathbf{x})\wedge(\mathbf{B}\cdot d\mathbf{x})\big) &= (\mathrm{curl}\mathbf{A})\cdot d\mathbf{S}\wedge\mathbf{B}\cdot d\mathbf{x} \\
&\quad -\mathbf{A}\cdot d\mathbf{x}\wedge(\mathrm{curl}\mathbf{B})\cdot d\mathbf{S} \\
&= \big(\mathbf{B}\cdot\mathrm{curl}\mathbf{A}-\mathbf{A}\cdot\mathrm{curl}\mathbf{B}\big)d^{\,3}x \\
&= d\big((\mathbf{A}\times\mathbf{B})\cdot d\mathbf{S}\big) \\
&= \mathrm{div}(\mathbf{A}\times\mathbf{B})\,d^{\,3}x\,.
\end{aligned}
$$

These calculations return the familiar formulas from vector calculus for quantities curl(grad), div(curl), $\mathrm{curl}(f\mathbf{A})$ and $\mathrm{div}(\mathbf{A}\times\mathbf{B})$.

★

Exercise. Integral calculus formulas
Show that Stokes theorem for the vector calculus formulas yields the following familiar results in $\mathbb{R}^3$:

(1) The **fundamental theorem of calculus**, upon integrating df along a curve in $\mathbb{R}^3$ starting at point a and ending at point b,

$$\int_a^b df \;=\; \int_a^b \nabla f\cdot d\mathbf{x} = f(b)-f(a)\,.$$

(2) **Classical Stokes theorem**, for a compact surface S with boundary ∂S:

$$\int_S(\mathrm{curl}\,\mathbf{v})\cdot d\mathbf{S} \;=\; \oint_{\partial S}\mathbf{v}\cdot d\mathbf{x}\,.$$

(For a planar surface $\Omega\in\mathbb{R}^2$, this is **Green's theorem**.)

(3) The **Gauss divergence theorem**, for a compact spatial domain D with boundary ∂D:

$$\int_D(\mathrm{div}\,\mathbf{A})\,d^{\,3}x \;=\; \oint_{\partial D}\mathbf{A}\cdot d\mathbf{S}\,.$$

★

These exercises illustrate the following.

Theorem

9.33 (Stokes theorem)**.**
Suppose M is a compact oriented k-dimensional manifold with boundary ∂M and α is a smooth $(k-1)$-form on M. Then

$$\int_M d\alpha = \oint_{\partial M} \alpha \,.$$

9.9 Dynamic definition of Lie derivative

Definition

9.34 (Dynamic definition of Lie derivative)**.** *Let α be a k-form on a manifold M and let X be a vector field with flow ϕ_t on M. The Lie derivative of α along X is defined as*

$$\pounds_X \alpha = \frac{d}{dt}\Big|_{t=0} (\phi_t^* \alpha)\,. \tag{9.13}$$

Remark

9.35. *This is the definition we have been using all along in defining vector fields by their characteristic equations.*

Definition

9.36 (Cartan's formula for Lie derivative). *Cartan's formula defines the Lie derivative of the k-form α with respect to a vector field X in terms of the operations d and $\lrcorner$ as*

$$\pounds_X \alpha = X \lrcorner d\alpha + d(X \lrcorner \alpha)\,. \tag{9.14}$$

The proof of the equivalence of these two definitions of the Lie derivative of an arbitrary k-form is straightforward, but too cumbersome to be given here. We shall investigate the equivalence of these two definitions in a few individual cases, instead.

9.10 Poincaré's theorem

By Cartan's formula, the **Lie derivative** of a differential form ω by a Hamiltonian vector field X_H is defined by

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

Proposition

9.37. *Poincaré's Theorem 8.11 for preservation of the symplectic form ω may be rewritten using Lie derivative notation as*

$$\begin{aligned} 0 = \frac{d}{dt}\phi_t^*\omega\Big|_{t=0} = \pounds_{X_H}\omega \quad &:= \quad d(X_H \lrcorner \omega) + X_H \lrcorner d\omega \\ &=: \quad (\mathrm{div} X_H)\,\omega\,. \end{aligned} \tag{9.15}$$

The last equality defines the divergence of the vector field X_H, which vanishes by virtue of $d(X_H \lrcorner \omega) = d^2H = 0$ and $d\omega = 0$.

Remark

9.38.

- *Relation (9.15) expresses Hamiltonian dynamics as the symplectic flow in phase space of the divergenceless Hamiltonian vector field X_H.*

- *The Lie derivative operation defined in (9.15) is equivalent to the time derivative along the characteristic paths (flow) of the first order linear partial differential operator X_H, which are obtained from its characteristic equations,*

$$dt = \frac{dq}{H_p} = \frac{dp}{-H_q} \,.$$

*This equivalence instills the **dynamical meaning** of the Lie derivative. Namely,*

$$\pounds_{X_H}\omega = \frac{d}{dt}\phi_t^*\omega\Big|_{t=0}$$

is the evolution operator for the symplectic flow ϕ_t in phase space.

Theorem

9.39 (Poincaré theorem for N degrees of freedom)**.**
For a system of N degrees of freedom, the flow of a Hamiltonian vector field $X_H = \{\,\cdot\,,\,H\}$ preserves each subvolume in the phase space $T^*\mathbb{R}^N$. That is, let $\omega_n \equiv dq_n \wedge dp_n$ be the symplectic form expressed in terms of the position and momentum of the n-th particle. Then

$$\frac{d\omega_M}{dt}\Big|_{t=0} = \pounds_{X_H}\omega_M = 0\,, \quad \textit{for} \quad \omega_M = \Pi_{n=1}^M \omega_n\,,\ \textit{for all } M \le N\,.$$

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

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

Thus,

$$X_H \,\lrcorner\, \omega_n = dH := H_{p_n}\,dp_n + H_{q_n}\,dq_n$$

with $\omega_n = dq_n \wedge dp_n$ and one uses

$$\partial_{q_m} \lrcorner\, dq_n = \delta_{mn} = \partial_{p_m} \lrcorner\, dp_n$$

and

$$\partial_{q_m} \,\lrcorner\, dp_n = 0 = \partial_{p_m} \,\lrcorner\, dq_n$$

to compute

$$\frac{d\omega_n}{dt}\Big|_{t=0} = \pounds_{X_H}\omega_n := \underbrace{d(X_H \,\lrcorner\, \omega_n)}_{d(dH)=0} + \underbrace{X_H \,\lrcorner\, d\omega_n}_{=0} = 0\,, \tag{9.16}$$

where $\omega_n \equiv dq_n \wedge dp_n$ is closed $(d\omega_n = 0)$ for all n. $\square$

Remark

9.40. *Many of the following exercises may be solved (or checked) by equating the dynamical definition of Lie derivative in equation (9.13) with its geometrical definition by Cartan's formula (9.14)*

$$\begin{aligned} \pounds_X \alpha &= \frac{d}{dt}\Big|_{t=0} (\phi_t^* \alpha) \\ &= X \,\lrcorner\, d\alpha + d(X \,\lrcorner\, \alpha)\,, \end{aligned}$$

where α is a k-form on a manifold M and X is a smooth vector field with flow ϕ_t on M. Informed by this equality, one may derive various Lie-derivative relations by differentiating the properties of the pull-back ϕ_t^*, which commutes with exterior derivative as in (9.3), wedge product as in (9.7) and contraction as in (9.10). That is, for $m \in M$,

$$\begin{aligned} d(\phi_t^* \alpha) &= \phi_t^* d\alpha\,, \\ \phi_t^*(\alpha \wedge \beta) &= \phi_t^* \alpha \wedge \phi_t^* \beta\,, \\ \phi_t^*(X(m) \,\lrcorner\, \alpha) &= X(\phi_t(m)) \,\lrcorner\, \phi_t^* \alpha\,. \end{aligned}$$

9.11 Lie derivative exercises

Exercise. Lie derivative of forms in $\mathbb{R}^3$
Show that both the dynamic definition and Cartan's formula imply the following Lie derivative relations in vector notation,

(a) $\pounds_X f = X \,\lrcorner\, df = \mathbf{X}\cdot\nabla f\,,$

(b) $\pounds_X\,(\mathbf{v}\cdot d\mathbf{x}) = \big(-\ \mathbf{X}\times\mathrm{curl}\,\mathbf{v} + \nabla(\mathbf{X}\cdot\mathbf{v})\big)\cdot d\mathbf{x}\,,$

(c) $\pounds_X(\boldsymbol{\omega}\cdot d\mathbf{S}) = \big(-\ \mathrm{curl}\,(\mathbf{X}\times\boldsymbol{\omega}) + \mathbf{X}\,\mathrm{div}\,\boldsymbol{\omega}\big)\cdot d\mathbf{S}\,,$

(d) $\pounds_X(f\,d^{\,3}x) = (\mathrm{div}\,f\mathbf{X})\,d^{\,3}x\,.$

★

Exercise. Lie derivative identities for k-forms
Show that both the dynamic definition and Cartan's formula imply the following Lie derivative identities for a k-form α:

(a) $\pounds_{fX}\alpha = f\pounds_X\alpha + df\wedge(X\,\lrcorner\,\alpha)\,,$

(b) $\pounds_X d\alpha = d\big(\pounds_X\alpha\big)\,,$

(c) $\pounds_X(X\,\lrcorner\,\alpha) = X\,\lrcorner\,\pounds_X\alpha\,,$

(d) $\pounds_X(Y\,\lrcorner\,\alpha) = (\pounds_X Y)\,\lrcorner\,\alpha + Y\,\lrcorner\,(\pounds_X\alpha).$

(e) When $k=1$ so that α is a 1-form ($\alpha = d\mathbf{x}$), show that the previous exercise **(d)** implies a useful relation for $(\pounds_X Y)$. Namely,

$$\pounds_X(Y\,\lrcorner\,d\mathbf{x}) = \pounds_X Y\,\lrcorner\,d\mathbf{x} + Y\,\lrcorner\,\pounds_X d\mathbf{x}\,, \tag{9.17}$$

which implies the relation,

$$\pounds_X Y = [X,\,Y]\,, \tag{9.18}$$

where $[X,\,Y]$ is the Jacobi-Lie bracket (**??**) of vector fields X and Y.

(f) Use the two properties

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

for contraction ($\lrcorner$) and

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

for exterior derivative (d), along with Cartan's formula, to verify the product rule for Lie derivative of the wedge product,

$$\pounds_X(\alpha \wedge \beta) = (\pounds_X \alpha) \wedge \beta + \alpha \wedge \pounds_X \beta \,. \tag{9.19}$$

The product rule for Lie derivative (9.19) also follows immediately from its dynamical definition (9.13).

(g) Use

$$[X\,,\,Y] \,\lrcorner\, \alpha = \pounds_X(Y \,\lrcorner\, \alpha) - Y \,\lrcorner\, (\pounds_X \alpha)\,, \tag{9.20}$$

as verified in part **(d)** in concert with the definition(s) of Lie derivative to show,

$$\pounds_{[X\,,\,Y]}\alpha = \pounds_X \pounds_Y \alpha - \pounds_Y \pounds_X \alpha \,. \tag{9.21}$$

(h) Use the result of **(g)** to verify the **Jacobi identity for the Lie derivative**,

$$\pounds_{[\,Z\,,\,[X\,,\,Y]]}\,\alpha + \pounds_{[\,X\,,\,[Y\,,\,Z]]}\,\alpha + \pounds_{[\,Y\,,\,[Z\,,\,X]]}\,\alpha = 0\,.$$

★

10 Formulations of ideal fluid dynamics

10.1 Euler's fluid equations

Euler's equations for the incompressible motion of an ideal flow of a fluid of unit density and velocity $\mathbf{u}$ satisfying $\mathrm{div}\mathbf{u} = 0$ in a rotating frame with Coriolis parameter $\mathrm{curl}\mathbf{R} = 2\boldsymbol{\Omega}$ are given in the form of Newton's Law of Force by

$$\underbrace{\partial_t\,\mathbf{u} + \mathbf{u}\cdot\nabla\mathbf{u}}_{\text{Acceleration}} = \underbrace{\mathbf{u}\times 2\boldsymbol{\Omega}}_{\text{Coriolis}} - \underbrace{\nabla p}_{\text{Pressure}} \,. \tag{10.1}$$

Requiring preservation of the divergence-free (volume preserving) constraint $\nabla\cdot\mathbf{u} = 0$ results in a Poisson equation for pressure p, which may be written in several equivalent forms,

$$\begin{aligned} -\Delta p &= \mathrm{div}\big(\mathbf{u}\cdot\nabla\mathbf{u} - \mathbf{u}\times 2\boldsymbol{\Omega}\big)\,, \\ &= u_{i,j}u_{j,i} - \mathrm{div}\,\big(\mathbf{u}\times 2\boldsymbol{\Omega}\big)\,, \\ &= \mathrm{tr}\,\mathsf{S}^2 - \frac{1}{2}|\mathrm{curl}\mathbf{u}|^2 - \mathrm{div}\,\big(\mathbf{u}\times 2\boldsymbol{\Omega}\big)\,, \end{aligned} \tag{10.2}$$

where $\mathsf{S} = \frac{1}{2}(\nabla \mathbf{u} + \nabla \mathbf{u}^T)$ is the **strain-rate tensor**.

The Newton's Law equation for Euler fluid motion in (10.1) may be rearranged into an alternative form,

$$\partial_t \mathbf{v} - \mathbf{u} \times \boldsymbol{\omega} + \nabla\Big(p + \frac{1}{2}|\mathbf{u}|^2\Big) = 0\,, \tag{10.3}$$

where we denote

$$\mathbf{v} \equiv \mathbf{u} + \mathbf{R}\,, \quad \boldsymbol{\omega} = \mathrm{curl}\mathbf{v} = \mathrm{curl}\mathbf{u} + 2\boldsymbol{\Omega}\,, \tag{10.4}$$

and introduce the **Lamb vector**,

$$\boldsymbol{\ell} := -\,\mathbf{u} \times \boldsymbol{\omega}\,, \tag{10.5}$$

which represents the nonlinearity in Euler's fluid equation (10.3). The Poisson equation (10.2) for pressure p may now be expressed in terms of the divergence of the Lamb vector,

$$-\,\Delta\big(p + \tfrac{1}{2}|\mathbf{u}|^2\,\big) = \mathrm{div}(-\,\mathbf{u} \times \mathrm{curl}\,\mathbf{v}) = \mathrm{div}\,\boldsymbol{\ell}\,. \tag{10.6}$$

Remark

10.1 (Boundary Conditions).

Because the velocity $\mathbf{u}$ *must be tangent to any fixed boundary, the normal component of the motion equation must vanish. This requirement produces a Neumann condition for pressure given by*

$$\partial_n\big(p + \tfrac{1}{2}|\mathbf{u}|^2\big) + \hat{\mathbf{n}} \cdot \boldsymbol{\ell} = 0\,, \tag{10.7}$$

at a fixed boundary with unit outward normal vector $\hat{\mathbf{n}}$.

Remark

10.2 (Helmholtz vorticity dynamics).

Taking the curl of the Euler fluid equation (10.3) yields the ***Helmholtz vorticity equation***

$$\partial_t\,\boldsymbol{\omega} - \mathrm{curl}\,(\mathbf{u} \times \boldsymbol{\omega}) = 0\,, \tag{10.8}$$

whose geometrical meaning will emerge in discussing Stokes Theorem 10.5 for the vorticity of a rotating fluid.

The rotation terms have now been fully integrated into both the dynamics and the boundary conditions. In this form, the **Kelvin circulation theorem** and the **Stokes vorticity theorem** will emerge naturally together as geometrical statements.

Theorem

10.3 (Kelvin's circulation theorem)**.**
The Euler equations (10.1) preserve the circulation integral $I(t)$ defined by

$$I(t) = \oint_{c(\mathbf{u})} \mathbf{v}\cdot d\mathbf{x}\,, \tag{10.9}$$

where $c(\mathbf{u})$ is a closed circuit moving with the fluid at velocity $\mathbf{u}$.

Proof. The dynamical definition of Lie derivative (9.13) yields the following for the time rate of change of this circulation integral,

$$\begin{aligned}\frac{d}{dt}\oint_{c(\mathbf{u})} \mathbf{v}\cdot d\mathbf{x} &= \oint_{c(\mathbf{u})} \Big(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\Big)(\mathbf{v}\cdot d\mathbf{x}) \\ &= \oint_{c(\mathbf{u})} \Big(\frac{\partial \mathbf{v}}{\partial t} + \frac{\partial \mathbf{v}}{\partial x^j}u^j + v_j\frac{\partial u^j}{\partial \mathbf{x}}\Big)\cdot d\mathbf{x} \\ &= -\oint_{c(\mathbf{u})} \nabla\Big(p + \frac12|\mathbf{u}|^2 - \mathbf{u}\cdot\mathbf{v}\Big)\cdot d\mathbf{x} \\ &= -\oint_{c(\mathbf{u})} d\Big(p + \frac12|\mathbf{u}|^2 - \mathbf{u}\cdot\mathbf{v}\Big) = 0\,.\end{aligned} \tag{10.10}$$

The Cartan formula (9.14) defines the Lie derivative of the circulation integrand in an equivalent form that we need for the third step and will also use in a moment for the Stokes theorem,

$$\begin{aligned}\pounds_{\mathbf{u}}(\mathbf{v}\cdot d\mathbf{x}) &= \big(\mathbf{u}\cdot\nabla\mathbf{v} + v_j\nabla u^j\big)\cdot d\mathbf{x} \\ &= u \,\lrcorner\, d(\mathbf{v}\cdot d\mathbf{x}) + d(u \,\lrcorner\, \mathbf{v}\cdot d\mathbf{x}) \\ &= u \,\lrcorner\, d(\operatorname{curl}\mathbf{v}\cdot d\mathbf{S}) + d(\mathbf{u}\cdot\mathbf{v}) \\ &= \big(-\,\mathbf{u}\times\operatorname{curl}\mathbf{v} + \nabla(\mathbf{u}\cdot\mathbf{v})\big)\cdot d\mathbf{x}\,.\end{aligned} \tag{10.11}$$

This identity recasts Euler's equation into the following geometric form,

$$\begin{aligned}\Big(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\Big)(\mathbf{v}\cdot d\mathbf{x}) &= \big(\partial_t\mathbf{v} -\,\mathbf{u}\times\operatorname{curl}\mathbf{v} + \nabla(\mathbf{u}\cdot\mathbf{v})\big)\cdot d\mathbf{x} \\ &= -\nabla\Big(p + \frac12|\mathbf{u}|^2 - \mathbf{u}\cdot\mathbf{v}\Big)\cdot d\mathbf{x} \\ &= -\,d\Big(p + \frac12|\mathbf{u}|^2 - \mathbf{u}\cdot\mathbf{v}\Big).\end{aligned} \tag{10.12}$$

This finishes the last step in the proof (10.10), because the integral of an exact differential around a closed loop vanishes. □

The exterior derivative of the Euler fluid equation in the form (10.12) yields Stokes theorem, after using the commutativity of the exterior and Lie derivatives $[d,\, \pounds_{\mathbf{u}}] = 0$,

$$\begin{aligned}
d\pounds_{\mathbf{u}}(\mathbf{v}\cdot d\mathbf{x}) &= \pounds_{\mathbf{u}}\, d(\mathbf{v}\cdot d\mathbf{x}) \\
&= \pounds_{\mathbf{u}}(\mathrm{curl}\mathbf{v}\cdot d\mathbf{S}) \\
&= -\,\mathrm{curl}\,\big(\mathbf{u}\times\mathrm{curl}\mathbf{v}\big)\cdot d\mathbf{S} \\
&= \big[\mathbf{u}\cdot\nabla\mathrm{curl}\mathbf{v} + \mathrm{curl}\mathbf{v}(\mathrm{div}\mathbf{u}) - (\mathrm{curl}\mathbf{v})\cdot\nabla\mathbf{u}\big]\cdot d\mathbf{S} \\
\text{(by } \mathrm{div}\mathbf{u}=0) \quad &= \big[\mathbf{u}\cdot\nabla\mathrm{curl}\mathbf{v} - (\mathrm{curl}\mathbf{v})\cdot\nabla\mathbf{u}\big]\cdot d\mathbf{S} \\
&=: \,[\,u,\, \mathrm{curl}v\,]\cdot d\mathbf{S}\,,
\end{aligned} \tag{10.13}$$

where $[\,u,\,\mathrm{curl}v\,]$ denotes the **Jacobi-Lie bracket** (??) of the vector fields u and $\mathrm{curl}v$. This calculation proves the following.

Theorem

10.4. *Euler's fluid equations (10.3) imply that*

$$\frac{\partial\omega}{\partial t} = -\,[\,u,\,\omega\,] \tag{10.14}$$

where $[\,u,\,\omega\,]$ denotes the Jacobi-Lie bracket (??) of the divergenceless vector fields u and $\omega := \mathrm{curl}v$.

The exterior derivative of Euler's equation in its geometric form (10.12) is equivalent to the curl of its vector form (10.3). That is,

$$d\Big(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\Big)(\mathbf{v}\cdot d\mathbf{x}) = \Big(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\Big)(\mathrm{curl}\mathbf{v}\cdot d\mathbf{S}) = 0\,. \tag{10.15}$$

Hence from the calculation in (10.13) and the Helmholtz vorticity equation (10.15) we have

$$\Big(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\Big)(\mathrm{curl}\mathbf{v}\cdot d\mathbf{S}) = \Big(\partial_t\,\boldsymbol{\omega} - \mathrm{curl}\,(\mathbf{u}\times\boldsymbol{\omega})\Big)\cdot d\mathbf{S} = 0\,, \tag{10.16}$$

in which one denotes $\boldsymbol{\omega} := \mathrm{curl}\mathbf{v}$. This Lie-derivative version of the Helmholtz vorticity equation may be used to prove the following form of Stokes theorem for the Euler equations in a rotating frame.

Theorem

10.5 (Stokes theorem for vorticity of a rotating fluid)**.**

$$\begin{aligned}\frac{d}{dt}\iint_{S(\mathbf{u})} \mathrm{curl}\mathbf{v}\cdot d\mathbf{S} &= \iint_{S(\mathbf{u})}\Big(\frac{\partial}{\partial t}+\pounds_{\mathbf{u}}\Big)(\mathrm{curl}\mathbf{v}\cdot d\mathbf{S}) \\ &= \iint_{S(\mathbf{u})}\Big(\partial_t\,\boldsymbol{\omega}-\mathrm{curl}\,(\mathbf{u}\times\boldsymbol{\omega})\Big)\cdot d\mathbf{S}=0\,,\end{aligned}\tag{10.17}$$

where the surface $S(\mathbf{u})$ is bounded by an arbitrary circuit $\partial S = c(\mathbf{u})$ moving with the fluid.

10.2 Steady solutions: Lamb surfaces

According to Theorem 10.4, Euler's fluid equations (10.3) imply that

$$\frac{\partial\omega}{\partial t} = -\,[\,u,\,\omega\,]\,.\tag{10.18}$$

Consequently, the vector fields u, ω in *steady* Euler flows, which satisfy $\partial_t\omega = 0$, also satisfy the condition necessary for the Frobenius theorem to hold[1] – namely, that their Jacobi-Lie bracket vanishes. That is, in smooth steady, or equilibrium, solutions of Euler's fluid equations, the flows of the two divergenceless vector fields u and ω commute with each other and lie on a surface in three dimensions.

A sufficient condition for this commutation relation is that the **Lamb vector** $\boldsymbol{\ell} := -\,\mathbf{u}\times\mathrm{curl}\mathbf{v}$ in (10.5) satisfies

$$\boldsymbol{\ell} := -\,\mathbf{u}\times\mathrm{curl}\mathbf{v} = \nabla H(\mathbf{x})\,,\tag{10.19}$$

for some smooth function $H(\mathbf{x})$. This condition means that the flows of vector fields u and $\mathrm{curl}v$ (which are *steady* flows of the Euler equations) are both confined to the same surface $H(\mathbf{x}) = const$. Such a surface is called a **Lamb surface**.

The vectors of velocity $(\mathbf{u})$ and total vorticity $(\mathrm{curl}\mathbf{v})$ for a steady Euler flow are both perpendicular to the normal vector to the Lamb surface along $\nabla H(\mathbf{x})$. That is, the Lamb surface is invariant under the flows of both vector fields, *viz*

$$\pounds_u H = \mathbf{u}\cdot\nabla H = 0 \quad\text{and}\quad \pounds_{\mathrm{curl}v}H = \mathrm{curl}\mathbf{v}\cdot\nabla H = 0\,.\tag{10.20}$$

The Lamb surface condition (10.19) has the following coordinate-free representation [HaMe1998].

[1] For a precise statement and proof of the Frobenius Theorem with applications to differential geometry, see [La1999].

Theorem

10.6 (Lamb surface condition [HaMe1998])**.**
The Lamb surface condition (10.19) is equivalent to the following double substitution of vector fields into the volume form,

$$dH = u \lrcorner \operatorname{curl} v \lrcorner d^3x\,. \tag{10.21}$$

Proof. Recall that the contraction of vector fields with forms yields the following useful formula for the surface element:

$$\nabla \lrcorner d^3x = d\mathbf{S}\,. \tag{10.22}$$

Then using results from previous exercises in vector calculus operations one finds by direct computation that

$$\begin{aligned} u \lrcorner \operatorname{curl} v \lrcorner d^3x &= u \lrcorner (\operatorname{curl}\mathbf{v}\cdot d\mathbf{S}) \\ &= -\big(\mathbf{u}\times \operatorname{curl}\mathbf{v}\big)\cdot d\mathbf{x} \\ &= \nabla H \cdot d\mathbf{x} \\ &= dH\,. \end{aligned} \tag{10.23}$$

□

Remark

10.7. *Formula (10.23)*

$$u \lrcorner (\operatorname{curl}\mathbf{v}\cdot d\mathbf{S}) = dH$$

is to be compared with

$$X_h \lrcorner \omega = dH\,,$$

in the definition of a Hamiltonian vector field in equation (8.4) of Theorem 8.6. Likewise, the stationary case of the Helmholtz vorticity equation (10.15), namely,

$$\pounds_{\mathbf{u}}(\operatorname{curl}\mathbf{v}\cdot d\mathbf{S}) = 0\,. \tag{10.24}$$

is to be compared with the proof of Poincaré's theorem in Proposition 9.37

$$\pounds_{X_h}\omega = d(X_h \lrcorner \omega) = d^2H = 0\,.$$

Thus, the 2-form $\operatorname{curl}\mathbf{v}\cdot d\mathbf{S}$ *plays the same role for stationary Euler fluid flows as the symplectic form* $dq\wedge dp$ *plays for canonical Hamiltonian flows.*

Definition

10.8. *The **Clebsch representation** of the 1-form* $\mathbf{v}\cdot d\mathbf{x}$ *is defined by*

$$\mathbf{v}\cdot d\mathbf{x} = -\,\Pi\, d\Xi + d\Psi\,. \tag{10.25}$$

The functions Ξ, Π *and* Ψ *are called **Clebsch potentials** for the vector* $\mathbf{v}$.[2]

In terms of the Clebsch representation (10.25) of the 1-form $\mathbf{v}\cdot d\mathbf{x}$, the total vorticity flux $\mathrm{curl}\mathbf{v}\cdot d\mathbf{S} = d(\mathbf{v}\cdot d\mathbf{x})$ is the exact 2-form,

$$\mathrm{curl}\mathbf{v}\cdot d\mathbf{S} = d\Xi\wedge d\Pi\,. \tag{10.26}$$

This amounts to writing the flow lines of the vector field of the total vorticity $\mathrm{curl}v$ as the intersections of level sets of surfaces $\Xi = const$ and $\Pi = const$. In other words,

$$\mathrm{curl}\mathbf{v} = \nabla\Xi\times\nabla\Pi\,, \tag{10.27}$$

with the assumption that these level sets foliate $\mathbb{R}^3$. That is, one assumes that any point in $\mathbb{R}^3$ along the flow of the total vorticity vector field $\mathrm{curl}v$ may be assigned to a regular intersection of these level sets. To justify this assumption, we shall refer without attempting a proof to the following theorem.

Theorem

10.9 (Geometry of Lamb surfaces [ArKh1992])**.**
In general, closed Lamb surfaces are tori foliating $\mathbb{R}^3$.

Hence, the symmetry $[\,u,\,\mathrm{curl}v\,] = 0$ that produces the Lamb surfaces for the steady incompressible flow of the vector field u on a three-dimensional manifold $M\in\mathbb{R}^3$ affords a reduction to a family of two-dimensional total vorticity flux surfaces. These surfaces are coordinatised by formula (10.26) and they may be envisioned along with the flow lines of the vector field $\mathrm{curl}v$ in $\mathbb{R}^3$ by using formula (10.27). The main result is the following.

Theorem

10.10 (Lamb surfaces are symplectic manifolds)**.**

[2] The Clebsch representation is another example of a momentum map. For more discussion of this aspect of fluid flows, see [MaWe83, HoMa2004].

The steady flow of the vector field u *satisfying the symmetry relation given by vanishing of the commutator* $[\,u,\, \mathrm{curl} v\,] = 0$ *on a three-dimensional manifold* $M \in \mathbb{R}^3$ *reduces to incompressible flow on a two-dimensional symplectic manifold whose canonically conjugate coordinates* $(\Xi,\, \Pi)$ *are provided by the total vorticity flux*

$$\mathrm{curl} v \,\lrcorner\, d^{\,3}x = \mathrm{curl}\mathbf{v} \cdot d\mathbf{S} = d\Xi \wedge d\Pi \,.$$

The reduced flow is canonically Hamiltonian on this symplectic manifold. Furthermore, the reduced Hamiltonian is precisely the restriction of the invariant H *onto the reduced phase space.*

Proof. Restricting formula (10.23) to coordinates on a total vorticity flux surface (10.26) yields the exterior derivative of the Hamiltonian,

$$\begin{aligned} dH(\Xi,\, \Pi) &= u \,\lrcorner\, (\mathrm{curl}\mathbf{v} \cdot d\mathbf{S}) \\ &= u \,\lrcorner\, (d\Xi \wedge d\Pi) \\ &= (\mathbf{u} \cdot \nabla \Xi)\, d\Pi - (\mathbf{u} \cdot \nabla \Pi)\, d\Xi \\ &=: \frac{d\Xi}{dT}\, d\Pi - \frac{d\Pi}{dT}\, d\Xi \\ &= \frac{\partial H}{\partial \Pi}\, d\Pi + \frac{\partial H}{\partial \Xi}\, d\Xi \,, \end{aligned} \tag{10.28}$$

where $T \in \mathbb{R}$ is the time parameter along the flow lines of the steady vector field u, which carries the Lagrangian fluid parcels. On identifying corresponding terms, the steady flow of the fluid velocity $\mathbf{u}$ is found to obey the canonical Hamiltonian equations,

$$(\mathbf{u} \cdot \nabla \Xi) = \pounds_u \Xi \;=:\; \frac{d\Xi}{dT} = \frac{\partial H}{\partial \Pi} = \big\{\Xi,\, H\big\}\,, \tag{10.29}$$

$$(\mathbf{u} \cdot \nabla \Pi) = \pounds_u \Pi \;=:\; \frac{d\Pi}{dT} = -\,\frac{\partial H}{\partial \Xi} = \big\{\Pi,\, H\big\}\,, \tag{10.30}$$

where $\{\,\cdot\,,\,\cdot\,\}$ is the canonical Poisson bracket for the symplectic form $d\Xi \wedge d\Pi$. □

Corollary

10.11. *The vorticity flux* $d\Xi \wedge d\Pi$ *is invariant under the flow of the velocity vector field* u.

Proof. By (10.28), one verifies

$$\pounds_u (d\Xi \wedge d\Pi) = d\big(u \,\lrcorner\, (d\Xi \wedge d\Pi)\big) = d^2 H = 0\,.$$

This is the standard computation in the proof of Poincaré's theorem in Proposition 9.37 for the preservation of a symplectic form by a canonical transformation. Its interpretation here is that the steady Euler flows preserve the total vorticity flux, $\mathrm{curl}\mathbf{v} \cdot d\mathbf{S} = d\Xi \wedge d\Pi$. □

10.3 Helicity in incompressible fluids

Definition

10.12 (Helicity).
The helicity $\Lambda[\mathrm{curl}\mathbf{v}]$ *of a divergence-free vector field* $\mathrm{curl}\mathbf{v}$ *that is tangent to the boundary* ∂D *of a simply connected domain* $D \in \mathbb{R}^3$ *is defined as*

$$\Lambda[\mathrm{curl}\mathbf{v}] = \int_D \mathbf{v} \cdot \mathrm{curl}\mathbf{v}\, d^{\,3}x \,, \tag{10.31}$$

where $\mathbf{v}$ *is a divergence-free vector-potential for the field* $\mathrm{curl}\mathbf{v}$.

Remark

10.13. *The helicity is unchanged by adding a gradient to the vector* $\mathbf{v}$. *Thus,* $\mathbf{v}$ *is not unique and* $\mathrm{div}\mathbf{v} = 0$ *is not a restriction for simply connected domains in* $\mathbb{R}^3$, *provided* $\mathrm{curl}\mathbf{v}$ *is tangent to the boundary* ∂D.

The helicity of a vector field $\mathrm{curl}\mathbf{v}$ measures the average linking of its field lines, or their relative winding. (For details and mathematical history, see Arnold and Khesin [ArKh1998].) The idea of helicity goes back to Helmholtz [He1858] and Kelvin [Ke1869] in the 19th century. Interest in helicity of fluids was rekindled in magnetohydrodynamics (MHD) by Woltjer [Wo1958] and later in ideal hydrodynamics by Moffatt [Mo1969] who first applied the name *helicity* and emphasised its topological character. Refer to [Mo1981, MoTs1992, ArKh1998] for excellent historical surveys. The principal feature of this concept for fluid dynamics is embodied in the following theorem.

Theorem

10.14 (Euler flows preserve helicity).
When homogeneous or periodic boundary conditions are imposed, Euler's equations for an ideal incompressible fluid flow in a rotating frame with Coriolis parameter $\mathrm{curl}\mathbf{R} = 2\boldsymbol{\Omega}$ *preserves the helicity*

$$\Lambda[\mathrm{curl}\mathbf{v}] = \int_D \mathbf{v} \cdot \mathrm{curl}\mathbf{v}\, d^{\,3}x \,, \tag{10.32}$$

with $\mathbf{v} = \mathbf{u} + \mathbf{R}$, *for which* $\mathbf{u}$ *is the divergenceless fluid velocity* $(\mathrm{div}\mathbf{u} = 0)$ *and* $\mathrm{curl}\mathbf{v} = \mathrm{curl}\mathbf{u} + 2\boldsymbol{\Omega}$ *is the total vorticity.*

Proof. Rewrite the geometric form of the Euler equations (10.12) for rotating incompressible flow with unit mass density in terms of the circulation 1-form $v := \mathbf{v} \cdot d\mathbf{x}$ as

$$(\partial_t + \pounds_u)v = -\,d\left(p + \frac{1}{2}|\mathbf{u}|^2 - \mathbf{u}\cdot\mathbf{v}\right) =: -\,d\,\varpi \,, \tag{10.33}$$

and $\pounds_u\, d^{\,3}x = 0$, where ϖ is an augmented pressure variable,

$$\varpi := p + \frac{1}{2}|\mathbf{u}|^2 - \mathbf{u}\cdot\mathbf{v}\,. \tag{10.34}$$

The fluid velocity vector field is denoted as $u = \mathbf{u}\cdot\nabla$ with $\mathrm{div}\mathbf{u} = 0$. Then the ***helicity density***, defined as

$$v\wedge dv = \mathbf{v}\cdot\mathrm{curl}\mathbf{v}\, d^{\,3}x = \lambda\, d^{\,3}x\,, \quad \text{with} \quad \lambda = \mathbf{v}\cdot\mathrm{curl}\mathbf{v}\,, \tag{10.35}$$

obeys the dynamics it inherits from the Euler equations,

$$\big(\partial_t + \pounds_u\big)(v\wedge dv) = -d\varpi\wedge dv - v\wedge d^2\varpi = -d(\varpi\, dv)\,, \tag{10.36}$$

after using $d^2\varpi = 0$ and $d^2v = 0$. In vector form, this result may be expressed as a conservation law,

$$\big(\partial_t\lambda + \mathrm{div}\lambda\mathbf{u}\big)d^{\,3}x = -\,\mathrm{div}(\varpi\,\mathrm{curl}\mathbf{v})\, d^{\,3}x\,. \tag{10.37}$$

Consequently, the time derivative of the integrated helicity in a domain D obeys

$$\begin{aligned}
\frac{d}{dt}\Lambda[\mathrm{curl}\mathbf{v}] &= \int_D \partial_t\lambda\, d^{\,3}x = -\int_D \mathrm{div}(\lambda\mathbf{u} + \varpi\,\mathrm{curl}\mathbf{v})\, d^{\,3}x \\
&= -\oint_{\partial D} (\lambda\mathbf{u} + \varpi\,\mathrm{curl}\mathbf{v})\,\cdot\, d\mathbf{S}\,,
\end{aligned} \tag{10.38}$$

which vanishes when homogeneous or periodic boundary conditions are imposed on ∂D. □

Remark

10.15. *This result means the* ***helicity integral***

$$\Lambda[\mathrm{curl}\mathbf{v}] = \int_D \lambda\, d^{\,3}x\,,$$

is conserved in periodic domains, or in all of $\mathbb{R}^3$ with vanishing boundary conditions at spatial infinity. However, if either the velocity or total vorticity at the boundary possesses a nonzero normal component, then the boundary is a source of helicity. For a fixed impervious boundary, the normal component of velocity does vanish, but no such condition is imposed on the total vorticity by the physics of fluid flow. Thus, we have the following.

Corollary

10.16. *A flux of total vorticity* curl$\mathbf{v}$ *into the domain is a source of helicity.*

Exercise. Use Cartan's formula (9.14) to compute $\pounds_u(v \wedge dv)$ in equation (10.36). ★

Exercise. Compute the helicity for the 1-form $v = \mathbf{v} \cdot d\mathbf{x}$ in the Clebsch representation (10.25). What does this mean for the linkage of the vortex lines that admit the Clebsch representation? ★

Remark

10.17 (Helicity as Casimir)**.**
The helicity turns out to be a Casimir for the Hamiltonian formulation of the Euler fluid equations [ArKh1998]. Namely, $\{\Lambda,\, H\} = 0$ *for every Hamiltonian functional of the velocity, not just the kinetic energy. The Hamiltonian formulation of ideal fluid dynamics is beyond our present scope. However, the plausibility that the helicity is a Casimir may be confirmed by the following.*

Theorem

10.18 (Diffeomorphisms preserve helicity)**.**
The helicity $\Lambda[\xi]$ *of any divergenceless vector field* ξ *is preserved under the action on* ξ *of any volume-preserving diffeomorphism of the manifold* M *[ArKh1998].*

Remark

10.19 (Helicity is a topological invariant)**.**
The helicity $\Lambda[\xi]$ *is a topological invariant, not a dynamical invariant, because its invariance is independent of which diffeomorphism acts on* ξ*. This means the invariance of helicity is independent of which Hamiltonian flow produces the diffeomorphism. This is the*

hallmark of a Casimir function. Although it is defined above with the help of a metric, every volume-preserving diffeomorphism carries a divergenceless vector field ξ into another such field with the same helicity. However, independently of any metric properties, the action of diffeomorphisms does not create or destroy linkages of the characteristic curves of divergenceless vector fields.

Definition

10.20 (Beltrami flows).
*Equilibrium Euler fluid flows whose velocity and total vorticity are collinear are called **Beltrami flows**.*

Theorem

10.21 (Helicity and Beltrami flows).
Critical points of the conserved sum of fluid kinetic energy and a constant $\varkappa$ times helicity are Beltrami flows of an Euler fluid.

Proof. A critical point of the sum of fluid kinetic energy and a constant $\varkappa$ times helicity satisfies

$$
\begin{aligned}
0 = \delta H_\Lambda &= \int_D \frac{1}{2}|\mathbf{u}|^2 \, d^{\,3}x + \varkappa \int_D \mathbf{v}\cdot \mathrm{curl}\mathbf{v}\, d^{\,3}x \\
&= \int_D \big(\mathbf{u} + 2\varkappa\, \mathrm{curl}\mathbf{v}\big)\cdot \delta\mathbf{u}\, d^{\,3}x\,,
\end{aligned}
$$

after an integration by parts with either homogeneous, or periodic boundary conditions. Vanishing of the integrand for an arbitrary variation in fluid velocity $\delta\mathbf{u}$ implies the Beltrami condition that the velocity and total vorticity are collinear. □

Remark

10.22 (No conclusion about Beltrami stability).
The second variation of H_Λ is given by

$$
\delta^2 H_\Lambda = \int_D |\delta\mathbf{u}|^2 + 2\varkappa\, \delta\mathbf{u}\cdot \mathrm{curl}\, \delta\mathbf{u}\, d^{\,3}x\,.
$$

This second variation is indefinite in sign unless $\varkappa$ vanishes, which corresponds to a trivial motionless fluid equilibrium. Hence, no conclusion is offered by the energy-Casimir method for the stability of a Beltrami flow of an Euler fluid.

10.4 Silberstein-Ertel theorem for potential vorticity

Euler-Boussinesq equations. The Euler-Boussinesq equations for the incompressible motion of an ideal flow of a stratified fluid and velocity $\mathbf{u}$ satisfying $\mathrm{div}\mathbf{u} = 0$ in a rotating frame with Coriolis parameter $\mathrm{curl}\mathbf{R} = 2\mathbf{\Omega}$ are given by

$$\underbrace{\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u}}_{\text{Acceleration}} = \underbrace{-\, gb\nabla z}_{\text{Buoyancy}} + \underbrace{\mathbf{u} \times 2\mathbf{\Omega}}_{\text{Coriolis}} - \underbrace{\nabla p}_{\text{Pressure}} \tag{10.39}$$

where $-g\nabla z$ is the constant downward acceleration of gravity and b is the bouyancy, which satisfies the **advection relation**,

$$\partial_t b + \mathbf{u} \cdot \nabla b = 0\,. \tag{10.40}$$

As for Euler's equations without buoyancy, requiring preservation of the divergence-free (volume preserving) constraint $\nabla \cdot \mathbf{u} = 0$ results in a Poisson equation for pressure p,

$$-\,\Delta\Big(p + \frac{1}{2}|\mathbf{u}|^2\Big) = \mathrm{div}(-\,\mathbf{u} \times \mathrm{curl}\,\mathbf{v}) + g\partial_z b\,, \tag{10.41}$$

which satisfies a Neumann boundary condition because the velocity $\mathbf{u}$ must be tangent to the boundary.

The Newton's Law form of the Euler-Boussinesq equations (10.39) may be rearranged as

$$\partial_t \mathbf{v} -\, \mathbf{u} \times \mathrm{curl}\,\mathbf{v} + gb\nabla z + \nabla\Big(p + \frac{1}{2}|\mathbf{u}|^2\Big) = 0\,, \tag{10.42}$$

where $\mathbf{v} \equiv \mathbf{u} + \mathbf{R}$ and $\nabla \cdot \mathbf{u} = 0$. Geometrically, this is

$$(\partial_t + \pounds_u)v + gbdz + d\varpi = 0\,, \tag{10.43}$$

where ϖ is defined in (10.33). In addition, the buoyancy satisfies

$$(\partial_t + \pounds_u)b = 0\,, \quad \text{with} \quad \pounds_u\, d^3x = 0\,. \tag{10.44}$$

The fluid velocity vector field is denoted as $u = \mathbf{u} \cdot \nabla$ and the circulation 1-form as $v = \mathbf{v} \cdot d\mathbf{x}$. The exterior derivatives of the two equations in (10.43) are written as

$$(\partial_t + \pounds_u)dv = -gdb \wedge dz \quad \text{and} \quad (\partial_t + \pounds_u)db = 0\,. \tag{10.45}$$

Consequently, one finds from the product rule for Lie derivatives (9.19) that

$$(\partial_t + \pounds_u)(dv \wedge db) = 0 \quad \text{or} \quad \partial_t q + \mathbf{u} \cdot \nabla q = 0\,, \tag{10.46}$$

in which the quantity

$$q = \nabla b \cdot \mathrm{curl}\mathbf{v} \tag{10.47}$$

is called **potential vorticity** and is abbreviated as PV. The potential vorticity is an important diagnostic for many processes in geophysical fluid dynamics. Conservation of PV on fluid parcels is called **Ertel's theorem** [Er1942], although it was probably known much earlier, at least by Silberstein, who presented it in his textbook [Si1913].

Remark

10.23 (Silberstein-Ertel theorem).
The constancy of the scalar quantities b and q on fluid parcels implies conservation of the spatially integrated quantity,

$$C_\Phi = \int_D \Phi(b,q)\, d^{\,3}x\,, \tag{10.48}$$

for any smooth function Φ for which the integral exists.

Remark

10.24 (Energy conservation).
In addition to C_Φ, the Euler-Boussinesq fluid equations (10.42) also conserve the ***total energy***

$$E = \int_D \frac{1}{2}|\mathbf{u}|^2 + gbz\;d^{\,3}x\,, \tag{10.49}$$

which is the sum of the kinetic and potential energies. We do not develop the Hamiltonian formulation of the 3D stratified rotating fluid equations here. However, one may imagine that the quantity C_Φ would be its Casimir, as the notation indicates. With this understanding, we shall prove the following.

Theorem

10.25 (Energy-Casimir criteria for equilibria).
Critical points of the conserved sum $E_\Phi = E + C_\Phi$, namely,

$$E_\Phi = \int_D \frac{1}{2}|\mathbf{u}|^2 + gbz\;d^{\,3}x + \int_D \Phi(b,q) + \varkappa q\, d^{\,3}x\,, \tag{10.50}$$

are equilibrium solutions of the Euler-Boussinesq fluid equations in (10.42). The function Φ in the Casimir and the Bernoulli function K in (10.54) for the corresponding fluid equilibrium are related by $q\Phi_q - \Phi = K$.

Proof. The last term in (10.50) was separated out for convenience in dealing with the boundary terms that arise on taking the variation. The variation of E_Φ is given by

$$
\begin{aligned}
\delta E_\Phi \;=\; & \int_D \Big(\mathbf{u}_e - \Phi_{qq}\nabla b_e \times \nabla q_e\Big)\cdot\delta\mathbf{u}\; d^{\,3}x \\
& + \int_D \Big(gz + \Phi_b - \mathrm{curl}\mathbf{v}_e\cdot\nabla\Phi_q\Big)\,\delta b\; d^{\,3}x \\
& + \Big(\Phi_q\Big|_{\partial D} + \varkappa\Big)\oint_{\partial D}\Big(\delta b\,\mathrm{curl}\mathbf{v}_e + b_e\,\mathrm{curl}\,\delta\mathbf{u}\Big)\cdot\hat{\mathbf{n}}\,dS\,,
\end{aligned}
\tag{10.51}
$$

in which the surface terms arise from integrating by parts and $\hat{\mathbf{n}}$ is the outward normal of the domain boundary, ∂D. Here, the partial derivatives Φ_b, Φ_q and Φ_{qq} are evaluated at the critical point b_e, q_e, $\mathbf{v}_e$ and $\Phi_q|_{\partial D}$ takes the critical point values on the boundary. The critical point conditions are obtained by setting $\delta E_\Phi = 0$. These conditions are,

$$
\begin{aligned}
\delta\mathbf{u}: &\qquad \mathbf{u}_e = \Phi_{qq}\nabla b_e\times\nabla q_e\,,\\
\delta b: &\qquad gz + \Phi_b = \mathrm{curl}\mathbf{v}_e\cdot\nabla\Phi_q\,,\\
\partial D: &\qquad \Phi_q\Big|_{\partial D} + \varkappa = 0\,.
\end{aligned}
\tag{10.52}
$$

The critical point condition on the boundary ∂D holds automatically for tangential velocity and plays no further role. The critical point condition for $\delta\mathbf{u}$ satisfies the steady flow conditions,

$$
\mathbf{u}_e\cdot\nabla q_e = 0 = \mathbf{u}_e\cdot\nabla b_e\,.
$$

An important steady flow condition derives from the motion equation (10.42)

$$
\mathbf{u}_e\times\mathrm{curl}\,\mathbf{v}_e = -\,gz\nabla b_e + \nabla K\,,
\tag{10.53}
$$

which summons the ***Bernoulli function***,

$$
K(b_e,q_e) = p_e + \frac{1}{2}|\mathbf{u}_e|^2 + gb_e z\,,
\tag{10.54}
$$

and forces it to be a function of (b_e,q_e). When taken in concert with the previous relation for K, the vector product of ∇b_e with (10.53) yields

$$
\mathbf{u}_e = \frac{1}{q_e}\nabla b_e\times\nabla K(b_e,q_e) = \nabla b_e\times\nabla\Phi_q(b_e,q_e)\,,
\tag{10.55}
$$

where the last relation uses the critical point condition arising from the variations of velocity, $\delta\mathbf{u}$. By equation (10.55), critical points of E_Φ are steady solutions of the Euler-Boussinesq fluid equations (10.42) and the function Φ in the Casimir is related to the Bernoulli function K in (10.54) for the corresponding steady solution by

$$q_e \Phi_{qq}(b_e, q_e) = K_q(b_e, q_e)\,. \tag{10.56}$$

This equation integrates to

$$q_e \Phi_q - \Phi = K + F(b_e)\,. \tag{10.57}$$

The vector product of ∇q_e with the steady flow relation (10.53) yields

$$\Phi_{qq}(b_e, q_e)(\nabla q_e \cdot \mathrm{curl}\mathbf{v}_e) = gz - K_b(b_e, q_e)\,. \tag{10.58}$$

Combining this result with the critical point condition for δb in (10.52) yields

$$q_e \Phi_q - \Phi = K + G(q_e)\,. \tag{10.59}$$

Subtracting the two relations (10.59) and (10.56) eliminates the integration functions F and G, and establishes

$$q_e \Phi_q(b_e, q_e) - \Phi(b_e, q_e) = K(b_e, q_e)\,, \tag{10.60}$$

as the relation between critical points of E_Φ and equilibrium solutions of the Euler-Boussinesq fluid equations. $\square$

Remark

10.26.
The energy-Casimir stability method was implemented for Euler-Boussinesq fluid equilibria in [AbHoMaRa1986]. See also [HoMaRaWe1985] for additional examples.

11 Hodge star operator on $\mathbb{R}^3$

Definition

11.1. *The* ***Hodge star operator*** *establishes a linear correspondence between the space of k-forms and the space of $(3-k)$-forms, for $k = 0, 1, 2, 3$. This correspondence may be defined by its usage:*

$$
\begin{aligned}
*1 &:= d^3x = dx^1 \wedge dx^2 \wedge dx^3 \,, \\
*d\mathbf{x} &:= d\mathbf{S} \,, \\
(*dx^1, *dx^2, *dx^3) &:= (dS_1, dS_2, dS_3) \,, \\
&:= (dx^2 \wedge dx^3,\, dx^3 \wedge dx^1,\, dx^1 \wedge dx^2) \,, \\
*d\mathbf{S} &= d\mathbf{x} \,, \\
(*dS_1, *dS_2, *dS_3) &:= (dx^1, dx^2, dx^3) \,, \\
*d^3x &= 1 \,,
\end{aligned}
$$

in which each formula admits cyclic permutations of the set $\{1, 2, 3\}$.

Remark

11.2. *Note that $**\,\alpha = \alpha$ for these k-forms.*

Definition

11.3 (L^2 inner product of forms)**.**
The Hodge star induces an ***inner product*** $(\,\cdot\,,\,\cdot\,) : \Lambda^k(M) \times \Lambda^k(M) \to \mathbb{R}$ *on the space of k-forms. Given two k-forms α and β defined on a smooth manifold M, one defines their L^2 inner product as*

$$
(\alpha,\, \beta) := \int_M \alpha \wedge *\beta = \int_M \langle \alpha, \beta \rangle \; d^3x \,, \tag{11.1}
$$

where d^3x is the volume form. The main examples of the inner product are for $k = 0, 1$. These are given by the L^2 pairings,

$$
\begin{aligned}
(f,\, g) = \int_M f \wedge *g &:= \int_M fg \, d^3x \,, \\
(\mathbf{u} \cdot d\mathbf{x},\, \mathbf{v} \cdot d\mathbf{x}) = \int_M \mathbf{u} \cdot d\mathbf{x} \wedge *(\mathbf{v} \cdot d\mathbf{x}) &:= \int_M \mathbf{u} \cdot \mathbf{v} \, d^3x \,.
\end{aligned}
$$

Exercise. Show that combining the Hodge star operator with the exterior derivative yields the following vector calculus operations:

$$
\begin{aligned}
\,d(\mathbf{v}\cdot d\mathbf{x}) &= \operatorname{div}\mathbf{v}\,,\\
*\,d(\mathbf{v}\cdot d\mathbf{x}) &= (\operatorname{curl}\mathbf{v})\cdot d\mathbf{x}\,,\\
d*d*(\mathbf{v}\cdot d\mathbf{x}) &= (\nabla\operatorname{div}\mathbf{v})\cdot d\mathbf{x}\,,\\
*\,d*d(\mathbf{v}\cdot d\mathbf{x}) &= \operatorname{curl}(\operatorname{curl}\mathbf{v})\cdot d\mathbf{x}\,.
\end{aligned}
$$

★

The Hodge star on manifolds is used to define the **codifferential**.

Definition

11.4. *The* ***codifferential****, denoted as δ, is defined for a k-form $\alpha\in\Lambda^k$ as*

$$
\delta\alpha = (-1)^{k+1+k(3-k)} * d * \alpha\,. \tag{11.2}
$$

Note that the sign is positive for $k=1$ and negative for $k=2$.

Exercise. Verify that $\delta^2=0$. ★

Remark

11.5. *Introducing the notation δ for codifferential* ***cannot cause any confusion*** *with other familiar uses of the same notation, for example, to denote Kronecker delta, or the variational derivative delta. All these standard usages of the notation (δ) are easily recognised in their individual contexts.*

Proposition

11.6. *The codifferential is the adjoint of the exterior derivative, in that*

$$(\delta\alpha,\, \beta) = (\alpha,\, d\beta)\,. \tag{11.3}$$

Exercise. Verify that the codifferential is the adjoint of the exterior derivative by using the definition of the Hodge star inner product.
(Hint: Why may one use $\int_M d(*\alpha \wedge \beta) = 0$ when integrating by parts?) ★

Definition

11.7. *The* ***Laplace-Beltrami operator*** *on smooth functions is defined to be* $\nabla^2 = \operatorname{div}\operatorname{grad} = \delta\, d$*. Thus, one finds,*

$$\nabla^2 f = \delta\, df = *\, d * df\,, \tag{11.4}$$

for a smooth function f*.*

Definition

11.8. *The* ***Laplace-deRham operator*** *is defined by*

$$\Delta := d\delta + \delta d\,. \tag{11.5}$$

Exercise. Show that the **Laplace-deRham operator** on a 1-form $\mathbf{v}\cdot d\mathbf{x}$ expresses the **Laplacian of a vector**,

$$(d\delta + \delta d)(\mathbf{v}\cdot d\mathbf{x}) = (\nabla \operatorname{div}\mathbf{v} - \operatorname{curl}\operatorname{curl}\mathbf{v})\cdot d\mathbf{x} =: (\Delta\mathbf{v})\cdot d\mathbf{x}\,.$$

Use this expression to define the inverse of the curl operator applied to a divergenceless vector function as

$$\operatorname{curl}^{-1}\mathbf{v} = \operatorname{curl}(-\Delta^{-1}\mathbf{v}) \quad \text{when} \quad \operatorname{div}\mathbf{v} = 0. \tag{11.6}$$

This is the **Biot-Savart Law** often used in electo-magnetism and incompressible fluid dynamics. ★

Remark

11.9. *Identifying this formula for $\Delta\mathbf{v}$ as the* ***vector Laplacian*** *on a differentiable manifold agrees with the definition of the Laplacian of a vector in any curvilinear coordinates.*

Exercise. Compute the components of the Laplace-deRham operator $\Delta\mathbf{v}$ for a 1-form $\mathbf{v}\cdot d\mathbf{x}$ defined on a sphere of radius R. How does this differ from the Laplace-Beltrami operator ($\nabla^2\mathbf{v} = \mathrm{div}\,\mathrm{grad}\mathbf{v}$) in spherical curvilinear coordinates? ★

Exercise. Show that the Laplace-deRham operator $-\Delta := d\delta + \delta d$ is symmetric with respect to the Hodge star inner product, that is,

$$(\Delta\alpha,\, \beta) = (\alpha,\, \Delta\beta)\,.$$

★

Exercise. In coordinates, symmetry of Δ with respect to the Hodge star inner product is expressed as

$$\begin{aligned}\int -\Delta\alpha\cdot\beta d^{\,3}x &= \int(-\nabla\mathrm{div}\alpha + \mathrm{curlcurl}\alpha)\cdot\beta d^{\,3}x\\ &= \int(\mathrm{div}\alpha\cdot\mathrm{div}\beta + \mathrm{curl}\alpha\cdot\mathrm{curl}\beta d^{\,3}x\,.\end{aligned}$$

Conclude that the Laplace-deRham operator $-\Delta$ is non-negative, by setting $\alpha=\beta$.

★

Exercise. Use formula (11.2) for the definition of codifferential $\delta = *d*$ to express in vector notation,

$$\delta(\pounds_u v) = -\delta dp - g\delta(bdz)\,,$$

for the 1-form $v = \mathbf{v} \cdot d\mathbf{x}$, vector field $u = \mathbf{u} \cdot \nabla$, functions p, b and constant g. How does this expression differ from the Poisson equation for pressure p in (10.41)? ★

12 Poincaré's Lemma: Closed vs exact differential forms

Definition

12.1 (Closed and exact differential forms)**.**
*A k-form α is **closed** if $d\alpha = 0$.*
*The k-form is **exact** if there exists a $(k-1)$-form β for which $\alpha = d\beta$.*

Definition

12.2 (Co-closed and co-exact differential forms)**.** *A k-form α is **co-closed** if $\delta\alpha = 0$ and is **co-exact** if there exists a $(3-k)$-form β for which $\alpha = \delta\beta$.*

Proposition

12.3. *Exact and co-exact forms are **orthogonal** with respect to the L^2 inner product on $\Lambda^k(M)$.*

Proof. Let $\alpha = \delta\beta$ be a co-exact form and let $\zeta = d\eta$ be an exact form. Their L^2 inner product defined in (11.1) is computed as

$$(\alpha,\, \zeta) = (\delta\beta,\, d\eta) = (\beta,\, d^2\eta) = 0\,.$$

This vanishes, because δ is dual to d, that is, $(\delta\beta,\, \zeta) = (\beta,\, d\zeta)$ by Proposition 11.6. □

Remark

12.4. *Not all closed forms are globally exact on a given manifold M.*

Example

12.5 (Helicity example).
As an example, the one-form

$$v = f dg + \psi d\phi$$

for smooth functions f, g, ψ, ϕ on $\mathbb{R}^3$ may be used to create the closed three-form (helicity)

$$v \wedge dv = (\psi df - f d\psi) \wedge dg \wedge d\phi \in \mathbb{R}^3 \,.$$

This three-form is closed because it is a "top form" in $\mathbb{R}^3$. However, it is exact only when the combination $\psi df - f d\psi$ is exact, and this fails whenever ψ and f are functionally independent. Thus, ***some closed forms are not exact.***

However, it turns out that all closed forms may be shown to be ***locally exact****. This is the content of the following Lemma.*

Definition

12.6 (Locally exact differential forms).
A closed differential form α that satisfies $d\alpha = 0$ on a manifold M is ***locally exact****, when a neighbourhood exists around each point in M in which $\alpha = d\beta$.*

Lemma

12.7 (Poincaré's lemma).
Any closed form on a manifold M is locally exact.

Remark

12.8. *Rather than give the standard proof appearing in most texts in this subject, let us illustrate Poincaré's Lemma in an example, then use it to contrast the closed versus exact properties.*

Example

12.9. *In the example of **helicity** above, the one-form $v = fdg + \psi d\phi$ may always be written locally as $v = fdg + cd\phi$ in a neighbourhood defined on a level surface $\psi = c$. In that neighbourhood, $v \wedge dv = c\,(df \wedge dg \wedge d\phi)$, which is exact because c is a constant.*

Remark

12.10. *Closed forms that are not globally exact have **topological content**. For example, the spatial integral of the three-form $v \wedge dv \in \mathbb{R}^3$ is the **degree-of-mapping formula** for the Hopf map $S^3 \mapsto S^2$. It also measures the number of **self-linkages** (also known as **helicity**) of the divergenceless vector field associated with the two-form dv. See [ArKh1998] for in-depth discussions of the topological content of differential forms that are closed, but only locally exact, in the context of geometric mechanics.*

Example

12.11 (A locally closed and exact two-form in $\mathbb{R}^3$). *The transformation in $\mathbb{R}^3$ from 3D Cartesian coordinates (x, y, z) to spherical coordinates (r, θ, ϕ) is given by*

$$(x, y, z) = (r \sin\theta \cos\phi, r \sin\theta \sin\phi, r \cos\theta)\,.$$

As is well known, the volume form transforms into spherical coordinates as

$$\mathrm{dVol} = d^{\,3}x = dx \wedge dy \wedge dz = r^2 dr \wedge d\phi \wedge d\cos\theta\,.$$

Exercise. *Compute the transformation in the previous equation explicitly.* ★

*In general, contraction of a vector field into a volume form produces a two-form $X \mathbin{\lrcorner} d^{\,3}x = \mathbf{X} \cdot \hat{\mathbf{n}}\, dS$, where dS is the surface area element with unit normal vector $\hat{\mathbf{n}}$. Consider the two-form $\beta \in \Lambda^2$ obtained by substituting the **radial vector field**,*

$$X = \mathbf{x} \cdot \nabla = x\partial_x + y\partial_y + z\partial_z = r\partial_r\,,$$

into the volume form dVol. *This may be computed in various ways,*

$$
\begin{aligned}
\beta &= X \,\lrcorner\, d^3x = \mathbf{x}\cdot\hat{\mathbf{n}}\,dS \\
&= xdy\wedge dz + ydz\wedge dx + zdx\wedge dy \\
&= \frac{1}{2}\epsilon_{abc}x^a\,dx^b\wedge dx^c \\
&= r\partial_r \,\lrcorner\, r^2dr\wedge d\phi\wedge d\cos\theta = r^3d\phi\wedge d\cos\theta\,.
\end{aligned}
$$

One computes the exterior derivative

$$
\begin{aligned}
d\beta &= d(X \,\lrcorner\, d^3x) = d(\mathbf{x}\cdot\hat{\mathbf{n}}\,dS) = \operatorname{div}\mathbf{x}\,d^3x \\
&= 3\,d^3x = 3\,r^2dr\wedge d\phi\wedge d\cos\theta \neq 0\,.
\end{aligned}
$$

So the two-form β is not closed. Hence it cannot be exact. When evaluated on the spherical level surface $r=1$ (which is normal to the radial vector field X) the 2-form β becomes the area element on the sphere.

Remark

12.12.

- *The 2-form β in the previous example is closed on the level surface $r=1$, but it is not exact everywhere. This is because singularities occur at the poles, where the coordinate ϕ is not defined.*
- *This is an example of Poincaré's theorem, in which a differential form is closed, but is only locally exact.*
- *If β were exact on $r=1$, its integral $\int_{S^2}\beta$ would give zero for the area of the unit sphere instead of 4π!*

Example

12.13. *Instead of the radial vector field, let us choose an arbitrary three-dimensional vector field $\mathbf{n}(\mathbf{x})$ in which $\mathbf{n}:\mathbb{R}^3\to\mathbb{R}^3$. As for the radial vector field, we may compute the two-form,*

$$
\begin{aligned}
\beta = X \,\lrcorner\, d^3n := \mathbf{n}\cdot\frac{\partial}{\partial\mathbf{n}} \,\lrcorner\, d^3n &= \frac{1}{2}\epsilon_{abc}\,n^a\,dn^b\wedge dn^c \\
&= \frac{1}{2}\epsilon_{abc}\,n^a\nabla n^b\times\nabla n^c\cdot d\mathbf{S}(\mathbf{x})\,.
\end{aligned}
$$

One computes the exterior derivative once again,

$$d\beta = d(X \intprod d^3x) = \operatorname{div} \mathbf{n}\, d^3n = \det[\nabla \mathbf{n}]\, d^3x\,.$$

Now suppose $\mathbf{n}$ *is a unit vector, satisfying the relation* $|\mathbf{n}(\mathbf{x})|^2 = 1$*. Then* $\mathbf{n} : \mathbb{R}^3 \to S^2$ *and the rows of its Jacobian will be functionally dependent, so the determinant* $\det[\nabla \mathbf{n}]$ *must vanish.*

Consequently, $d\beta$ *vanishes and the two-form* β *is closed. In this case, Poincaré's Lemma states that a one-form* α *exists locally such that the closed two-form* β *satisfies* $\beta = d\alpha$*. In fact, the unit vector in spherical coordinates,*

$$\mathbf{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)\,,$$

does produce a closed two-form β *that is expressible as an exterior derivative,*

$$\beta = d\phi \wedge d\cos\theta = d(\phi\, d\cos\theta)\,. \tag{12.1}$$

However, we know from the previous example that β *could only be locally exact. The obstructions to being globally exact are indicated by the singularities of the polar coordinate representation, in which the azimuthal angle* ϕ *is undefined at the North and South poles of the unit sphere.*

Remark

12.14 (Hopf fibration)**.**
These considerations introduce the ***Hopf map*** *in which the unit vector* $\mathbf{n}(\mathbf{x})$ *maps* $\mathbf{x} \in S^3$ *to the spherical surface* S^2 *given by* $|\mathbf{n}(\mathbf{x})|^2 = 1$ *locally as* $S^3 \simeq S^2 \times S^1$*. Such a direct-product map that holds locally, but does not hold globally, is called a* ***fibration****. Here* S^2 *is called the* ***base space*** *and* S^1 *is called the* ***fibre****. The integral* $\int_{S^2} \beta$ *is called the* ***degree of mapping*** *of the* ***Hopf fibration****. This integral is related to the self-linkage or* ***helicity*** *discussed earlier in this section. For more details, see [ArKh1998, Fl1963, Is1999, Ur2003].*

13 Euler's equations in Maxwell form

Exercise. (Maxwell form of Euler's fluid equations)
Show that by making the following identifications

$$\begin{aligned} \mathbf{B} &:= \boldsymbol{\omega} + \operatorname{curl}\mathbf{A}_0 \\ \mathbf{E} &:= \boldsymbol{\ell} + \nabla\big(p + \tfrac{1}{2}|\mathbf{u}|^2\big) + \big(\nabla\phi - \partial_t\mathbf{A}_0\big) \\ \mathbf{D} &:= \boldsymbol{\ell} \\ \mathbf{H} &:= \nabla\psi\,, \end{aligned} \tag{13.1}$$

the Euler fluid equations (10.3) and (10.6) imply the **Maxwell form**

$$\begin{aligned} \partial_t\mathbf{B} &= -\operatorname{curl}\mathbf{E} \\ \partial_t\mathbf{D} &= \operatorname{curl}\mathbf{H} + \mathbf{J} \\ \operatorname{div}\mathbf{B} &= 0 \\ \operatorname{div}\mathbf{E} &= 0 \\ \operatorname{div}\mathbf{D} &= \rho \;=\; -\Delta\big(p + \tfrac{1}{2}|\mathbf{u}|^2\big) \\ \mathbf{J} &= \mathbf{E}\times\mathbf{B} + (\operatorname{curl}^{-1}\mathbf{E})\times\operatorname{curl}\mathbf{B}\,, \end{aligned} \tag{13.2}$$

provided the (smooth) gauge functions ϕ and $\mathbf{A}_0$ satisfy $\Delta\phi - \partial_t\mathrm{div}\mathbf{A}_0 = 0$ with $\partial_n\phi = \hat{\mathbf{n}}\cdot\partial_t\mathbf{A}_0$ at the boundary and ψ may be arbitrary, because $\operatorname{curl}\mathbf{H} = 0$ removes $\mathbf{H}$ from further consideration in the dynamics. ★

Remark

13.1.

- The first term in the current density $\mathbf{J}$ in the Maxwell form of Euler's fluid equations in (13.2) is reminiscent of the ***Poynting vector*** in electromagnetism [BoWo1965]. The second term in $\mathbf{J}$ contains the inverse of the curl operator acting on the divergenceless vector function $\mathbf{E}$. This inverse-curl operation may be defined via the Laplace-DeRham theory that leads to the Biot-Savart Law (11.6).

- The divergence of the **D**-equation in the Maxwell form (13.2) of the Euler fluid equations implies a conservation equation, given by

$$\partial_t \rho = \operatorname{div} \mathbf{J}\,. \tag{13.3}$$

Thus, the total "charge" $\int \rho\, d^3x$ is conserved, provided the current density $\mathbf{J}$ (or, equivalently, the partial time derivative of the Lamb vector) has no normal component at the boundary.

- The conservation equation (13.3) for $\rho = \operatorname{div}\boldsymbol{\ell}$ is potentially interesting in applications. For example, it may be interesting to use the divergence of the Lamb vector as a diagnostic quantity in turbulence experiments.

- The equation for the curl of the Lamb vector is of course also easily accessible, if needed.

14 Euler's equations in Hodge-star form in $\mathbb{R}^4$

Definition

14.1. *The **Hodge star operator** on $\mathbb{R}^4$ establishes a linear correspondence between the space of k-forms and the space of $(4-k)$-forms, for $k = 0, 1, 2, 3, 4$. This correspondence may be defined by its usage:*

$$\begin{aligned}
*1 \;&:=\; d^4x = dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3\,,\\
(*dx^1, *dx^2, *dx^3) \;&:=\; (dS_1 \wedge dx^0, -dS_2 \wedge dx^0, dS_3 \wedge dx^0)\,,\\
&:= (dx^2 \wedge dx^3 \wedge dx^0,\, dx^3 \wedge dx^0 \wedge dx^1,\, dx^0 \wedge dx^1 \wedge dx^2)\,,\\
(*dS_1, *dS_2, *dS_3) \;&:=\; \big(*(dx^2 \wedge dx^3), *(dx^3 \wedge dx^1), *(dx^1 \wedge dx^2)\big)\,,\\
&=\; (dx^0 \wedge dx^1, dx^0 \wedge dx^2, dx^3 \wedge dx^0)\,,\\
*d^3x \;&=\; *(dx^1 \wedge dx^2 \wedge dx^3) = dx^0\,,\\
*d^4x \;&=\; 1\,,
\end{aligned}$$

in which each formula admits cyclic permutations of the set $\{0, 1, 2, 3\}$.

Remark

14.2. *Note that $**\,\alpha = \alpha$ for these k-forms.*

Exercise. Prove that

$$*(dx^\mu \wedge dx^\nu) = \tfrac{1}{2}\varepsilon_{\mu\nu\sigma\gamma} dx^\sigma \wedge dx^\gamma \,,$$

where $\varepsilon_{\mu\nu\sigma\gamma} = +1$ (resp. -1) when $\{\mu\nu\sigma\gamma\}$ is an even (resp. odd) permutation of the set $\{0,1,2,3\}$ and it vanishes if any of its indices are repeated. ★

Exercise. Introduce the $\mathbb{R}^4$-vectors for fluid velocity and vorticity with components $u_\mu = (1, \mathbf{u})$ and $\omega_\nu = (0, \boldsymbol{\omega})$. Let $dx^0 = dt$ and prove that

$$F = *u_\mu \omega_\nu dx^\mu \wedge dx^\nu = \boldsymbol{\ell} \cdot d\mathbf{x} \wedge dt + \boldsymbol{\omega} \cdot d\mathbf{S} \,. \tag{14.1}$$

★

Exercise. Show that Euler's fluid equations (10.3) imply

$$F = d\Big(\mathbf{v} \cdot d\mathbf{x} - (p + \tfrac{1}{2}u^2) dt\Big) \,, \tag{14.2}$$

in the Euler fluid notation of equation (10.4). ★

After the preparation of having solved these exercises, it is an easy computation to show that the Helmholtz vorticity equation (10.15) follows from the compatibility condition for F. Namely,

$$0 = dF = \Big(\partial_t \,\boldsymbol{\omega} - \mathrm{curl}\,(\mathbf{u} \times \boldsymbol{\omega})\Big) \cdot d\mathbf{S} \wedge dt + \mathrm{div}\,\boldsymbol{\omega}\, d^3x \,.$$

This Hodge-star version of the Helmholtz vorticity equation brings us a step closer to understanding the electromagnetic analogy in the Maxwell form of Euler's fluid equations (13.2). This is because Faraday's Law in Maxwell's equations has a similar formulation, but for 4-vectors in Minkowski space-time instead of $\mathbb{R}^4$ [Fl1963]. The same concepts from the calculus of differential forms still apply, but with the Minkowski metric.

Next, introduce the 2-form in $\mathbb{R}^4$

$$G = \boldsymbol{\ell} \cdot d\mathbf{S} + d\chi \wedge dt, \tag{14.3}$$

representing the flux of the Lamb vector though a *fixed* spatial surface element $d\mathbf{S}$. Two more brief computations recover the other formulas in the Maxwell representation of fluid dynamics in (13.2). First, the exterior derivative of G, given by

$$\begin{aligned} dG &= \partial_t \boldsymbol{\ell} \cdot d\mathbf{S} \wedge dt + \operatorname{div} \boldsymbol{\ell}\, d^3x & (14.4)\\ &= \mathbf{J} \cdot d\mathbf{S} \wedge dt + \rho\, d^3x =: \boldsymbol{J}\,, & (14.5)\end{aligned}$$

recovers the two relations $\partial_t \boldsymbol{\ell} = \mathbf{J}$ and $\operatorname{div} \boldsymbol{\ell} = \rho$ in (13.2). Here $\boldsymbol{J}$ is the current density 3-form with components $(\rho, \mathbf{J})$. The second calculation we need is the compatibility condition for G, namely

$$d^2G = \Big(\operatorname{div} \mathbf{J} - \partial_t \rho\Big) d^3x \wedge dt = 0\,.$$

This recovers the conservation law in (13.3) for the Maxwell form of Euler's fluid equations.

Thus, the differential-form representation of Euler's fluid equations in $\mathbb{R}^4$ reduces to two elegant relations,

$$dF = 0 \quad \text{and} \quad dG = \boldsymbol{J}\,, \tag{14.6}$$

where the 2-forms F, G and the 3-form $\boldsymbol{J}$ are given in (14.1), (14.3) and (14.5), respectively.

Exercise. Show that equations (14.6) for the differential representation of Euler's fluid equations in $\mathbb{R}^4$ may be written as a pair of partial differential equations,

$$\partial_\mu F^{\mu\nu} = 0 \quad \text{and} \quad \partial_\mu G^{\mu\nu} = J^\nu, \tag{14.7}$$

written in terms of the $\mathbb{R}^4$-vector $J^\nu = (-\mathbf{J}, \rho)^T$ and the 4×4 antisymmetric tensors $F^{\mu\nu} = u^\mu \omega^\nu - u^\nu \omega^\mu$. In matrix form $F^{\mu\nu}$ is given by

$$F^{\mu\nu} = \begin{bmatrix} 0 & \ell_3 & -\ell_2 & \omega_1 \\ -\ell_3 & 0 & \ell_1 & \omega_2 \\ \ell_2 & -\ell_1 & 0 & \omega_3 \\ -\omega_1 & -\omega_2 & -\omega_3 & 0 \end{bmatrix},$$

and $G^{\mu\nu}$ is given by

$$G^{\mu\nu} = \begin{bmatrix} 0 & \chi_{,3} & -\chi_{,2} & \ell_1 \\ -\chi_{,3} & 0 & \chi_{,1} & \ell_2 \\ \chi_{,2} & -\chi_{,1} & 0 & \ell_3 \\ -\ell_1 & -\ell_2 & -\ell_3 & 0 \end{bmatrix},$$

where $\mu, \nu = 1, 2, 3, 4$, with notation $\partial_\mu = \partial/\partial x^\mu$ with $x^\mu = (\mathbf{x}, t)^T$, $u^\mu = (\mathbf{u}, 1)^T$ and $\omega^\mu = (\boldsymbol{\omega}, 0)^T$. Also, $\partial_\mu u^\mu \equiv u^\mu_{,\mu} = \nabla \cdot \mathbf{u} = 0$ and $\omega^\mu_{,\mu} = \nabla \cdot \boldsymbol{\omega} = 0$. ★

Exercise. Write Maxwell's equations for the propagation of electromagnetic waves in Hodge-star form (14.6) in Minkowski space. Discuss the role of the Minkowski metric in defining Hodge-star and the effects of curl$\mathbf{H} \neq 0$ on the solutions, in comparison to the treatment of Euler's fluid equations in $\mathbb{R}^4$ presented here. ★

References

[AbHoMaRa1986] Abarbanel, H. D. I., Holm, D. D., Marsden, J. E. and Ratiu, T. S. [1986] Nonlinear stability analysis of stratified ideal fluid equilibria. *Phil. Trans. R. Soc. Lond. A* **318**, 349–409.

[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.

[ArKh1992] Arnold, V. I. and Khesin, B. A. [1992] Topological methods in hydrodynamics. *Annu. Rev. Fluid Mech.* **24**, 145–166.

[ArKh1998] Arnold, V. I and Khesin, B. A. [1998] *Topological Methods in Hydrodynamics.* Springer: New York.

[BoWo1965] Born, M. and Wolf, E. [1965] *Principles of Optics*. Oxford: Pergamon Press, p. 31.

[Bl1965] Bloembergen, N. [1965] *Nonlinear Optics*. New York: Benjamin.

[BoMa2009] Bou-Rabee, N. and Marsden, J. E. [2009]

[BoWo1965] Born, M. and Wolf, E. [1965] *Principles of Optics*. Oxford: Pergamon Press, p. 31.

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

[CuBa1997] Cushman, R. H. and Bates, L. M. [1997] *Global Aspects of Integrable Systems*. Basel: Birkhäuser.

[Di1933] Dirac, P. A. M. [1933] The Lagrangian in quantum mechanics. *Physikalische Zeitschrift der Sowjetunion* **3**, 64–72.

[Di1981] Dirac, P. A. M. [1981] *The Principles of Quantum Mechanics*, 4th revised ed. Oxford: Oxford University Press.

[EfMoSa2005] Efstathiou, K., Morrison, R. and Sadovskií, D. [2005] No polar coordinates: Lectures by Richard Cushman. 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.

[Er1942] Ertel, H. [1942] Ein neuer hydrodynamischer Wirbelsatz. *Met. Z., Braunschweig* **59**, 277–281.

[Fe1948] Feynman, R. P. [1948] Space-time approach to non-relativistic quantum mechanics. *Rev. Mod. Phys.* **20**, 367–387.

[FeHi1965] Feynman, R. P. and Hibbs, A. R. [1965] *Quantum Mechanics and Path Integrals*. New York: McGraw-Hill.

[FeVeHe1957] Feynman, R., Vernon, F. and Hellwarth, R. [1957] Geometrical representation of the Schrödinger equation for solving maser problems. *J. Appl. Phys.* **28**, 49–52.

[Fl1963] Flanders, H. [1963] *Differential Forms with Applications to the Physical Sciences*. New York: Academic Press.

[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.

[HaMe1998] Haller, G. and Mezic, I. [1998] Reduction of three-dimensional, volume-preserving flows with symmetry. *Nonlinearity* **11**, 319–339.

[He1858] Helmholtz, H. [1858] On integrals of the hydrodynamical equations which express vortex motion. Transl. Tait, P. G. [1867] *Philos. Mag.* **33**, 485–512.

[Ho2002] Holm, D. D. [2002] Euler–Poincaré dynamics of perfect complex fluids. In *Geometry, Mechanics, and Dynamics: In Honor of the 60th Birthday of Jerrold E. Marsden*, edited by P. Newton, P. Holmes and A. Weinstein. New York: Springer, pp. 113–167.

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

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

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

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

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

[HoMa2004] Holm, D. D. and Marsden, J. E. [2004] Momentum maps and measure valued solutions (peakons, filaments, and sheets) the EPDiff equation. In *The Breadth of Symplectic and Poisson Geometry: A Festshrift in Honor of Alan Weinstein* edited by J.E. Marsden and T.S. Ratiu. Boston: Birkhäuser. pp. 203–235. http://arxiv.org/abs/nlin.CD/0312048.

[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.

[HoMaRaWe1985] Holm, D. D., Marsden, J. E., Ratiu, T. S. and Weinstein, A. [1985] Nonlinear stability of fluid and plasma equilibria. *Phys. Rep.* **123**, 1–116.

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

[Is1999] Isham, C. J. [1999] *Differential Geometry for Physicists*. World Scientific Lecture Notes in Physics, Vol. 61. Singapore: World Scientific.

[Ke1869] Kelvin, L. [1869] On vortex motion. *Trans. Roy. Soc. Edin.* **25**, 217–260; Maximum and minimum energy in vortex motion. In *Mathematical and Physical Papers*, Vol. 4. Cambridge: Cambridge University Press, pp. 172–183.

[KnHj2001] Knudsen, J. M. and Hjorth, P. G. [2001] *Elements of Newtonian Mechanics: Including Nonlinear Dynamics*, 3rd ed. New York: Springer-Verlag.

[KoSc2001] Kosmann-Schwarzbach, Y. [2011] *The Noether Theorems: Invariance and Conservation Laws in the Twentieth Century*. Translated by B. L. Schwarzbach, New York: Springer-Verlag.

[La1999] Lang, S. [1999] *Fundamentals of Differential Geometry*, Graduate Texts in Mathematics, Vol. 191. New York: Springer-Verlag.

[Le2003] Lee, J. [2003] *Introduction to Smooth Manifolds*. New York: Springer-Verlag.

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

[MaWe83] Marsden, J. E. and Weinstein, A. [1983] Coadjoint orbits, vortices and Clebsch variables for incompressible fluids. *Physica D.* **7**, 305–323.

[Mo1969] Moffatt, H. K. [1969] The degree of knottedness of tangled vortex lines. *J. Fluid. Mech.* **106**, 117–129.

[Mo1981] Moffatt, H. K. [1981] Some developments in the theory of turbulence. *J. Fluid Mech.* **106**, 27–47.

[MoTs1992] Moffatt, H. K. and Tsinober, A. [1992] Helicity in laminar and turbulent flow. *Annu. Rev. Fluid Mech.* **24**, 281–312.

[Na1973] Nambu, Y. [1973] Generalized Hamiltonian mechanics, *Phys. Rev. D* **7**, 2405–2412.

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

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

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

[Sh1984] Shen, Y. R. [1984] *The Principles of Nonlinear Optics*. New York: Wiley-Interscience.

[Si1913] Silberstein, L. [1913] *Vectorial Mechanics*. London: Macmillan and Co.

[St1852] Stokes, G. G. [1852] On the composition and resolution of streams of polarized light from different sources.
Trans. Camb. Phil. Soc. **9**, 399–423.

[Ur2003] Urbantke, H. K. [2003] The Hopf fibration – seven times in physics. *J. Geom. Phys.* **46**, 125–150.

[Wo1958] Woltjer, L. [1958] A theorem on force-free magnetic fields. *Proc. Natl. Acad. Sci. USA* **44**, 489–491.

Index