Geometric Mechanics
Part I: Dynamics and Symmetry

Darryl D Holm
Mathematics Department
Imperial College London

2nd Edition

To Justine, for her love, kindness and patience.
Thanks for letting me think this was important.

Contents

Preface

Introduction to the 2^{nd} edition

This is the text for a course in geometric mechanics taught by the author for undergraduates in their third year of mathematics at Imperial College London. The course has now been taught five times, to different groups of students with quite diverse interests and preparation. After each class, the students were requested to turn in a response sheet on which they answered two questions.

These two questions were:

1. "What was this class about?" and

2. "What question about the class would *you* like to see pursued further in the course?"

The answers to these questions helped the instructor keep the lectures aligned with the interests and understanding of the students, and it enfranchised the students because they themselves selected the topics for many of the lectures. Responding to this written dialogue with the class required the instructor to develop a flexible approach that would allow a certain amount of shifting from one topic to another, in response to the associations arising in the minds of the students, but without losing sight of the fundamental concepts. This flexibility seemed useful in keeping the students engaged and meeting their diverse needs. It also showed the students the breadth, power and versatility of geometric mechanics.

In an attempt to meet this need for flexibility in teaching, the introduction of each chapter of the 2^{nd} edition has been rewritten to start at a rather elementary level that does not assume proficiency with the previous material. This means the chapters are sufficiently self-contained that the instructor may select the material of most value to an individual class.

A brief history of geometric mechanics

The ideas underlying geometric mechanics first emerged in the principles of optics formulated by Galileo, DesCartes, Fermat and Huygens. These underlying ideas were developed in optics and particle mechanics by Newton, Euler, Lagrange and Hamilton, with added contributions from Gauss, Poisson, Jacobi, Riemann, Maxwell and Lie, for example, then later by Poincaré, Noether, Cartan and others. In many of these contributions, optics and mechanics held equal sway.

Fermat's principle (that the light ray passing from one point to another in an optical medium takes the path of stationary optical length) is complementary to Huygens principle (that a later wave front emerges as the envelope of wavelets emitted from the present wave front). Both principles are only models of reality, but they are models in the best sense. Both are transcendent fabrications that intuited the results of a later more fundamental principle (Maxwell's equations) and gave accurate predictions at the level of physical perception in their time. Without being the full truth by being physically tenable themselves, they fulfilled the tasks for which they were developed and they laid the foundations for more fundamental theories. Light rays do not exist and points along a light wave do not emit light. However, both principles work quite well in the design of optical instruments! In addition, both principles are still interesting now as the mathematical *definitions* of rays and wave fronts, respectively, although neither one fully represents the physical principles of optics.

The duality between tangents to stationary paths (Fermat) and normals to wave fronts (Huygens) in classical optics corresponds in geometric mechanics to the duality between velocities and momenta. This duality between ray paths and wave fronts may remind us of the duality between complementary descriptions of particles and waves in quantum mechanics. The bridge from the wave description to the ray description is crossed in the geometric-optical high-wavenumber limit ($k \to \infty$). The bridge from quantum mechanics to classical mechanics is crossed in another type of geometric-optical limit ($\hbar \to 0$) as Planck's constant tends to zero. In this course we arrive at the threshold of the bridge to quantum mechanics when we write the Bloch equations for the MASER and the two-level qubit of quantum computing in Chapter 4, for example, and the Maxwell-Bloch equations for self-induced transparency in Chapter 6. Although we do not cross over this bridge, its presence reminds us that the conceptual unity in the historical developments of geometrical optics and classical mechanics is still of interest today. Indeed, Hamilton's formulations of optics and mechanics were guiding lights in the development of the quantum mechanics of atoms and molecules, especially the Hamilton-Jacobi equation discussed in Chapter 3, and the quantum version of the Hamiltonian approach is still used today in scientific research on the interactions of photons, electrons and beyond.

Building on the earlier work by Hamilton and Lie, in a series of famous studies during the 1890s, Poincaré laid the geometric foundations for the modern approach to classical mechanics. One study by Poincaré addressed the propagation of polarised optical beams. For us, Poincaré's representation of the oscillating polarisation states of light as points on a sphere turns out to inform the geometric mechanics of nonlinearly coupled resonant oscillators. Following Poincaré, we shall represent the dynamics of coupled resonant oscillators as flows along curves on manifolds whose points are resonantly oscillating motions. Such orbit manifolds are *fibrations* (local factorisations) of larger spaces.

Lie symmetry is the perfect method for applying Poincaré's geometric approach. In mechanics, a Lie symmetry is an invariance of the Lagrangian or Hamiltonian under a Lie group; that is, under a group of transformations that depend smoothly on a set of parameters. The effectiveness of Lie symmetry in mechanics is seen on the Lagrangian side in Noether's theorem. Noether's theorem states that each Lie symmetry of the Lagrangian in Hamilton's principle of least action implies a conservation law. On the Hamiltonian side, Noether's theorem states that each Lie symmetry of the Hamiltonian summons a momentum map, which maps the canonical phase space to the dual of the Lie symmetry algebra. When the Lie group is a symmetry of the Hamiltonian, the momentum map is conserved and the dynamics is confined to its level sets. Of course, there is much more geometry in this idea than the simple restriction of the dynamics to a level set. In particular, restriction to level sets of momentum maps culminates in the reduction of phase space to manifolds whose points are orbits of symmetries. In some cases, this geometric approach to the separation of motions may produce complete integrability of the original problem.

The language of Lie groups, especially Lie derivatives, is needed to take advantage of Poincaré's geometrical framework for mechanics. The text also provides an introduction to exterior differential calculus; so that the student will have the language to go further in geometric mechanics. The lessons here are only the first steps – the road to geometric mechanics is long and scenic, even beautiful, for those who may take it. It leads from finite to infinite dimensions. Anyone taking this road will need these basic tools and the language of Lie symmetries, in order to interpret the concepts that will be met along the way.

The approach of the text

The text surveys a small section of the road to geometric mechanics, by treating several examples in classical mechanics, all in the same geometric framework. These example problems include:

- Fermat's principle for ray optics to travelling waves propagating by self-induced transparency (the Maxwell-Bloch equations);
- Bifurcations in the behaviour of resonant oscillators and polarised travelling wave pulses in optical fibres;

- The bead sliding on a rotating hoop, the spherical pendulum and the elastic spherical pendulum. The approximate solution of the elastic spherical pendulum via a phase-averaged Lagrangian shares concepts with molecular oscillations of CO_2 and with 2nd harmonic generation in nonlinear laser optics;
- Divergenceless vector fields and stationary patterns of fluid flow on invariant surfaces.

In each case, the results of the geometric analysis eventually reduce to divergence free flow in $\mathbb{R}^3$ along intersections of level surfaces of constants of the motion. On these level surfaces, the motion is symplectic, as guaranteed by the Marsden-Weinstein theorem [MaWe74].

How to read this book

The book is organised into six chapters and two appendices.

Chapter 1 treats Fermat's principle for ray optics in refractive media as a detailed example that lays out the strategy of Lie symmetry reduction in geometric mechanics that will be applied in the remainder of the text.

Chapter 2 summarises the contributions of Newton, Lagrange and Hamilton to geometric mechanics. The key examples are the rigid body and the spherical pendulum.

Chapter 3 discusses Lie symmetry reduction in the language of the exterior calculus of differential forms. The main example is ideal incompressible fluid flow.

Chapters 4, 5 and 6 illustrate the ideas laid out in the previous chapters through a variety of applications. In particular, the strategy of Lie symmetry reduction laid out in Chapter 1 in the example of ray optics is applied to resonant oscillator dynamics in Chapter 4, then to the elastic spherical pendulum in Chapter 5 and finally to a special case of the Maxwell-Bloch equations for laser excitation of matter in Chapter 6.

The two Appendices contain a compendium of example problems which may be used as topics for homework and enhanced coursework.

The first chapter treats Fermat's principle for ray optics as an example that lays out the strategy of ***Lie symmetry reduction*** for all of the other applications of geometric mechanics discussed in the course. This strategy begins by deriving the dynamical equations from a variational principle and Legendre transforming from the Lagrangian to the Hamiltonian formulation. Then the implications of Lie symmetries are considered. For example, when the medium is symmetric under rotations about the axis of optical propagation, the corresponding symmetry of the Hamiltonian for ray optics yields a conserved quantity called *skewness*, which was first discovered by Lagrange.

The second step in the strategy of Lie symmetry reduction is to transform to invariant variables. This transformation is called the ***quotient map***. In particular, the quotient map for ray optics is obtained by writing the Hamiltonian as a function of the axisymmetric invariants constructed from the optical phase space variables. This construction has the effect of quotienting out the angular dependence, as an alternative to the polar coordinate representation. The transformation to axisymmetric variables in the quotient map takes the four dimensional optical phase space into the three dimensional real space $\mathbb{R}^3$. The image of the quotient map may be represented as the zero level set of a function of the invariant variables. This zero level set is called the ***orbit manifold***, although in some cases it may have singular points. For ray optics, the orbit manifold is a two dimensional level set of the conserved skewness whose value depends on the initial conditions, and each point on it represents a circle in phase space corresponding to the orbit of the axial rotations.

In the third step, the canonical Poisson brackets of the axisymmetric invariants with the phase space coordinates produce ***Hamiltonian vector fields***, whose flows on optical phase space yield the diagonal action of the symplectic Lie group $Sp(2,\mathbb{R})$ on the optical position and momentum. The Poisson brackets of the axisymmetric invariants *close among themselves* as linear functions of these invariants, thereby yielding a ***Lie-Poisson bracket*** dual to the symplectic Lie algebra $sp(2, \mathbb{R})$, represented as divergenceless vector fields on $\mathbb{R}^3$. The Lie-Poisson bracket reveals the geometry of the solution behaviour in axisymmetric ray optics as flows along the intersections of the level sets of the Hamiltonian and the orbit manifold in the $\mathbb{R}^3$ space of axisymmetric invariants. This is ***coadjoint motion***.

In the final step, the angle variable is reconstructed. This angle turns out to be the sum of two parts: one part is called *dynamic*, because it depends on the Hamiltonian. The other part is called *geometric* and is equal to the area enclosed by the solution on the orbit manifold.

The geometric-mechanics treatment of Fermat's principle by Lie symmetry reduction identifies two ***momentum maps*** admitted by axisymmetric ray optics. The first is the map from optical phase space (position and momentum of a ray on an image screen) to their associated area on the screen. This phase-space area is Lagrange's invariant in axisymmetric ray optics; it takes the same value on each image screen along the optical axis. The second momentum map transforms from optical phase space to the bilinear axisymmetric invariants by means of the quotient map. Because this transformation is a momentum map, the quotient map yields a valid Lie-Poisson bracket among the bilinear axisymmetric invariants. The evolution then proceeds by coadjoint motion on the orbit manifold.

Chapter 2 treats the geometry of rigid-body motion from the viewpoints of Newton, Lagrange and Hamilton, respectively. This is the classical problem of geometric mechanics, which makes a natural counterpoint to the treatment in Chapter 1 of ray optics by Fermat's principle. Chapter 2 also treats the Lie symmetry reduction of the spherical pendulum. The treatments of the rigid body and the spherical pendulum by these more familiar approaches also sets the stage for the introduction of the flows of Hamiltonian vector fields and their Lie-derivative actions on differential forms in Chapter 3.

The problem of a single, polarised, optical laser pulse propagating as a travelling wave in an anisotropic, cubically nonlinear, lossless medium is investigated in Chapter 4. This is a Hamiltonian system in $\mathbb{C}^2$ for the dynamics of two complex oscillator modes (the two polarisations). Since the two polarisations of a single optical pulse must have the same natural frequency, they are in $1:1$ resonance. An S^1 phase invariance of the Hamiltonian for the interaction of the optical pulse with the optical medium in which it propagates will reduce the phase space to the Poincaré sphere, S^2, on which the problem is completely integrable. In Chapter 4, the fixed points and bifurcations sequences of the phase portrait of this system on S^2 are studied as the beam intensity and medium parameters are varied. The corresponding Lie-symmetry reductions for the $n:m$ resonances is also discussed in detail.

Chapter 5 treats the swinging spring, or elastic spherical pendulum, from the viewpoint of Lie-symmetry reduction. In this case, averaging the Lagrangian for the system over its rapid elastic oscillations introduces the additional symmetry needed to reduce the problem to an integrable Hamiltonian system. This reduction results in the three-wave surfaces in $\mathbb{R}^3$ and thereby sets up the framework for predicting the characteristic feature of the elastic spherical pendulum, which is the step-wise precession of its swing plane.

Chapter 6 treats the Maxwell-Bloch laser-matter equations for self-induced transparency. The Maxwell-Bloch equations arise from a variational principle obtained by averaging the Lagrangian for the Maxwell-Schrödinger equations. As for the swinging spring, averaging the Lagrangian introduces the Lie symmetry needed for reducing the dimensions of the dynamics and thereby making it more tractable. The various Lie-symmetry reductions of the real Maxwell-Bloch equations to two dimensional orbit manifolds are discussed and their corresponding geometric phases are determined in Chapter 6.

Exercises are sprinkled liberally throughout the text, often with hints or even brief explicit solutions. These are indented and marked with ★ and ▲, respectively. Moreover, the careful reader will find that many of the exercises are answered in passing somewhere later in the text in a more developed context.

> Key theorems, results and remarks are placed into frames (like this one).

Appendix A contains additional worked problems in geometric mechanics. These problems include a bead sliding on a rotating hoop, the spherical pendulum in polar coordinates, a charged particle in a magnetic field and the Kepler problem. Appendix A also contains descriptions of the dynamics of a rigid body coupled to a flywheel, complex phase space, a single harmonic oscillator, two resonant coupled oscillators and a cubically nonlinear oscillator, all treated from the viewpoint of geometric mechanics. Finally, Appendix B contains ten potential homework problems whose solutions are not given, but which we hope may challenge the serious student.

Acknowledgements. I am enormously grateful to my students, who have been partners in this endeavour, for their comments, questions and inspiration. I am also grateful to many friends and collaborators for their camaraderie in first forging these tools. I am especially grateful to Peter Lynch, who shared my delight and obsession over the step-wise shifts of the swing plane of the elastic spherical pendulum discussed in Chapter 5. I am also grateful to Tony Bloch, Dorje Brody, Alex Dragt, Poul Hjorth (who drew many of the figures), Daniel Hook (who drew the Kummer orbit manifolds), Jerry Marsden, Peter Olver, Tudor Ratiu, Tanya Schmah, Cristina Stoica and Bernardo Wolf, who all gave their help, comments, encouragement and astute recommendations about the work underlying the text.

Chapter 1

Fermat's ray optics

1.1 Fermat's principle

Fermat's principle (1662) states that *the path between two points taken by a ray of light leaves the optical length stationary under variations in a family of nearby paths.*

Fermat

This principle accurately describes the properties of light that is reflected by mirrors, refracted at a boundary between different media, or transmitted through a medium with a continuously varying index of refraction. Fermat's principle *defines* a light ray and provides an example that will guide us in recognising the principles of geometric mechanics.

The ***optical length*** of a path $\mathbf{r}(s)$ taken by a ray of light in passing from a point A to a point B in three-dimensional space is defined by

$$\mathsf{A} := \int_A^B n(\mathbf{r}(s))\, ds\,, \tag{1.1.1}$$

where $n(\mathbf{r})$ is the index of refraction at the spatial point $\mathbf{r} \in \mathbb{R}^3$ and

$$ds^2 = d\mathbf{r}(s) \cdot d\mathbf{r}(s) \tag{1.1.2}$$

is the element of arc length ds along the ray path $\mathbf{r}(s)$ through that point.

Definition 1.1.1 (Fermat's principle for ray paths)
The path $\mathbf{r}(s)$ taken by a ray of light passing from A to B in three-dimensional space leaves the optical length ***stationary*** *under variations in a family of nearby paths $\mathbf{r}(s, \varepsilon)$ depending smoothly on a parameter ε. That is, the path $\mathbf{r}(s)$ satisfies*

$$\delta \mathsf{A} = 0 \quad with \quad \delta \mathsf{A} := \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \int_A^B n(\mathbf{r}(s, \varepsilon))\, ds\,, \tag{1.1.3}$$

where the deviations of the ray path $\mathbf{r}(s, \varepsilon)$ from $\mathbf{r}(s)$ are assumed to vanish when $\varepsilon = 0$, and to leave its endpoints fixed.

Fermat's principle of stationary ray paths is dual to Huygens' principle (1678) of constructive interference of waves. According to Huygens, among all possible paths from an object to an image, the waves corresponding to the stationary path contribute most to the image because of their constructive interference. Both principles are models that approximate more fundamental physical results derived from Maxwell's equations.

The Fermat and Huygens principles for geometric optics are also foundational ideas in mechanics. Indeed, the founders of mechanics, Newton, Lagrange and Hamilton all worked seriously in optics as well. We start with Fermat and Huygens, whose works preceded Newton's *Principia Mathematica* by twenty five years, Lagrange's *Méchanique Analytique* by more than a century and Hamilton's papers *On a General Method in Dynamics*, by more than 150 years.

After briefly discussing the geometric ideas underlying the principles of Fermat and Huygens and using them to derive and interpret the eikonal equation for ray optics, this chapter shows that Fermat's principle naturally introduces Hamilton's principle for the Euler-Lagrange equations, as well as the concepts of phase space, Hamiltonian formulation, Poisson brackets, Hamiltonian vector fields, symplectic transformations and momentum maps arising from reduction by symmetry.

In his time, Fermat discovered the geometric foundations of ray optics. This is our focus in the chapter.

1.1.1 Three-dimensional eikonal equation

Stationary paths Consider the possible ray paths in Fermat's principle leading from a point A to a point B in 3D space as belonging to a family of C^2 curves $\mathbf{r}(s, \varepsilon) \in \mathbb{R}^3$ depending smoothly on a real parameter ε in an interval that includes $\varepsilon = 0$. This ε-family of paths $\mathbf{r}(s, \varepsilon)$ defines a set of smooth transformations of the ray path $\mathbf{r}(s)$. These transformations are taken to satisfy

$$\mathbf{r}(s, 0) = \mathbf{r}(s)\,, \quad \mathbf{r}(s_A, \varepsilon) = \mathbf{r}(s_A)\,, \quad \mathbf{r}(s_B, \varepsilon) = \mathbf{r}(s_B)\,. \tag{1.1.4}$$

That is, $\varepsilon = 0$ is the identity transformation of the ray path and its endpoints are left fixed. An infinitesimal variation of the path $\mathbf{r}(s)$ is denoted by $\delta\mathbf{r}(s)$, defined by the ***variational derivative***,

$$\delta\mathbf{r}(s) := \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \mathbf{r}(s, \varepsilon)\,, \tag{1.1.5}$$

where the fixed endpoint conditions in (1.1.4) imply $\delta\mathbf{r}(s_A) = 0 = \delta\mathbf{r}(s_B)$.

With these definitions, Fermat's principle in Definition 1.1.1 implies the fundamental equation for the ray paths, as follows.

Theorem 1.1.2 (Fermat's principle and the eikonal equation)

Stationarity of the ***optical length****, or* ***action*** A*, under variations of the ray paths*

$$0 = \delta\mathsf{A} = \delta \int_A^B n(\mathbf{r}(s)) \sqrt{\frac{d\mathbf{r}}{ds} \cdot \frac{d\mathbf{r}}{ds}}\, ds\,, \tag{1.1.6}$$

defined using arclength parameter s, satisfying $ds^2 = d\mathbf{r}(s)\cdot d\mathbf{r}(s)$ and $|\dot{\mathbf{r}}| = 1$, implies the equation for the ray path $\mathbf{r}\in\mathbb{R}^3$ as

$$\frac{d}{ds}\left(n(\mathbf{r})\frac{d\mathbf{r}}{ds}\right) = \frac{\partial n}{\partial \mathbf{r}}\,. \tag{1.1.7}$$

*In ray optics, this is called the **eikonal equation**.*

1

Remark 1.1.3 (Invariance under reparameterisation)
The integral in (1.1.8) is invariant under reparameterisation of the ray path. In particular, it is invariant under transforming $\mathbf{r}(s)\to\mathbf{r}(\tau)$ from arc-length ds to optical length $d\tau = n(\mathbf{r}(s))ds$. That is,

$$\int_A^B n(\mathbf{r}(s))\sqrt{\frac{d\mathbf{r}}{ds}\cdot\frac{d\mathbf{r}}{ds}}\,ds = \int_A^B n(\mathbf{r}(\tau))\sqrt{\frac{d\mathbf{r}}{d\tau}\cdot\frac{d\mathbf{r}}{d\tau}}\,d\tau\,, \tag{1.1.8}$$

in which $|d\mathbf{r}/d\tau|^2 = n^{-2}(\mathbf{r}(\tau))$.

We now prove Theorem 1.1.2.
Proof. The equation for determining the ray path that satisfies the stationary condition (1.1.8) may be computed by introducing the ε-family of paths into the ***action*** A, then differentiating it with respect to ε under the integral sign, setting $\varepsilon = 0$ and integrating by parts with respect to s, as follows

$$\begin{aligned}
0 &= \delta\int_A^B n(\mathbf{r}(s))\sqrt{\dot{\mathbf{r}}\cdot\dot{\mathbf{r}}}\,ds \\
&= \int_A^B \left[|\dot{\mathbf{r}}|\frac{\partial n}{\partial \mathbf{r}}\cdot\delta\mathbf{r} + \left(n(\mathbf{r}(s))\frac{\dot{\mathbf{r}}}{|\dot{\mathbf{r}}|}\right)\cdot\delta\dot{\mathbf{r}}\right]ds \\
&= \int_A^B \left[|\dot{\mathbf{r}}|\frac{\partial n}{\partial \mathbf{r}} - \frac{d}{ds}\left(n(\mathbf{r}(s))\frac{\dot{\mathbf{r}}}{|\dot{\mathbf{r}}|}\right)\right]\cdot\delta\mathbf{r}\,ds\,,
\end{aligned}$$

where we have exchanged the order of the derivatives in s and ε, and used the homogeneous endpoint conditions in (1.1.4). We

[1]The term ***eikonal*** (from the Greek $\epsilon\iota\kappa o\nu\alpha$ meaning image) was introduced into optics in [Br1895].

choose the arclength variable $ds^2 = d\mathbf{r} \cdot d\mathbf{r}$, so that $|\dot{\mathbf{r}}| = 1$ for $\dot{\mathbf{r}} := d\mathbf{r}/ds$. (This means that $d|\dot{\mathbf{r}}|/ds = 0$.) Consequently, the 3D eikonal equation (1.1.7) emerges for the ray path $\mathbf{r} \in \mathbb{R}^3$. ■

Remark 1.1.4 *From the viewpoint of historical contributions in classical mechanics, the eikonal equation (1.1.7) is the 3D*

$$\frac{d}{ds}\left(\frac{\partial L}{\partial \dot{\mathbf{r}}}\right) = \frac{\partial L}{\partial \mathbf{r}} \,. \tag{1.1.9}$$

arising from ***Hamilton's principle***

$$\delta \mathsf{A} = 0 \quad \textit{with} \quad \mathsf{A} = \int_A^B L((\mathbf{r}(s), \dot{\mathbf{r}}(s))\, ds$$

for the ***Lagrangian*** *function*

$$L((\mathbf{r}(s), \dot{\mathbf{r}}(s)) = n(\mathbf{r}(s))\, \sqrt{\dot{\mathbf{r}} \cdot \dot{\mathbf{r}}} \,, \tag{1.1.10}$$

in Euclidean space.

Exercise. Verify that the *same* three-dimensional eikonal equation (1.1.7) *also* follows from Fermat's principle in the form

$$0 = \delta \mathsf{A} = \delta \int_A^B \frac{1}{2} n^2(\mathbf{r}(\tau))\, \frac{d\mathbf{r}}{d\tau} \cdot \frac{d\mathbf{r}}{d\tau}\, d\tau \,, \tag{1.1.11}$$

after transforming to a new arc-length parameter τ given by $d\tau = nds$. ★

Answer. Denoting $\mathbf{r}'(\tau) = d\mathbf{r}/d\tau$ one computes,

$$0 = \delta \mathsf{A} = \int_A^B \frac{ds}{d\tau} \left[\frac{nds}{d\tau} \frac{\partial n}{\partial \mathbf{r}} - \frac{d}{ds}\left(\frac{nds}{d\tau} n \frac{d\mathbf{r}}{ds} \right) \right] \cdot \delta \mathbf{r}\, d\tau \,, \tag{1.1.12}$$

which agrees with the previous calculation upon reparameterising $d\tau = nds$. ▲

Remark 1.1.5 (Finsler geometry and singular Lagrangians)
The Lagrangian function in (1.1.10) for the 3D eikonal equation (1.1.7)

$$L(\mathbf{r},\dot{\mathbf{r}}) = n(\mathbf{r})\,\sqrt{\delta_{ij}\dot{r}^i\dot{r}^j}\,, \tag{1.1.13}$$

is homogeneous of degree 1 in $\dot{\mathbf{r}}$. That is, $L(\mathbf{r},\lambda\dot{\mathbf{r}}) = \lambda L(\mathbf{r},\dot{\mathbf{r}})$ for any $\lambda > 0$. Homogeneous functions of degree 1 satisfy ***Euler's relation***,

$$\dot{\mathbf{r}}\cdot\frac{\partial L}{\partial\dot{\mathbf{r}}} - L = 0\,. \tag{1.1.14}$$

Hence, Fermat's principle may be written as stationarity of the integral

$$\mathsf{A} = \int_A^B n(\mathbf{r}(s))\,ds = \int_A^B \mathbf{p}\cdot d\mathbf{r} \quad \text{for the quantity} \quad \mathbf{p} := \frac{\partial L}{\partial\dot{\mathbf{r}}}\,.$$

The quantity $\mathbf{p}$ defined here will be interpreted later as the ***canonical momentum*** for ray optics.

Taking another derivative of Euler's relation (1.1.14) yields

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

so the Hessian of the Lagrangian L with respect to the tangent vectors is singular (has zero determinant). A singular Lagrangian might become problematic in some situations. However, there is a simple way of obtaining a regular Lagrangian whose ray paths are the same as those for the singular Lagrangian arising in Fermat's principle.

The Lagrangian function in the integrand of action (1.1.11) is related to Fermat Lagrangian in (1.1.13) by

$$\tfrac12 n^2(\mathbf{r})\,\delta_{ij}\dot{r}^i\dot{r}^j = \tfrac12 L^2$$

Computing the Hessian with respect to the tangent vector yields the Riemannian metric for the regular Lagrangian in (1.1.11),

$$\frac12\frac{\partial^2 L^2}{\partial\dot{r}^i\partial\dot{r}^j} = n^2(\mathbf{r})\,\delta_{ij}\,.$$

The emergence of a Riemannian metric from the Hessian of the square of a homogenous function of degree 1 is the hallmark of

Finsler geometry, of which Riemannian geometry is a special case. Finsler geometry, however, is beyond our present scope. For more discussion of the ideas underlying Finsler geometry, see, e.g., [Ru1959, ChChLa1999].

In the present case, the variational principle for the regular Lagrangian in (1.1.11) leads to the same three-dimensional eikonal equation as that arising from Fermat's principle in (1.1.8), but parameterised by optical length, rather than arc-length. This will be sufficient for the purposes of studying ray trajectories because in geometric optics one is only concerned with the trajectories of the the ray paths in space, not their parameterisation.

Remark 1.1.6 (Newton's Law form of the eikonal equation) *Reparameterising the ray path in terms of a variable $d\sigma = n^{-1}ds$ transforms the eikonal equation (1.1.7) into the form of* ***Newton's Law***

$$\frac{d^2\mathbf{r}}{d\sigma^2} = \frac{1}{2}\frac{\partial n^2}{\partial \mathbf{r}}\,. \tag{1.1.15}$$

Thus, in terms of the parameter σ ray trajectories are governed by Newtonian dynamics. Interestingly, thus equation has a conserved ***energy integral****,*

$$E = \frac{1}{2}\left|\frac{d\mathbf{r}}{d\sigma}\right|^2 - \frac{1}{2}n^2(\mathbf{r})\,. \tag{1.1.16}$$

> **Exercise.** Propose various forms of the squared index of refraction, e.g., cylindrically or spherically symmetric, then solve the eikonal equation in Newtonian form (3.2.17) for the ray paths and interpret the results as optical devices, e.g., lenses.
>
> What choices of the index of refraction lead to *closed ray paths*? ★

1.1.2 Three-dimensional Huygens wave fronts

Ray vector Fermat's ray optics is complementary to Huygens' wavelets. According to the Huygens wavelet assumption, a level set of the wave front, $S(\mathbf{r})$, moves along the ***ray vector***, $\mathbf{n}(\mathbf{r})$, so that its incremental change over a distance $d\mathbf{r}$ along the ray is given by

$$\nabla S(\mathbf{r}) \cdot d\mathbf{r} = \mathbf{n}(\mathbf{r}) \cdot d\mathbf{r} = n(\mathbf{r})\, ds\,. \tag{1.1.17}$$

The geometric relationship between wave fronts and ray paths is illustrated in Figure 1.1.

Theorem 1.1.7 (Huygens-Fermat complementarity)
Fermat's eikonal equation (1.1.7) follows from the Huygens wavelet equation (1.1.17)

$$\nabla S(\mathbf{r}) = n(\mathbf{r})\frac{d\mathbf{r}}{ds} \qquad \textbf{\textit{(Huygens equation)}} \tag{1.1.18}$$

by differentiating along the ray path.

Corollary 1.1.8 *The wave front level sets $S(\mathbf{r}) = constant$ and the ray paths $\mathbf{r}(s)$ are mutually orthogonal.*

Proof. The corollary follows once equation (3.2.20) is proved, because $\nabla S(\mathbf{r})$ is along the ray vector and is perpendicular to the level set of $S(\mathbf{r})$. ■

Proof. Theorem 1.1.7 may be proved by a direct calculation applying the operation

$$\frac{d}{ds} = \frac{d\mathbf{r}}{ds} \cdot \nabla = \frac{1}{n}\nabla S \cdot \nabla$$

to Huygens equation (3.2.20). This yields the eikonal equation (1.1.7), by the following reasoning:

$$\frac{d}{ds}\left(n\frac{d\mathbf{r}}{ds}\right) = \frac{1}{n}\nabla S \cdot \nabla(\nabla S) = \frac{1}{2n}\nabla|\nabla S|^2 = \frac{1}{2n}\nabla n^2 = \nabla n\,.$$

In this chain of equations, the first step substitutes

$$d/ds = n^{-1}\nabla S \cdot \nabla\,.$$

The second step exchanges order of derivatives. The third step uses the modulus of the Huygens equation (3.2.20) and invokes the property that $|d\mathbf{r}/ds|^2 = 1$. ■

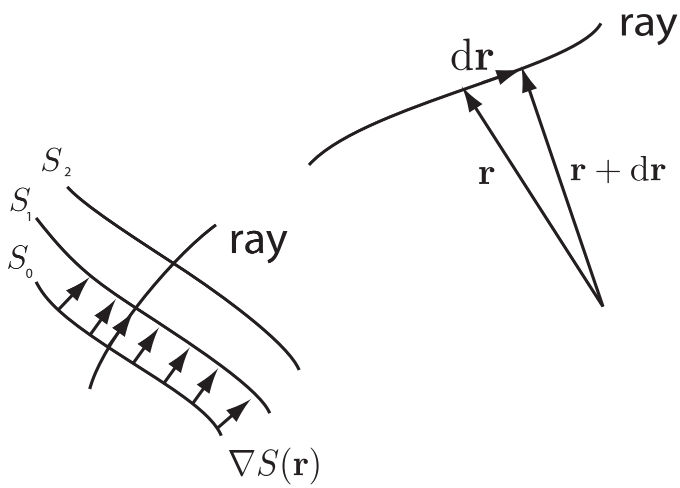


Figure 1.1: Huygens wave front and one of its corresponding ray paths. The wave fronts and ray paths form mutually orthogonal families of curves. The gradient $\nabla S(\mathbf{r})$ is normal to the wave front and tangent to the ray through it at the point $\mathbf{r}$.

Corollary 1.1.9 *The modulus of the Huygens equation (3.2.20) yields*

$$|\nabla S|^2(\mathbf{r}) = n^2(\mathbf{r}) \qquad \textit{\textbf{(scalar eikonal equation)}} \qquad (1.1.19)$$

which follows because $d\mathbf{r}/ds = \hat{\mathbf{s}}$ *in equation (1.1.7) is a unit vector.*

Remark 1.1.10 (The Hamilton-Jacobi equation)
Corollary 1.1.9 arises as an algebraic result in the present considerations. However, it also follows at a more fundamental level from Maxwell's equations for electrodynamics in the slowly varying amplitude approximation of geometric optics, cf. [BoWo1965], Chapter 3. See Keller [Ke1962] for the modern extension of geometric optics to include diffraction. The scalar eikonal equation (1.1.19) is also known as the steady ***Hamilton-Jacobi equation****. The time dependent Hamilton-Jacobi equation (3.2.11) is discussed in Chapter 3.*

Theorem 1.1.11 (Ibn Sahl-Snell law of refraction)
The gradient in Huygens equation (3.2.20) defines the ***ray vector****,*

$$\mathbf{n} = \nabla S = n(\mathbf{r})\hat{\mathbf{s}} \tag{1.1.20}$$

of magnitude $|\mathbf{n}| = n$. *Integration of this gradient around a closed path vanishes, thereby yielding*

$$\oint_P \nabla S(\mathbf{r}) \cdot d\mathbf{r} = \oint_P \mathbf{n}(\mathbf{r}) \cdot d\mathbf{r} = 0\,. \tag{1.1.21}$$

Let's consider the case in which the closed path P *surrounds a boundary separating two different media. If we let the sides of the loop perpendicular to the interface shrink to zero, then only the parts of the line integral tangential to the interface path will contribute. Since these contributions must sum to zero, the tangential components of the ray vectors must be preserved. That is,*

$$(\mathbf{n} - \mathbf{n}') \times \hat{\mathbf{z}} = 0\,, \tag{1.1.22}$$

where the primes refer to the side of the boundary into which the ray is transmitted, whose normal vector is $\hat{\mathbf{z}}$. *Now imagine a ray piercing the boundary and passing into the region enclosed by the integration loop. If* θ *and* θ' *are the angles of incidence and transmission, measured from the direction* $\hat{\mathbf{z}}$ *normal to the boundary, then preservation of the tangential components of the ray vector means that*

$$n \sin\theta = n' \sin\theta'\,. \tag{1.1.23}$$

This is the ***Ibn Sahl-Snell law of refraction****, credited to Ibn Sahl (984) and Willebrord Snellius (1621). A similar analysis may be applied in the case of a reflected ray to show that the angle of incidence must equal the angle of reflection.*

Remark 1.1.12 [Momentum form of Ibn Sahl-Snell law]
The phenomenon of refraction may be seen as a break in the direction $\widehat{\mathsf{s}}$ of the ray vector $\mathbf{n}(\mathbf{r}(s)) = n\widehat{\mathsf{s}}$ at a finite discontinuity

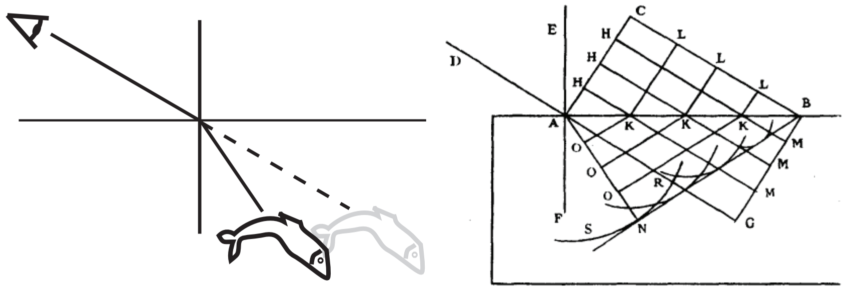


Figure 1.2: Ray tracing version (left) and Huygens version (right) of the Ibn Sahl-Snell law of refraction that $n\sin\theta = n'\sin\theta'$. This law is implied in ray optics by preservation of the components of the ray vector tangential to the interface. According to Huygens principle the law of refraction is implied by slower wave propagation in media of higher refractive index below the horizontal interface.

in the refractive index $n = |\mathbf{n}|$ along the ray path $\mathbf{r}(s)$. According to the eikonal equation (1.1.7) the jump (denoted by Δ) in three-dimensional canonical momentum across the discontinuity must satisfy

$$\Delta\left(\frac{\partial L}{\partial(d\mathbf{r}/ds)}\right) \times \frac{\partial n}{\partial \mathbf{r}} = 0\,.$$

This means the projections $\mathbf{p}$ and $\mathbf{p}'$ of the ray vectors $\mathbf{n}(\mathbf{q}, z)$ and $\mathbf{n}'(\mathbf{q}, z)$ which lie tangent to the plane of the discontinuity in refractive index will be invariant. In particular, the lengths of these projections will be preserved. Consequently,

$$|\mathbf{p}| = |\mathbf{p}'| \quad \text{implies} \quad n\sin\theta = n'\sin\theta' \quad \text{at } z = 0.$$

This is again the ***Ibn Sahl-Snell law***, now written in terms of canonical momentum.

Exercise. How do the canonical momenta differ in the two versions of Fermat's principle in (1.1.8) and (1.1.11)? Do their Ibn Sahl-Snell laws differ? Do their Hamiltonian formulations differ? ★

Answer. The first stationary principle (1.1.8) gives $n(\mathbf{r}(s))d\mathbf{r}/ds$ for the optical momentum, while the second one (1.1.11) gives its reparameterised version $n^2(\mathbf{r}(\tau))d\mathbf{r}/d\tau$. Because $d/ds = nd/d\tau$, the values of the two versions of optical momentum agree in either parameterisation. Consequently, their Ibn Sahl-Snell laws agree. ▲

Remark 1.1.13 As Newton discovered in his famous prism experiment, the propagation of a Huygens wave front depends on the light frequency, ω, through the frequency dependence $n(\mathbf{r}, \omega)$ of the index of refraction of the medium. Having noted this possibility now, in what follows, we shall treat monochromatic light of fixed frequency ω and ignore effects of frequency dispersion. We will also ignore finite-wavelength effects such as interference and diffraction of light. These effects were discovered in a sequence of pioneering scientific investigations during the 350 years after Fermat.

1.1.3 Eikonal equation for axial ray optics

Most optical instruments are designed to possess a line of sight (or primary direction of propagation of light) called the ***optical axis***. Choosing coordinates so that the z-axis coincides with the optical axis expresses the arc-length element ds in terms of the increment along the optical axis, dz, as

$$\begin{aligned} ds &= [(dx)^2 + (dy)^2 + (dz)^2]^{1/2} \\ &= [1 + \dot{x}^2 + \dot{y}^2]^{1/2}\, dz = \frac{1}{\gamma}\, dz\,, \end{aligned} \tag{1.1.24}$$

in which the added notation defines $\dot{x} := dx/dz$, $\dot{y} := dy/dz$ and $\gamma := dz/ds$.

For such optical instruments, formula (1.1.24) for the element of arc length may be used to express the optical length in Fermat's principle as an integral called the ***optical action,***

$$\mathsf{A} := \int_{z_A}^{z_B} L(x, y, \dot{x}, \dot{y}, z)\, dz\,. \tag{1.1.25}$$

Definition 1.1.14 (Trajectories and tangent vectors: tangent bundle)

*The coordinates $(x, y, \dot{x}, \dot{y}) \in \mathbb{R}^2 \times \mathbb{R}^2$ designate points along the ray path through **configuration space** and the **tangent space** of possible vectors along the ray trajectory. The position on a ray passing through an image plane at a fixed value of z is denoted $(x, y) \in \mathbb{R}^2$. The notation $(x, y, \dot{x}, \dot{y}) \in T\mathbb{R}^2 \simeq \mathbb{R}^2 \times \mathbb{R}^2$ for the combined space of positions and tangent vectors designates the **tangent bundle of** $\mathbb{R}^2$. The space $T\mathbb{R}^2$ consists of the union of all the position vectors $(x, y) \in \mathbb{R}^2$ and all the possible tangent vectors $(\dot{x}, \dot{y}) \in \mathbb{R}^2$ at each position (x, y).*

The integrand in the optical action A is expressed as $L : T\mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$, or explicitly

$$L(x, y, \dot{x}, \dot{y}, z) = n(x, y, z)[1 + \dot{x}^2 + \dot{y}^2]^{1/2} = \frac{n(x, y, z)}{\gamma}\,.$$

This is the ***optical Lagrangian,*** in which

$$\gamma := \frac{1}{\sqrt{1 + \dot{x}^2 + \dot{y}^2}} \leq 1\,.$$

We may write the coordinates equivalently as $(x, y, z) = (\mathbf{q}, z)$ where $\mathbf{q} = (x, y)$ is a vector with components in the plane perpendicular to the optical axis at displacement z.

The possible ray paths from a point A to a point B in space may be parameterised for axial ray optics as a family of C^2 curves $\mathbf{q}(z, \varepsilon) \in \mathbb{R}^2$ depending smoothly on a real parameter ε in an interval that includes $\varepsilon = 0$. The ε-family of paths $\mathbf{q}(z, \varepsilon)$ defines a set of smooth transformations of the ray path $\mathbf{q}(z)$. These transformations are taken to satisfy

$$\mathbf{q}(z, 0) = \mathbf{q}(z)\,, \quad \mathbf{q}(z_A, \varepsilon) = \mathbf{q}(z_A)\,, \quad \mathbf{q}(z_B, \varepsilon) = \mathbf{q}(z_B)\,, \tag{1.1.26}$$

so $\varepsilon = 0$ is the identity transformation of the ray path and its endpoints are left fixed. Define the ***variation*** of the optical action (1.1.25) using this parameter as

$$\begin{aligned} \delta \mathsf{A} &= \delta \int_{z_A}^{z_B} L(\mathbf{q}(z), \dot{\mathbf{q}}(z)) dz \\ &:= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \int_{z_A}^{z_B} L(\mathbf{q}(z,\varepsilon), \dot{\mathbf{q}}(z,\varepsilon)) dz \,. \end{aligned} \tag{1.1.27}$$

In this formulation, Fermat's principle is expressed as the ***stationary condition*** $\delta \mathsf{A} = 0$ under infinitesimal variations of the path. This condition implies the ***axial eikonal equation***, as follows.

Theorem 1.1.15 (Fermat's principle for axial eikonal equation)

Stationarity under variations of the ***action*** A,

$$0 = \delta \mathsf{A} = \delta \int_{z_A}^{z_B} L(\mathbf{q}(z), \dot{\mathbf{q}}(z)) dz \,, \tag{1.1.28}$$

for the ***optical Lagrangian***

$$L(\mathbf{q}, \dot{\mathbf{q}}, z) = n(\mathbf{q}, z)[1 + |\dot{\mathbf{q}}|^2]^{1/2} =: \frac{n}{\gamma} \,, \tag{1.1.29}$$

with

$$\gamma := \frac{dz}{ds} = \frac{1}{\sqrt{1 + |\dot{\mathbf{q}}|^2}} \leq 1 \,, \tag{1.1.30}$$

implies the ***axial eikonal equation***

$$\gamma \frac{d}{dz}\Big(n(\mathbf{q}, z)\, \gamma \frac{d\mathbf{q}}{dz} \Big) = \frac{\partial n}{\partial \mathbf{q}} \,, \quad \textit{with} \quad \frac{d}{ds} = \gamma \frac{d}{dz} \,. \tag{1.1.31}$$

Proof. As in the derivation of the eikonal equation (1.1.7), differentiating with respect to ε under the integral sign, denoting the ***varia-***

tional derivative as

$$\delta \mathbf{q}(z) := \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \mathbf{q}(z, \varepsilon)\,, \tag{1.1.32}$$

and integrating by parts produces the following variation of the optical action,

$$\begin{aligned} 0 = \delta \mathsf{A} \;&=\; \delta \int L(\mathbf{q}, \dot{\mathbf{q}}, z)dz \qquad\qquad (1.1.33)\\ &=\; \int \left(\frac{\partial L}{\partial \mathbf{q}} - \frac{d}{dz}\frac{\partial L}{\partial \dot{\mathbf{q}}} \right) \cdot \delta \mathbf{q}\, dz + \left[\frac{\partial L}{\partial \dot{\mathbf{q}}} \cdot \delta \mathbf{q} \right]_{z_A}^{z_B} . \end{aligned}$$

In the second line, one assumes equality of cross derivatives, $\mathbf{q}_{z\varepsilon} = \mathbf{q}_{\varepsilon z}$ evaluated at $\varepsilon = 0$, and thereby exchanges the order of derivatives; so that

$$\delta \dot{\mathbf{q}} = \frac{d}{dz} \delta \mathbf{q}\,.$$

The endpoint terms vanish in the ensuing integration by parts, because $\delta \mathbf{q}(z_A) = 0 = \delta \mathbf{q}(z_B)$. That is, the variation in the ray path must vanish at the prescribed spatial points A and B at z_A and z_B along the optical axis. Since $\delta \mathbf{q}$ is otherwise arbitrary, the principle of stationary action expressed in equation (1.1.28) is equivalent to the following equation, written in a standard form later made famous by Euler and Lagrange,

$$\boxed{\frac{\partial L}{\partial \mathbf{q}} - \frac{d}{dz}\frac{\partial L}{\partial \dot{\mathbf{q}}} = 0 \qquad \textit{\textbf{(Euler-Lagrange equation).}} \qquad (1.1.34)}$$

After a short algebraic manipulation using the explicit form of the optical Lagrangian in (1.1.29), the Euler-Lagrange equation (1.1.34) for the light rays yields the eikonal equation (1.1.31), in which $\gamma\, d/dz = d/ds$ relates derivatives along the optical axis to derivatives in the arc-length parameter. ■

Exercise. Check that the eikonal equation (1.1.31) follows from the Euler-Lagrange equations (1.1.34) with Lagrangian (1.1.29). ★

Corollary 1.1.16 (Noether's theorem)

Each smooth symmetry of the Lagrangian in an action principle implies a conservation law for its Euler-Lagrange equations.

Proof. In a stationary action principle $\delta\mathsf{A} = 0$ for $\mathsf{A} = \int L\, dz$ as in (1.1.28) the Lagrangian L has a symmetry, if it is *invariant* under the transformations $\mathbf{q}(z,0) \to \mathbf{q}(z,\varepsilon)$. In this case, stationary $\delta\mathsf{A} = 0$ under the infinitesimal variations defined in (1.1.32) follows because of this invariance of the Lagrangian, even if these variations did *not* vanish at the endpoints in time. The variational calculation (1.1.33)

$$0 = \delta\mathsf{A} = \int \underbrace{\left(\frac{\partial L}{\partial \mathbf{q}} - \frac{d}{dz}\frac{\partial L}{\partial \dot{\mathbf{q}}}\right)}_{\text{Euler-Lagrange}} \cdot \delta\mathbf{q}\, dz + \underbrace{\left[\frac{\partial L}{\partial \dot{\mathbf{q}}} \cdot \delta\mathbf{q}\right]_{z_A}^{z_B}}_{\text{Noether}} \tag{1.1.35}$$

then shows that along solution paths of the Euler-Lagrange equations (1.1.34) any smooth symmetry of the Lagrangian L implies

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

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 equations) whenever $\delta\mathsf{A} = 0$, because of the symmetry of the Lagrangian L in the action $\mathsf{A} = \int L\, dt$. ■

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

$$\mathsf{A} = \int_{z_A}^{z_B} L(\dot{\mathbf{q}}(z))dz$$

★

Answer. In this case, $\partial L/\partial \mathbf{q} = 0$. This means the action A is invariant under translations $\mathbf{q}(z) \to \mathbf{q}(z) + \boldsymbol{\varepsilon}$ 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 equations (1.1.34). ▲

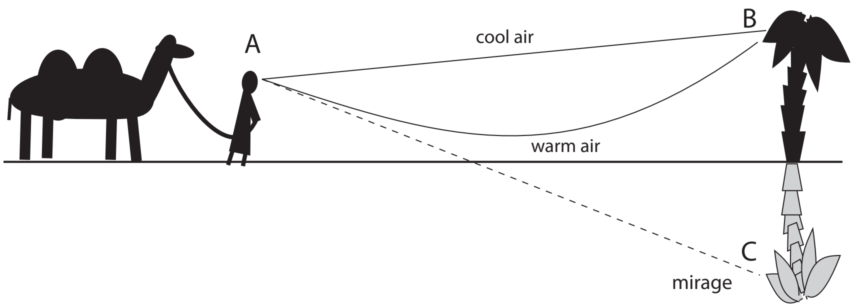


Figure 1.3: Fermat's principle states that the ray path from an observer at A to a point B in space is a stationary path of optical length. For example, along a sun-baked road, the temperature of the air is warmest near the road and decreases with height, so that the index of refraction, n, increases in the vertical direction. For an observer at A, the curved path has the *same* optical path length as the straight line. Therefore, he sees not only the direct line-of-sight image of the tree top at B, but it also appears to him that the tree top has a mirror image at C. If there is no tree, the observer sees a direct image of the sky and also its mirror image of the same optical length, thereby giving the impression, perhaps sadly, that he is looking at water, when there is none.

1.1.4 The eikonal equation for mirages

Air adjacent to a hot surface rises in temperature and becomes less dense. Thus over a flat hot surface, such as a desert expanse or a sun-baked roadway, air density locally increases with height and the average refractive index may be approximated by a linear variation of the form

$$n(x) = n_0(1 + \kappa x)\,,$$

where x is the vertical height above the planar surface, n_0 is the refractive index at ground level, and κ is a positive constant. We may use the eikonal equation (1.1.31) to find an equation for the approximate ray trajectory. This will be an equation for the ray height x as a function of ground distance z of a light ray launched from a height x_0 at an angle θ_0 with respect to the horizontal surface of the earth.

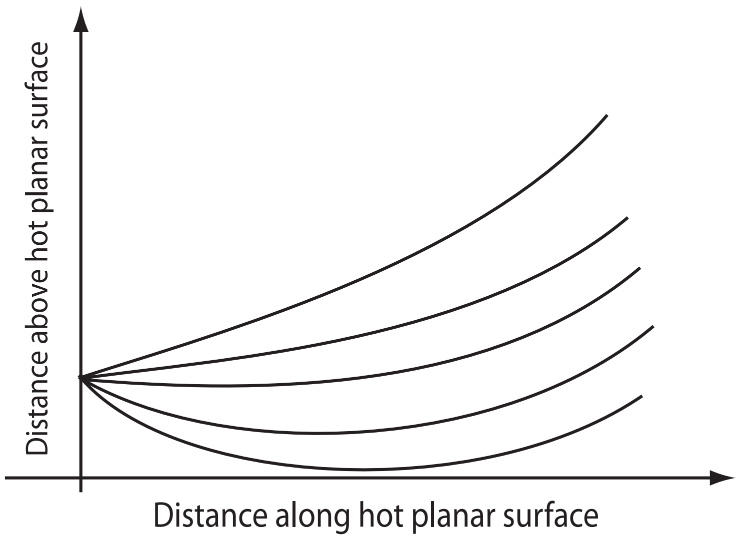


Figure 1.4: Ray trajectories are diverted in a spatially varying medium whose refractive index increases with height above a hot planar surface.

In this geometry, the eikonal equation (1.1.31) implies

$$\frac{1}{\sqrt{1+\dot{x}^2}}\,\frac{d}{dz}\left(\frac{(1+\kappa x)}{\sqrt{1+\dot{x}^2}}\,\dot{x}\right) = \kappa\,.$$

For nearly horizontal rays, $\dot{x}^2 \ll 1$, and if the variation in refractive index is also small then $\kappa x \ll 1$. In this case, the eikonal equation simplifies considerably to

$$\frac{d^2x}{dz^2} \approx \kappa \quad \text{for} \quad \kappa x \ll 1 \quad \text{and} \quad \dot{x}^2 \ll 1\,. \tag{1.1.36}$$

Thus, the ray trajectory is given approximately by

$$\begin{aligned}\mathbf{r}(z) &= x(z)\,\hat{\mathbf{x}} + z\,\hat{\mathbf{z}}\\ &= \Big(\frac{\kappa}{2}z^2 + \tan\theta_0 z + x_0\Big)\hat{\mathbf{x}} + z\,\hat{\mathbf{z}}\,.\end{aligned}$$

The resulting parabolic divergence of rays above the hot surface is shown in Figure 1.4.

> **Exercise.** Explain how the ray pattern would differ from the rays shown in Figure 1.4 if the refractive index were *decreasing* with height x above the surface, rather than increasing. ★

1.1.5 Paraxial optics and classical mechanics

Rays whose direction is nearly along the optical axis are called ***paraxial***. In a medium whose refractive index is nearly homogeneous, paraxial rays remain paraxial and geometric optics closely resembles classical mechanics. Consider the trajectories of paraxial rays through a medium whose refractive index may be approximated by

$$n(\mathbf{q},z) = n_0 - \nu(\mathbf{q},z)\,,\quad \text{with}\quad \nu(\mathbf{0},z)=0 \quad\text{and}\quad \nu(\mathbf{q},z)/n_0 \ll 1\,.$$

Being nearly parallel to the optical axis, paraxial rays satisfy $\theta \ll 1$ and $|\mathbf{p}|/n \ll 1$; so the optical Hamiltonian (1.2.9) may then be approximated by

$$H = -\,n\left[1-\frac{|\mathbf{p}|^2}{n^2}\right]^{1/2} \simeq\, -\,n_0 + \frac{|\mathbf{p}|^2}{2n_0} + \nu(\mathbf{q},z)\,.$$

The constant n_0 here is immaterial to the dynamics. This calculation shows the following.

Lemma 1.1.17 *Geometric ray optics in the paraxial regime corresponds to classical mechanics with a time dependent potential $\nu(\mathbf{q},z)$, upon identifying $z \leftrightarrow t$.*

Exercise. Show that the canonical equations for paraxial rays recover the mirage equation (1.1.36) when $n = n_s(1 + \kappa x)$ for $\kappa > 0$.

Explain what happens to the ray pattern when $\kappa < 0$.

★

1.2 Hamiltonian formulation of axial ray optics

Definition 1.2.1 (Canonical momentum)
The ***canonical momentum*** *(denoted as* $\mathbf{p}$*) associated to the ray path position* $\mathbf{q}$ *in an* ***image plane****, or* ***image screen****, at a fixed value of* z *along the optical axis is defined to be*

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}}\,, \quad \textit{with} \quad \dot{\mathbf{q}} := \frac{d\mathbf{q}}{dz}\,. \tag{1.2.1}$$

Remark 1.2.2 For the optical Lagrangian (1.1.29), the corresponding ***canonical momentum for axial ray optics*** is found to be

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}} = n\gamma\,\dot{\mathbf{q}}\,, \quad \text{which satisfies} \quad |\mathbf{p}|^2 = n^2(1-\gamma^2)\,. \tag{1.2.2}$$

Figure 1.5 illustrates the geometrical interpretation of this momentum for optical rays as the projection along the optical axis of the ray onto an image plane.

From the definition of optical momentum (1.2.2), the corresponding ***velocity*** $\dot{\mathbf{q}} = d\mathbf{q}/dz$ is found as a function of position and momentum $(\mathbf{q}, \mathbf{p})$ as

$$\dot{\mathbf{q}} = \frac{\mathbf{p}}{\sqrt{n^2(\mathbf{q}, z) - |\mathbf{p}|^2}}\,. \tag{1.2.3}$$

A Lagrangian admitting such an invertible relation between $\dot{\mathbf{q}}$ and $\mathbf{p}$ is said to be ***non-degenerate*** (or ***hyperregular*** [MaRa1994]). Moreover, the velocity is real-valued, provided

$$n^2 - |\mathbf{p}|^2 > 0\,. \tag{1.2.4}$$

The latter condition is explained geometrically, as follows.

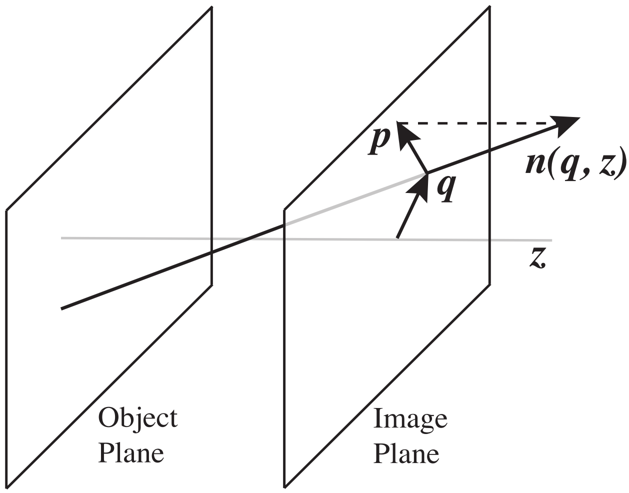


Figure 1.5: Geometrically, the momentum $\mathbf{p}$ associated to the coordinate $\mathbf{q}$ by equation (1.2.1) on the image plane at z turns out to be the projection onto the plane of the ray vector $\mathbf{n}(\mathbf{q}, z) = \nabla S = n(\mathbf{q}, z)d\mathbf{r}/ds$ passing through the point $\mathbf{q}(z)$. That is, $|\mathbf{p}| = n(\mathbf{q}, z)\sin\theta$, where $\cos\theta = dz/ds$ is the direction cosine of the ray with respect to the optical z-axis.

1.2.1 Geometry, phase space and the ray path

Huygens' equation (3.2.20) summons the following geometric picture of the ray path, as shown in Figure 1.5. Along the optical axis (the z-axis) each image plane normal to the axis is pierced at a point $\mathbf{q} = (x, y)$ by the ***ray vector***, defined as

$$\mathbf{n}(\mathbf{q}, z) = \nabla S = n(\mathbf{q}, z)\frac{d\mathbf{r}}{ds} \,.$$

The ray vector is tangent to the ray path and has magnitude $n(\mathbf{q}, z)$. This vector makes an angle $\theta(z)$ with respect to the z-axis at the point $\mathbf{q}$. Its direction cosine with respect to the z-axis is given by

$$\cos\theta := \hat{\mathbf{z}} \cdot \frac{d\mathbf{r}}{ds} = \frac{dz}{ds} = \gamma \,. \tag{1.2.5}$$

This definition of $\cos\theta$ leads by (1.2.2) to

$$|\mathbf{p}| = n\sin\theta \quad \text{and} \quad \sqrt{n^2 - |\mathbf{p}|^2} = n\cos\theta \,. \tag{1.2.6}$$

Thus, the projection of the ray vector $\mathbf{n}(\mathbf{q}, z)$ onto the image plane is the momentum $\mathbf{p}(z)$ of the ray. In three-dimensional vector notation, this is expressed as

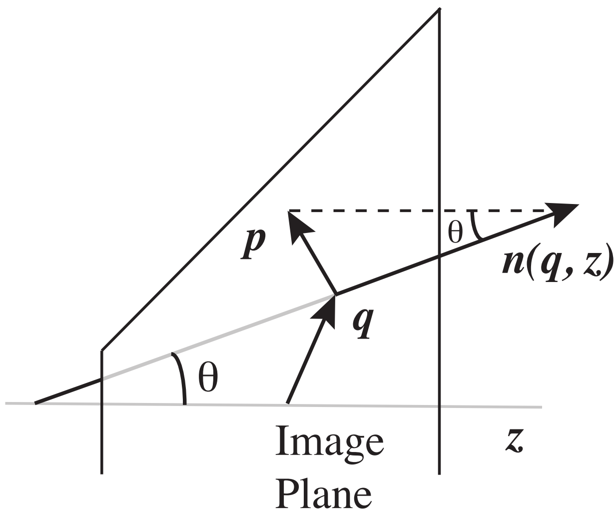


Figure 1.6: The canonical momentum $\mathbf{p}$ associated to the coordinate $\mathbf{q}$ by equation (1.2.1) on the image plane at z has magnitude $|\mathbf{p}| = n(\mathbf{q}, z)\sin\theta$, where $\cos\theta = dz/ds$ is the direction cosine of the ray with respect to the optical z-axis.

$$\mathbf{p}(z) = \mathbf{n}(\mathbf{q}, z) - \hat{\mathbf{z}}\Big(\hat{\mathbf{z}} \cdot \mathbf{n}(\mathbf{q}, z)\Big) . \tag{1.2.7}$$

The coordinates $(\mathbf{q}(z), \mathbf{p}(z))$ determine each ray's position and orientation completely as a function of propagation distance z along the optical axis.

Definition 1.2.3 (Optical phase space, or cotangent bundle)
*The coordinates $(\mathbf{q}, \mathbf{p}) \in \mathbb{R}^2 \times \mathbb{R}^2$ comprise the **phase space** of the ray trajectory. Position on an image plane is denoted $\mathbf{q} \in \mathbb{R}^2$. Phase space coordinates are denoted $(\mathbf{q}, \mathbf{p}) \in T^*\mathbb{R}^2$. The notation $T^*\mathbb{R}^2 \simeq \mathbb{R}^2 \times \mathbb{R}^2$ for phase space designates the **cotangent bundle of** $\mathbb{R}^2$. The space $T^*\mathbb{R}^2$ consists of the union of all the position vectors $\mathbf{q} \in \mathbb{R}^2$ and all the possible canonical momentum vectors $\mathbf{p} \in \mathbb{R}^2$ at each position $\mathbf{q}$.*

Remark 1.2.4 The phase space $T^*\mathbb{R}^2$ for ray optics is restricted to the disc,

$$|\mathbf{p}| < n(\mathbf{q}, z)\,,$$

so that $\cos\theta$ in (1.2.6) remains real. When $n^2 = |\mathbf{p}|^2$, the ray trajectory is tangent to the image screen and is said to have ***grazing incidence*** to the screen at a certain value of z. Rays of grazing incidence are eliminated by restricting the momentum in the phase space for ray optics to lie in a disc $|\mathbf{p}|^2 < n^2(\mathbf{q}, z)$. This restriction implies that the velocity will remain real, finite and of a single sign, which we may choose to be positive ($\dot{\mathbf{q}} > 0$) in the direction of propagation.

1.2.2 Legendre transformation

The passage from the description of the eikonal equation for ray optics in variables $(\mathbf{q}, \dot{\mathbf{q}}, z)$ to its ***phase space description*** in variables $(\mathbf{q}, \mathbf{p}, z)$ is accomplished by applying the ***Legendre transformation*** from the Lagrangian L to the Hamiltonian H, defined as,

$$H(\mathbf{q}, \mathbf{p}) = \mathbf{p} \cdot \dot{\mathbf{q}} - L(\mathbf{q}, \dot{\mathbf{q}}, z)\,. \tag{1.2.8}$$

For the Lagrangian (1.1.29) the Legendre transformation (1.2.8) leads to the following ***optical Hamiltonian***,

$$H(\mathbf{q}, \mathbf{p}) = n\gamma\,|\dot{\mathbf{q}}|^2 - n/\gamma = -n\gamma = -\left[n(\mathbf{q}, z)^2 - |\mathbf{p}|^2\right]^{1/2}, \tag{1.2.9}$$

upon using formula (1.2.3) for the velocity $\dot{\mathbf{q}}(z)$ in terms of the position $\mathbf{q}(z)$ at which the ray punctures the screen at z and its canonical momentum there $\mathbf{p}(z)$. Thus, in the geometric picture of canonical screen optics in Figure 1.5, the component of the ray vector along the optical axis is (minus) the Hamiltonian. That is,

$$\hat{\mathbf{z}} \cdot \mathbf{n}(\mathbf{q}, z) = n(\mathbf{q}, z)\cos\theta = -H\,. \tag{1.2.10}$$

Remark 1.2.5 The optical Hamiltonian in (1.2.9) takes real values, so long as the phase space for ray optics is restricted to the disc $|\mathbf{p}| \le n(\mathbf{q}, z)$. The boundary of this disc is the zero level set of the Hamiltonian, $H = 0$. Thus, flows that preserve the value of the optical Hamiltonian will remain inside its restricted phase space.

Theorem 1.2.6 *The phase space description of the ray path follows from* ***Hamilton's canonical equations****, which are defined as*

$$\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}}\,, \qquad \dot{\mathbf{p}} = -\,\frac{\partial H}{\partial \mathbf{q}}\,. \tag{1.2.11}$$

With the optical Hamiltonian $H(\mathbf{q},\mathbf{p}) = -\left[n(\mathbf{q},z)^2 - |\mathbf{p}|^2\right]^{1/2}$ *in (1.2.9), these are*

$$\dot{\mathbf{q}} = \frac{-1}{H}\mathbf{p}\,, \qquad \dot{\mathbf{p}} = \frac{-1}{2H}\frac{\partial n^2}{\partial \mathbf{q}}\,. \tag{1.2.12}$$

Proof. Hamilton's canonical equations are obtained by differentiating both sides of the Legendre transformation formula (1.2.8) to find

$$\begin{aligned} dH(\mathbf{q},\mathbf{p},z) &= \mathbf{0}\cdot d\dot{\mathbf{q}} + \frac{\partial H}{\partial \mathbf{p}}\cdot d\mathbf{p} + \frac{\partial H}{\partial \mathbf{q}}\cdot d\mathbf{q} + \frac{\partial H}{\partial z}\cdot dz \\ &= \left(\mathbf{p} - \frac{\partial L}{\partial \dot{\mathbf{q}}}\right)\cdot d\dot{\mathbf{q}} + \dot{\mathbf{q}}\cdot d\mathbf{p} - \frac{\partial L}{\partial \mathbf{q}}\cdot d\mathbf{q} - \frac{\partial L}{\partial z}\,dz\,. \end{aligned}$$

The coefficient of $(d\dot{\mathbf{q}})$ vanishes in this expression by virtue of the definition of canonical momentum. Vanishing of this coefficient is required for H to be independent of $\dot{\mathbf{q}}$. Identifying the other coefficients yields the relations

$$\frac{\partial H}{\partial \mathbf{p}} = \dot{\mathbf{q}}\,, \qquad \frac{\partial H}{\partial \mathbf{q}} = -\frac{\partial L}{\partial \mathbf{q}} = -\frac{d}{dz}\frac{\partial L}{\partial \dot{\mathbf{q}}} = -\,\dot{\mathbf{p}}\,, \tag{1.2.13}$$

and

$$\frac{\partial H}{\partial z} = -\,\frac{\partial L}{\partial z}\,, \tag{1.2.14}$$

in which ones uses the Euler-Lagrange equation to derive the second relation. Hence, one finds the canonical Hamiltonian formulas (1.2.11) in Theorem 1.2.6. ∎

Definition 1.2.7 (Canonical momentum)
The momentum $\mathbf{p}$ *defined in (1.2.1) that appears with the position* $\mathbf{q}$ *in Hamilton's canonical equations (1.2.11) is called the **canonical momentum**.*

1.3 Hamiltonian form of optical transmission

Proposition 1.3.1 (Canonical bracket)
Hamilton's canonical equations (1.2.11) arise from a bracket operation,

$$\big\{F,\, H\big\} = \frac{\partial F}{\partial \mathbf{q}} \cdot \frac{\partial H}{\partial \mathbf{p}} - \frac{\partial H}{\partial \mathbf{q}} \cdot \frac{\partial F}{\partial \mathbf{p}}\,, \tag{1.3.1}$$

expressed in terms of position $\mathbf{q}$ *and momentum* $\mathbf{p}$.

Proof. One directly verifies,

$$\dot{\mathbf{q}} = \big\{\mathbf{q},\, H\big\} = \frac{\partial H}{\partial \mathbf{p}} \quad \text{and} \quad \dot{\mathbf{p}} = \big\{\mathbf{p},\, H\big\} = -\,\frac{\partial H}{\partial \mathbf{q}}\,.$$

■

Definition 1.3.2 (Canonically conjugate variables)
The components q_i *and* p_j *of position* $\mathbf{q}$ *and momentum* $\mathbf{p}$ *satisfy*

$$\big\{q_i,\, p_j\big\} = \delta_{ij}\,, \tag{1.3.2}$$

*with respect to the canonical bracket operation (1.3.1). Variables that satisfy this relation are said to be **canonically conjugate**.*

Definition 1.3.3 (Dynamical systems in Hamiltonian form) *A dynamical system on the tangent space* TM *of a space* M

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

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

$$\dot{\mathbf{x}}(t) = \{\mathbf{x}, H\}\,, \quad \textit{for} \quad H : M \to \mathbb{R}\,, \tag{1.3.3}$$

in terms of a ***Poisson bracket*** *operation* $\{\cdot\,,\,\cdot\}$*, which is a map among smooth real functions* $\mathcal{F}(M):\ M\to\mathbb{R}$ *on* M*,*

$$\{\cdot\,,\,\cdot\}:\mathcal{F}(M)\times\mathcal{F}(M)\to\mathcal{F}(M)\,,\tag{1.3.4}$$

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

Definition 1.3.4 (Poisson bracket)
A ***Poisson bracket operation*** $\{\cdot\,,\,\cdot\}$ *is defined as possessing the following properties:*

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

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

for the product of any two functions F *and* G *on* M*, and*

4. *satisfies the* ***Jacobi identity***

$$\{F\,,\,\{G\,,\,H\}\}+\{G\,,\,\{H\,,\,F\}\}+\{H\,,\,\{F\,,\,G\}\}=0\,,\tag{1.3.5}$$

for any three functions F*,* G *and* H *on* M*.*

Remark 1.3.5 Definition 1.3.4 of Poisson bracket certainly includes the ***canonical Poisson bracket*** in (1.3.1) that produces Hamilton's canonical equations (1.2.11), in position $\mathbf{q}$ and conjugate momentum $\mathbf{p}$. However, this definition does not require the Poisson bracket to be expressed in the canonical form (1.3.1).

Exercise. Show that the defining properties of a Poisson bracket hold for the canonical bracket expression in (1.3.1). ★

Exercise. Compute the Jacobi identity (1.3.5) using the canonical Poisson bracket (1.3.1) in one dimension for $F = p$, $G = q$ and H arbitrary. ★

Exercise. What does the Jacobi identity (1.3.5) imply about $\{F\,,\,G\}$ when F and G are constants of motion, so that $\{F\,,\,H\} = 0$ and $\{G\,,\,H\} = 0$ for a Hamiltonian H?

★

Exercise. How do the Hamiltonian formulations differ in the two versions of Fermat's principle in (1.1.8) and (1.1.11)? ★

Answer.

The two Hamiltonian formulations differ, because the Lagrangian in (1.1.8) is homogeneous of degree one in its velocity, while the Lagrangian in (1.1.11) is homogeneous of degree two. Consequently, under the Legendre transformation, the Hamiltonian in the first formulation vanishes identically, while the other Hamiltonian is quadratic in its momentum, namely, $H = |\mathbf{p}|^2/(2n)^2$.
▲

Definition 1.3.6 (Hamiltonian vector fields and flows) *A **Hamiltonian vector field** X_F is a map from a function $F \in \mathcal{F}(M)$ on space M with Poisson bracket $\{\,\cdot\,,\,\cdot\,\}$ to a tangent vector on its tangent space TM given by the Poisson bracket. When M is the optical phase space $T^*\mathbb{R}^2$, this map is given by the partial differential operator obtained by inserting the phase space function F into the canonical Poisson bracket,*

$$X_F = \{\,\cdot\,,\,F\} = \frac{\partial F}{\partial \mathbf{p}} \cdot \frac{\partial}{\partial \mathbf{q}} - \frac{\partial F}{\partial \mathbf{q}} \cdot \frac{\partial}{\partial \mathbf{p}}\,.$$

The solution $\mathbf{x}(t) = \phi_t^F \mathbf{x}$ *of the resulting differential equation,*

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

*yields the **flow*** $\phi_t^F \,:\, \mathbb{R} \times M \,\to\, M$ *of the Hamiltonian vector field* X_F *on* M*. Assuming that it exists and is unique, the solution* $\mathbf{x}(t)$ *of the differential equation is a **curve** on* M *parameterised by* $t \in \mathbb{R}$*. The tangent vector* $\dot{\mathbf{x}}(t)$ *to the flow represented by the curve* $\mathbf{x}(t)$ *at time* t *satisfies*

$$\dot{\mathbf{x}}(t) = X_F \mathbf{x}(t)\,,$$

*which are the **characteristic equations** of the Hamiltonian vector field on manifold* M*.*

Remark 1.3.7 [Caution about caustics]
Caustics were discussed definitively in a famous unpublished paper written by Hamilton in 1823 at the age of 18. Hamilton later published what he called "supplements" to this paper in [Ha1830, Ha1837]. The optical singularities discussed by Hamilton form in bright ***caustic*** surfaces when light reflects off a curved mirror. In the present context, we shall avoid caustics. Indeed, we shall avoid reflection altogether and deal only with smooth Hamiltonian flows in media whose spatial variation in refractive index is smooth. For a modern discussion of caustics, see [Ar1994].

1.3.1 Translation invariant media

- If $n = n(\mathbf{q})$, so that the medium is ***invariant under translations*** along the optical axis with coordinate z, then

  $$\hat{\mathbf{z}} \cdot \mathbf{n}(\mathbf{q}, z) = n(\mathbf{q}, z) \cos\theta = -\, H\,,$$

  in (1.2.10) is ***conserved***. That is, the projection $\hat{\mathbf{z}} \cdot \mathbf{n}(\mathbf{q}, z)$ of the ray vector along the optical axis is constant in translation-invariant media.

- For translation-invariant media, the eikonal equation (1.1.31) simplifies via the canonical equations (1.2.12) to ***Newtonian dynamics***,

$$\boxed{\ddot{\mathbf{q}} = -\frac{1}{2H^2}\frac{\partial n^2}{\partial \mathbf{q}} \quad \text{for} \quad \mathbf{q} \in \mathbb{R}^2 . \qquad (1.3.6)}$$

- Thus, in translation-invariant media, geometric ray tracing formally reduces to Newtonian dynamics in z, with a potential $-n^2(\mathbf{q})$ and with "time" z rescaled along each path by the constant value of $\sqrt{2}\,H$ determined from the initial conditions for each ray at the ***object screen*** at $z = 0$.

Remark 1.3.8 In media for which the index of refraction is not translation-invariant, the optical Hamiltonian $n(\mathbf{q}, z)\cos\theta = -H$ is not generally conserved.

1.3.2 Axisymmetric, translation-invariant materials

In axisymmetric, translation-invariant media, the index of refraction may depend on the distance from the optical axis, $r = |\mathbf{q}|$, but does not depend on the azimuthal angle. As we have seen, translation invariance implies conservation of the optical Hamiltonian. Axisymmetry implies yet another constant of motion. This additional constant of motion allows the Hamiltonian system for the light rays to be reduced to phase plane analysis. For such media, the index of refraction satisfies

$$n(\mathbf{q}, z) = n(r) , \quad \text{where} \quad r = |\mathbf{q}| . \qquad (1.3.7)$$

Passing to polar coordinates (r, ϕ) yields

$$\begin{aligned} \mathbf{q} &= (x, y) = r(\cos\phi,\, \sin\phi) , \\ \mathbf{p} &= (p_x, p_y) \\ &= (p_r\cos\phi - p_\phi\sin\phi/r,\; p_r\sin\phi + p_\phi\cos\phi/r) , \end{aligned}$$

so that

$$|\mathbf{p}|^2 = p_r^2 + p_\phi^2/r^2\,. \tag{1.3.8}$$

Consequently, the optical Hamiltonian,

$$H = -\left[n(r)^2 - p_r^2 - p_\phi^2/r^2\right]^{1/2}, \tag{1.3.9}$$

is *independent* of the azimuthal angle ϕ. This independence of angle ϕ leads to conservation of its canonically conjugate momentum p_ϕ, whose interpretation will be discussed in a moment.

Exercise. Verify formula (1.3.9) for the optical Hamiltonian governing ray optics in axisymmetric, translation-invariant media by computing the Legendre transformation. ★

Answer. Fermat's principle $\delta S = 0$ for $S = \int L\,dz$ an axisymmetric, translation-invariant material may be written in polar coordinates using the Lagrangian

$$L = n(r)\sqrt{1 + \dot{r}^2 + r^2\dot{\phi}^2}\,, \tag{1.3.10}$$

from which one finds

$$p_r = \frac{\partial L}{\partial \dot{r}} = \frac{n(r)\dot{r}}{\sqrt{1 + \dot{r}^2 + r^2\dot{\phi}^2}}\,,$$

and

$$\frac{p_\phi}{r} = \frac{1}{r}\frac{\partial L}{\partial \dot{\phi}} = \frac{n(r)r\dot{\phi}}{\sqrt{1 + \dot{r}^2 + r^2\dot{\phi}^2}}\,.$$

Consequently, the velocities and momenta are related by

$$\frac{1}{\sqrt{1 + \dot{r}^2 + r^2\dot{\phi}^2}} = \sqrt{1 - \frac{p_r^2 + p_\phi^2/r^2}{n^2(r)}} = \sqrt{1 - |\mathbf{p}|^2/n^2(r)}\,,$$

which allows the velocities to be obtained from the momenta and positions. The Legendre transformation (1.2.9),

$$H(r, p_r, p_\phi) = \dot{r}p_r + \dot{\phi}p_\phi - L(r, \dot{r}, \dot{\phi})\,,$$

then yields formula (1.3.9) for the optical Hamiltonian.
▲

Exercise. Interpret the quantity p_ϕ in terms of the vector image-screen phase space variables $\mathbf{p}$ and $\mathbf{q}$. ★

Answer. The vector $\mathbf{q}$ points from the optical axis and lies in the optical (x, y) or (r, ϕ) plane. Hence, the quantity p_ϕ may be expressed in terms of the vector image-screen phase space variables $\mathbf{p}$ and $\mathbf{q}$ as

$$|\mathbf{p}\times\mathbf{q}|^2 = |\mathbf{p}|^2|\mathbf{q}|^2 - (\mathbf{p}\cdot\mathbf{q})^2 = p_\phi^2\,. \tag{1.3.11}$$

This may be obtained by using the relations

$$|\mathbf{p}|^2 = p_r^2 + \frac{p_\phi^2}{r^2}\,,\quad |\mathbf{q}|^2 = r^2 \quad\text{and}\quad \mathbf{q}\cdot\mathbf{p} = rp_r\,.$$

One interprets $p_\phi = \mathbf{p}\times\mathbf{q}$ as the ***oriented area spanned on the optical screen*** by the vectors $\mathbf{q}$ and $\mathbf{p}$. ▲

Exercise. Show that Theorem 1.1.16 (Noether's theorem) implies conservation of the quantity p_ϕ for the axisymmetric Lagrangian (1.3.10) in polar coordinates. ★

Answer. The Lagrangian (1.3.10) is invariant under $\phi \to \phi + \varepsilon$ for constant ε. Noether's theorem then implies conservation of $p_\phi = \partial L/\partial\dot{\phi}$. ▲

Exercise. What conservation law does Noether's theorem imply for the invariance of the Lagrangian (1.3.10) under translations in time, $t \to t + \epsilon$ for a real constant $\epsilon \in \mathbb{R}$. ★

1.3.3 Hamiltonian optics in polar coordinates

Hamilton's equations in polar coordinates are defined for axisymmetric, translation-invariant media by the canonical Poisson brackets with the optical Hamiltonian (1.3.9),

$$
\begin{aligned}
\dot r &= \{r,\,H\} = \frac{\partial H}{\partial p_r} = -\,\frac{p_r}{H}\,,\\
\dot p_r &= \{p_r,\,H\} = -\,\frac{\partial H}{\partial r} = -\,\frac{1}{2H}\frac{\partial}{\partial r}\left(n^2(r) - \frac{p_\phi^2}{r^2}\right)\,,\\
\dot \phi &= \{\phi,\,H\} = \frac{\partial H}{\partial p_\phi} = -\,\frac{p_\phi}{Hr^2}\,,\\
\dot p_\phi &= \{p_\phi,\,H\} = -\,\frac{\partial H}{\partial \phi} = 0\,.
\end{aligned}
\tag{1.3.12}
$$

In the Hamiltonian for axisymmetric ray optics (1.3.9), the constant of the motion p_ϕ may be regarded as a parameter that is set by the initial conditions. Consequently, the motion governed by (1.3.12) restricts to canonical Hamiltonian dynamics for $r(z)$, $p_r(z)$ in a reduced phase space.

Remark 1.3.9 (An equivalent reduction)
Alternatively, the level sets of the Hamiltonian $H = const$ and angular momentum $p_\phi = const$ may be regarded as two surfaces in $\mathbb{R}^3$ with coordinates $\boldsymbol{\chi} = (r, p_r, p_\phi)$. In these $\mathbb{R}^3$ coordinates, Hamilton's equations (1.3.12) may be written as

$$
\dot{\boldsymbol{\chi}}(t) = -\,\frac{\partial p_\phi}{\partial \boldsymbol{\chi}} \times \frac{\partial H}{\partial \boldsymbol{\chi}} \quad \text{with} \quad \boldsymbol{\chi} \in \mathbb{R}^3\,. \tag{1.3.13}
$$

This means that the evolution in $\mathbb{R}^3$ with coordinates $\boldsymbol{\chi} = (r, p_r, p_\phi)$ takes place along the intersections of the level sets of the constants of motion p_ϕ and H. On a level set of p_ϕ the $\mathbb{R}^3$ equations (1.3.13) restrict to the first two equations in the canonical system (1.3.12).

Remark 1.3.10 (Evolution of azimuthal angle)

The evolution of the azimuthal angle, or phase, $\phi(z)$ in polar coordinates for a given value of p_ϕ decouples from the rest of the equations in (1.3.12) and may be found separately, after solving the canonical equations for $r(z)$ and $p_r(z)$.

The polar canonical equation for $\phi(z)$ in (1.3.12) implies, for a given orbit $r(z)$, that the phase may be obtained as a ***quadrature***,

$$\Delta\phi(z) = \int^z \frac{\partial H}{\partial p_\phi}\, dz = -\frac{p_\phi}{H}\int^z \frac{1}{r^2(z)}\, dz\,, \tag{1.3.14}$$

where p_ϕ and H are constants of the motion. Because in this case the integrand is a square, the polar azimuthal angle, or phase, $\Delta\phi(z)$ must either increase or decrease monotonically in axisymmetric ray optics, depending on whether the sign of the conserved ratio p_ϕ/H is negative, or positive, respectively. Moreover, for a fixed value of the ratio p_ϕ/H, rays that are closer to the optical axis circulate around it faster.

The reconstruction of the phase for solutions of Hamilton's optical equations (1.3.12) for ray paths in an axisymmetric, translation-invariant medium has some interesting geometric features for periodic orbits in the radial (r, p_r) phase plane.

1.3.4 Geometric phase for Fermat's principle

One may decompose the total phase change around a closed periodic orbit of period Z in the phase space of radial variables (r, p_r) into the sum of the following two parts:

$$\oint p_\phi\, d\phi = p_\phi \Delta\phi = \underbrace{-\oint p_r\, dr}_{\text{Geometric}} + \underbrace{\oint \mathbf{p}\cdot d\mathbf{q}}_{\text{Dynamic}}\,. \tag{1.3.15}$$

On writing this decomposition of the phase as

$$\Delta\phi = \Delta\phi_{geom} + \Delta\phi_{dyn}\,,$$

one sees that

$$p_\phi \Delta\phi_{geom} = \frac{1}{H}\oint p_r^2 dz = -\iint dp_r \wedge dr \tag{1.3.16}$$

is the area enclosed by the periodic orbit in the radial phase plane. Thus, the name: ***geometric phase*** for $\Delta\phi_{geom}$, because this part of the phase only depends on the geometric area of the periodic orbit. The rest of the phase is given by

$$\begin{aligned} p_\phi \Delta\phi_{dyn} &= \oint \mathbf{p}\cdot d\mathbf{q} \\ &= \oint \left(p_r \frac{\partial H}{\partial p_r} + p_\phi \frac{\partial H}{\partial p_\phi} \right) dz \\ &= \frac{-1}{H} \oint \left(p_r^2 + \frac{p_\phi^2}{r^2} \right) dz \\ &= \frac{1}{H} \oint \big(H^2 - n^2(|\mathbf{q}(z)|) \big) dz \\ &= ZH - \frac{Z}{H} \big\langle n^2 \big\rangle \,, \end{aligned} \tag{1.3.17}$$

where the loop integral $\oint n^2(|\mathbf{q}(z)|)dz = Z\langle n^2\rangle$ defines the average $\langle n^2\rangle$ over the orbit of period Z of the squared index of refraction. This part of the phase depends on the Hamiltonian, orbital period and average of the squared index of refraction over the orbit. Thus, the name: ***dynamic phase*** for $\Delta\phi_{dyn}$, because this part of the phase depends on the dynamics of the orbit, not just its area.

1.3.5 Skewness

Definition 1.3.11 *The quantity*

$$p_\phi = \mathbf{p}\times\mathbf{q} = y p_x - x p_y \,, \tag{1.3.18}$$

is called the ***skewness function***.[2]

Remark 1.3.12 By (1.3.12) the skewness is conserved for rays in axisymmetric media.

[2]This is short notation for $p_\phi = \hat{\mathbf{z}}\cdot\mathbf{p}\times\mathbf{q}$. Scalar notation is standard for a vector normal to a plane that arises as a cross product of vectors in the plane. Of course, the notation for skewness S cannot be confused with the action S.

Remark 1.3.13 Geometrically, the skewness given by the cross product $S = \mathbf{p} \times \mathbf{q}$ is the area spanned on an image screen by the vectors $\mathbf{p}$ and $\mathbf{q}$. This geometric conservation law for screen optics was first noticed by Lagrange in paraxial lens optics and it is still called ***Lagrange's invariant*** in that field. On each screen, the angle, length and point of intersection of the ray vector with the screen may vary. However, the oriented area $S = \mathbf{p} \times \mathbf{q}$ will be the same on each screen, for rays propagating in an axisymmetric medium. This is the geometric meaning of Lagrange's invariant.

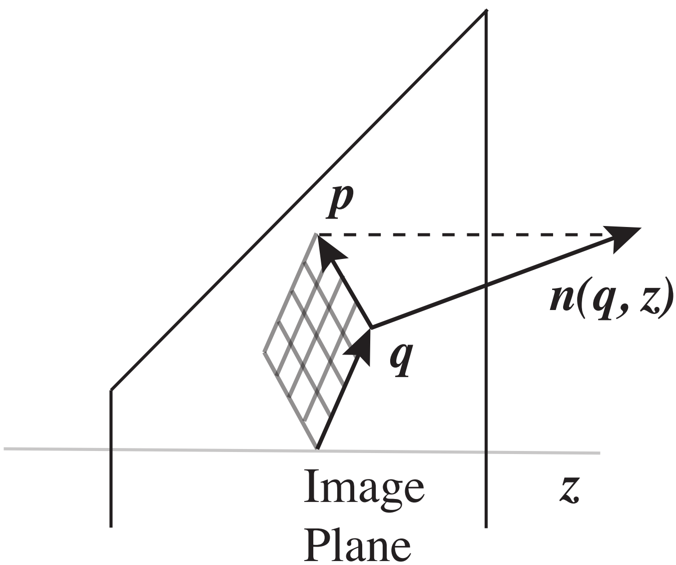


Figure 1.7: The skewness $S = \mathbf{p} \times \mathbf{q}$ of a ray $\mathbf{n}(\mathbf{q}, z)$ is an oriented area in an image plane. For axisymmetric media, skewness is preserved as a function of distance z along the optical axis. The projection $\hat{\mathbf{z}} \cdot \mathbf{n}(\mathbf{q}, z)$ is also conserved, provided the medium is invariant under translations along the optical axis.

Conservation of the skewness function $p_\phi = \mathbf{p} \times \mathbf{q}$ follows in the reduced system (1.3.12) by computing

$$\frac{dp_\phi}{dz} = \{p_\phi,\, H\} = -\frac{\partial H}{\partial \phi} = 0\,,$$

which vanishes because the optical Hamiltonian for an axisymmetric medium is independent of the azimuthal angle ϕ about the optical axis.

> **Exercise.** Check that Hamilton's canonical equations for ray optics (1.2.12) with the optical Hamiltonian (1.2.9) conserve the skewness function $p_\phi = \mathbf{p} \times \mathbf{q}$ when the refractive index satisfies (1.3.7). ★

Remark 1.3.14 The values of the skewness function characterise the various types of rays [Wo2004].

- Vanishing of $\mathbf{p} \times \mathbf{q}$ occurs for ***meridional rays***, for which $\mathbf{p} \times \mathbf{q} = 0$ implies that $\mathbf{p}$ and $\mathbf{q}$ are *collinear* in the image plane $(\mathbf{p} \parallel \mathbf{q})$.
- On the other hand, p_ϕ takes its maximum value for ***sagittal rays***, for which $\mathbf{p} \cdot \mathbf{q} = 0$, so that $\mathbf{p}$ and $\mathbf{q}$ are *orthogonal* in the image plane $(\mathbf{p} \perp \mathbf{q})$.
- Rays that are neither collinear nor orthogonal are said to be ***skew rays***.

Remark 1.3.15 (Sign conventions in optics and mechanics) *Unfortunately, the sign conventions differ between two fundamental ideas that are mathematically the same. Namely, in optics the skewness that characterises the rays is written in the 2D image plane as $S = \mathbf{p} \times \mathbf{q}$, while in mechanics the angular momentum of a particle with momentum $\mathbf{p}$ at position $\mathbf{q}$ in 3D is written as $\mathbf{L} = \mathbf{q} \times \mathbf{p}$. When working in either of these fields, it's probably best to adopt the customs of the natives. However, one must keep a sharp eye out for the difference in signs of rotations when moving between these fields.*

> **Exercise.** ***(Phase plane reduction)***
> ***(1)*** Solve Hamilton's canonical equations for axisymmetric, translation invariant media in the case of an optical

fibre with radially varying index of refraction in the following form:

$$n^2(r) = \lambda^2 + (\mu - \nu r^2)^2\,, \quad \lambda,\, \mu,\, \nu = \text{constants},$$

by reducing the problem to phase plane analysis. How does the phase space portrait differ between $p_\phi = 0$ and $p_\phi \neq 0$? What happens when ν changes sign?

(2) What regions of the phase plane admit real solutions? Is it possible for a phase point to pass from a region with real solutions to a region with complex solutions during its evolution? Prove it.

(3) Compute the dynamic and geometric phases for a periodic orbit of period Z in the (r, p_r) phase plane.

Hint: For $p_\phi \neq 0$ the problem reduces to a ***Duffing oscillator*** (Newtonian motion in a quartic potential) in a rotating frame, up to a rescaling of time by the value of the Hamiltonian on each ray "orbit".

See [HoKo1991] for a discussion of optical ray chaos under periodic perturbations of this solution. ★

1.3.6 Lagrange invariant: Poisson bracket relations

Under the canonical Poisson bracket (1.3.1), the skewness function, or Lagrange invariant,

$$S = p_\phi = \hat{\mathbf{z}} \cdot \mathbf{p} \times \mathbf{q} = y p_x - x p_y\,, \tag{1.3.19}$$

generates rotations of $\mathbf{q}$ and $\mathbf{p}$ jointly in the image plane. Both $\mathbf{q}$ and $\mathbf{p}$ are rotated by the same angle ϕ around the optical axis $\hat{\mathbf{z}}$. In other words, the equation $(d\mathbf{q}/d\phi, d\mathbf{p}/d\phi) = \{(\mathbf{q}, \mathbf{p})\,, S\}$ defined by the Poisson bracket,

$$\frac{d}{d\phi} = X_S = \Big\{\, \cdot\,,\, S \Big\} = \mathbf{q} \times \hat{\mathbf{z}} \cdot \frac{\partial}{\partial \mathbf{q}} + \mathbf{p} \times \hat{\mathbf{z}} \cdot \frac{\partial}{\partial \mathbf{p}}\,, \tag{1.3.20}$$

has the solution,

$$\begin{pmatrix} \mathbf{q}(\phi) \\ \mathbf{p}(\phi) \end{pmatrix} = \begin{pmatrix} R_z(\phi) & 0 \\ 0 & R_z(\phi) \end{pmatrix} \begin{pmatrix} \mathbf{q}(0) \\ \mathbf{p}(0) \end{pmatrix}. \tag{1.3.21}$$

Here the matrix

$$R_z(\phi) = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \tag{1.3.22}$$

represents rotation of both $\mathbf{q}$ and $\mathbf{p}$ by an angle ϕ about the optical axis.

Remark 1.3.16 The application of the Hamiltonian vector field X_S for skewness in (1.3.20) to the position vector $\mathbf{q}$ yields

$$X_S\mathbf{q} = -\,\hat{\mathbf{z}} \times \mathbf{q}\,. \tag{1.3.23}$$

Likewise, the application of the Hamiltonian vector field X_S for skewness in (1.3.20) to the momentum vector yields

$$X_S\mathbf{p} = -\,\hat{\mathbf{z}} \times \mathbf{p}\,. \tag{1.3.24}$$

Thus, the application of X_S to the vectors $\mathbf{q} \in \mathbb{R}^3$ and $\mathbf{p} \in \mathbb{R}^3$ rotates them both about $\hat{\mathbf{z}}$ by the same angle.

Definition 1.3.17 (Diagonal action & cotangent lift)
Together, formulas (1.3.23) - (1.3.24) comprise the ***diagonal action*** *on* $(\mathbf{q},\mathbf{p})$ *of axial rotations about* $\hat{\mathbf{z}}$. *The rotation of the momentum vector* $\mathbf{p}$ *that is* induced *by the rotation of the position vector* $\mathbf{q}$ *is called the* ***cotangent lift*** *of the action of the Hamiltonian vector field* X_S. *Namely, (1.3.24) is the lift of the action of rotation (1.3.23) from position vectors to momentum vectors.*

Remark 1.3.18 [Moment of momentum]
Applying the Hamiltonian vector field X_S for skewness in (1.3.20) to screen coordinates $\mathbf{q} = \mathbb{R}^2$ produces the infinitesimal action of rotations about $\hat{\mathbf{z}}$, as

$$X_S\mathbf{q} = \{\mathbf{q},\, S\} = -\,\hat{\mathbf{z}} \times \mathbf{q} = \frac{d\mathbf{q}}{d\phi}\bigg|_{\phi=0}\,.$$

The skewness function S in (1.3.19) may be expressed in terms of ***two different pairings,***

$$\begin{aligned} S &= \Big\langle\!\!\Big\langle \mathbf{p},\, X_S\mathbf{q} \Big\rangle\!\!\Big\rangle = \mathbf{p}\cdot(-\,\hat{\mathbf{z}}\times\mathbf{q}) \quad \text{and} \\ S &= (\mathbf{p}\times\mathbf{q})\cdot\hat{\mathbf{z}} = \Big\langle \mathbf{J}(\mathbf{p},\,\mathbf{q}),\,\hat{\mathbf{z}} \Big\rangle = J^z(\mathbf{p},\,\mathbf{q})\,. \end{aligned} \tag{1.3.25}$$

Although these pairings are both written as dot-products of vectors, strictly speaking they act on different spaces. Namely,

$$\begin{aligned}\Big\langle\!\!\Big\langle \cdot\,,\,\cdot \Big\rangle\!\!\Big\rangle : &\quad \text{(momentum)} \times \text{(velocity)} \to \mathbb{R}\,, \qquad (1.3.26)\\ \Big\langle \cdot\,,\,\cdot \Big\rangle : &\quad \text{(moment of momentum)} \times \text{(rotation rate)} \to \mathbb{R}\,.\end{aligned}$$

The first pairing $\langle\!\langle\,\cdot\,,\,\cdot\,\rangle\!\rangle$ is between two vectors that are tangent to an optical screen. These vectors represent the projection of the ray vector on the screen $\mathbf{p}$ and the rate of change of the position $\mathbf{q}$ with azimuthal angle, $d\mathbf{q}/d\phi$ in (1.3.23). This is also the pairing $\langle\!\langle\,\cdot\,,\,\cdot\,\rangle\!\rangle$ between velocity and momentum that appears in the Legendre transformation. The second pairing $\langle\,\cdot\,,\,\cdot\,\rangle$ is between the oriented area $\mathbf{p}\times\mathbf{q}$ and the normal to the screen $\hat{\mathbf{z}}$. Thus, as we knew, $J^z(\mathbf{p},\,\mathbf{q}) = S(\mathbf{p},\,\mathbf{q})$ is the Hamiltonian for an infinitesimal rotation about the $\hat{\mathbf{z}}$ axis in $\mathbb{R}^3$.

Definition 1.3.19 *Distinguishing between the pairings in (1.3.25) interprets the Lagrange invariant $S = J^z(\mathbf{p},\,\mathbf{q}) = \mathbf{p}\times\mathbf{q}\cdot\hat{\mathbf{z}}$ as the $\hat{\mathbf{z}}$-component of a map from phase space with coordinates $(\mathbf{p},\,\mathbf{q})$ to the oriented area $\mathbf{J}(\mathbf{p},\,\mathbf{q}) = \mathbf{p}\times\mathbf{q}$, or* ***moment of momentum****.*

Definition 1.3.20 (Momentum map for cotangent lift) *Formula (1.3.25) defines the momentum map for the* ***cotangent lift*** *of the action of rotations about $\hat{\mathbf{z}}$ from position vectors to their canonically conjugate momentum vectors in phase space. In general, a momentum map applies from phase space to the dual space of the Lie algebra of the Lie group whose action is involved. In this case, it is the map from phase space to the moment-of-momentum space, $\mathcal{M}$,*

$$\mathbf{J} : T^*\mathbb{R}^2 \to \mathcal{M}\,, \quad \textit{namely,} \quad \mathbf{J}(\mathbf{p},\,\mathbf{q}) = \mathbf{p}\times\mathbf{q}\,, \qquad (1.3.27)$$

and $\mathbf{p}\times\mathbf{q}$ is dual to the rotation rate about the axial direction $\hat{\mathbf{z}}$ under the pairing given by the three-dimensional scalar (dot) product. The corresponding Hamiltonian is the skewness

$$S = J^z(\mathbf{p},\,\mathbf{q}) = \mathbf{J}\cdot\hat{\mathbf{z}} = \mathbf{p}\times\mathbf{q}\cdot\hat{\mathbf{z}}$$

in (1.3.25). This is the real-valued phase space function whose Hamiltonian vector field X_S rotates a point $P = (\mathbf{q}, \mathbf{p})$ in phase space about the optical axis $\hat{\mathbf{z}}$ at its centre, according to

$$-\hat{\mathbf{z}} \times P = X_{\mathbf{J}\cdot\hat{\mathbf{z}}} P = \{\, P \,,\, \mathbf{J}\cdot\hat{\mathbf{z}}\} \,. \tag{1.3.28}$$

Remark 1.3.21 The skewness function S and its square, S^2 (called the ***Petzval invariant*** [Wo2004]) are conserved for ray optics in axisymmetric media. That is, the canonical Poisson bracket vanishes

$$\{S^2, H\} = 0 \,, \tag{1.3.29}$$

for optical Hamiltonians of the form,

$$H = -\left[n(|\mathbf{q}|^2)^2 - |\mathbf{p}|^2 \right]^{1/2} . \tag{1.3.30}$$

The Poisson bracket (1.3.29) vanishes because $|\mathbf{q}|^2$ and $|\mathbf{p}|^2$ in H both remain invariant under the simultaneous rotations of $\mathbf{q}$ and $\mathbf{p}$ about $\hat{\mathbf{z}}$ generated by S in (1.3.20).

1.4 Axisymmetric invariant coordinates

Transforming to axisymmetric coordinates and azimuthal angle in the optical phase space is similar to passing to polar coordinates (radius and angle) in the plane. Passing to polar coordinates by $(x, y) \to (r, \phi)$ decomposes the plane $\mathbb{R}^2$ into the product of the real line $r \in \mathbb{R}^+$ and the angle $\phi \in S^1$. Quotienting the plane by the angle leaves just the real line. The ***quotient map for the plane*** is

$$\pi : \mathbb{R}^2\{\mathbf{0}\} \to \mathbb{R}\backslash\{\mathbf{0}\} : \, (x, y) \to r \,. \tag{1.4.1}$$

The S^1 angle in optical phase space $T^*\mathbb{R}^2$ is the azimuthal angle. But how does one quotient the four-dimensional $T^*\mathbb{R}^2$ by the azimuthal angle?

As discussed in Section 1.3.3, azimuthal symmetry of the Hamiltonian summons the transformation to polar coordinates in phase space, as

$$(\mathbf{q},\mathbf{p}) \to (r, p_r; p_\phi, \phi)\,.$$

This transformation reduces the motion to phase planes of radial (r, p_r) position and momentum, defined on level surfaces of the skewness p_ϕ. The trajectories evolve along intersections of the the level sets of skewness (the planes $p_\phi = const$) with the level sets of the Hamiltonian $H(r, p_r, p_\phi) = const$. The motion along these intersections is independent of the *ignorable* phase variable $\phi \in S^1$, whose evolution thus decouples from that of the other variables. Consequently, the phase evolution may be reconstructed later by a quadrature, i.e., an integral that involves the parameters of the *reduced phase space*. Thus, in this case, azimuthal symmetry decomposes the phase space *exactly* as

$$T^*\mathbb{R}^2\backslash\{\mathbf{0}\} \simeq \big(T^*(\mathbb{R}\backslash\{0\}) \times \mathbb{R}\big) \times S^1\,. \tag{1.4.2}$$

The corresponding ***quotient map for azimuthal symmetry*** is

$$\pi : T^*\mathbb{R}^2\backslash\{\mathbf{0}\} \to T^*(\mathbb{R}\backslash\{0\}) \times \mathbb{R} : (\mathbf{q},\mathbf{p}) \to (r, p_r; p_\phi)\,. \tag{1.4.3}$$

An alternative procedure exists for quotienting out the angular dependence of an azimuthally symmetric Hamiltonian system, which is independent of the details of the Hamiltonian function. This alternative procedure involves transforming to quadratic azimuthally invariant functions.

Definition 1.4.1 (Quotient map to quadratic S^1 invariants)
The quadratic ***azimuthally invariant coordinates*** *in* $\mathbb{R}^3\backslash\{\mathbf{0}\}$ *are defined by the* ***quotient map***[3]

$$\pi : T^*\mathbb{R}^2\backslash\{0\} \to \mathbb{R}^3\backslash\{\mathbf{0}\} : (\mathbf{q},\mathbf{p}) \to \mathbf{X} = (X_1, X_2, X_3)\,, \tag{1.4.4}$$

given explicitly by the quadratic monomials,

$$X_1 = |\mathbf{q}|^2 \geq 0\,, \quad X_2 = |\mathbf{p}|^2 \geq 0\,, \quad X_3 = \mathbf{p}\cdot\mathbf{q}\,. \tag{1.4.5}$$

[3]The transformation $T^*\mathbb{R}^2 \to \mathbb{R}^3$ in (1.4.5) will be recognised later as another example of a ***momentum map***.

The quotient map (1.4.4) is written a bit more succinctly as

$$\pi\,(\mathbf{p},\,\mathbf{q}) = \mathbf{X}\,. \tag{1.4.6}$$

Theorem 1.4.2 *The vector* (X_1, X_2, X_3) *of quadratic monomials in phase space all Poisson-commute with skewness* S*,*

$$\{S, X_1\} = 0\,, \quad \{S, X_2\} = 0\,, \quad \{S, X_3\} = 0\,. \tag{1.4.7}$$

Proof. These three Poisson brackets with skewness S all vanish because dot products of vectors are preserved by the joint rotations of $\mathbf{q}$ and $\mathbf{p}$ that are generated by S. ■

Remark 1.4.3 The orbits of S in (1.3.21) are rotations of both $\mathbf{q}$ and $\mathbf{p}$ by an angle ϕ about the optical axis at a fixed position z. According to the relation $\{S,\,\mathbf{X}\} = 0$, the quotient map $\mathbf{X} = \pi\,(\mathbf{p},\,\mathbf{q})$ in (1.4.4) collapses each circular orbit of S on a given image screen in phase space $T^*\mathbb{R}^2\backslash\{0\}$ to a point in $\mathbb{R}^3\backslash\{\mathbf{0}\}$. The converse also holds. Namely, the inverse of the quotient map $\pi^{-1}\mathbf{X}$ for $\mathbf{X} \in \mathrm{Image}\,\pi$ consists of the circle (S^1) generated by the rotation of phase space about its centre by the flow of S.

Definition 1.4.4 (Orbit manifold)
The image in $\mathbb{R}^3$ *of the quotient map* $\pi : T^*\mathbb{R}^2\backslash\{0\} \to \mathbb{R}^3\backslash\{\mathbf{0}\}$ *in (1.4.4) is the* ***orbit manifold*** *for axisymmetric ray optics.*

Remark 1.4.5 [Orbit manifold for axisymmetric ray optics]
The image of the quotient map π in (1.4.4) may be conveniently displayed as the zero-level set of the the relation

$$C(X_1,\,X_2,\,X_3,\,S) = S^2 - (X_1X_2 - X_3^2) = 0\,, \tag{1.4.8}$$

among the axisymmetric variables in equation (1.4.5). Consequently, a level set of S in the quotient map $T^*\mathbb{R}^2\backslash\{0\} \to \mathbb{R}^3\backslash\{\mathbf{0}\}$ obtained by transforming to S^1 phase space invariants yields an orbit manifold defined by $C(X_1,\,X_2,\,X_3,\,S) = 0$ in $\mathbb{R}^3\backslash\{\mathbf{0}\}$.

For axisymmetric ray optics, the image of the quotient map π in $\mathbb{R}^3$ turns out to be a family of hyperboloids of revolution.

1.5 Geometry of invariant coordinates

In terms of the axially invariant coordinates (1.4.5), the Petzval invariant and the square of the optical Hamiltonian satisfy

$$|\mathbf{p}\times\mathbf{q}|^2 = |\mathbf{p}|^2|\mathbf{q}|^2-(\mathbf{p}\cdot\mathbf{q})^2 \quad \text{and} \quad H^2 = n^2(|\mathbf{q}|^2)-|\mathbf{p}|^2 \geq 0\,. \tag{1.5.1}$$

That is,

$$S^2 = X_1X_2 - X_3^2 \geq 0 \quad \text{and} \quad H^2 = n^2(X_1) - X_2 \geq 0\,. \tag{1.5.2}$$

The geometry of the solution is determined by the intersections of the level sets of the conserved quantities S^2 and H^2. The level sets of $S^2 \in \mathbb{R}^3$ are hyperboloids of revolution around the $X_1 = X_2$ axis in the horizontal plane defined by $X_3 = 0$. The level-set hyperboloids lie in the interior of the $S = 0$ cone with $X_1 > 0$ and $X_2 > 0$. The level sets of H^2 depend on the functional form of the index of refraction, but they are X_3-independent. The ray path in the S^1-invariant variables $\mathbf{X} = (X_1, X_2, X_3) \in \mathbb{R}^3$ must occur along intersections of S^2 and H^2, since both of these quantities are conserved along the ray path in axisymmetric translation-invariant media.

One would naturally ask how the quadratic phase space quantities (X_1, X_2, X_3) Poisson-commute among themselves. However, before addressing that question, let us ask the following.

Question 1.5.1
How does the Poisson bracket with each of the axisymmetric quantities (X_1, X_2, X_3) act as a derivative operation on functions of the phase space variables $\mathbf{q}$ and $\mathbf{p}$?

Remark 1.5.2
Answering this question introduces the concept of ***flows of Hamiltonian vector fields***.

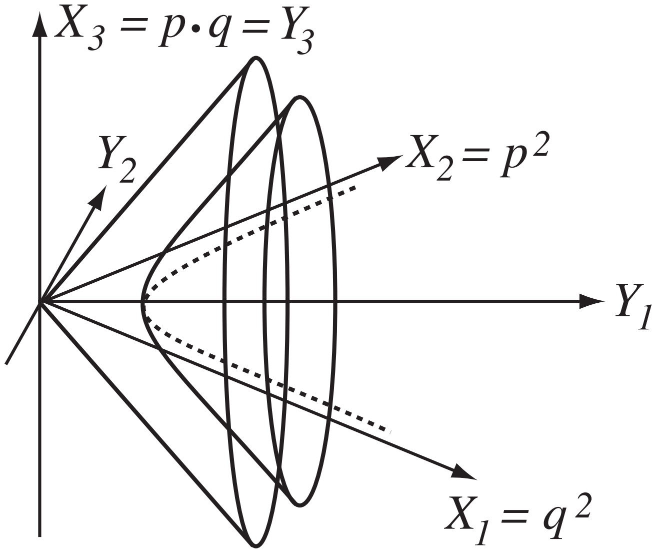


Figure 1.8: Level sets of the Petzval invariant $S^2 = X_1X_2 - X_3^2$ are hyperboloids of revolution around the $X_1 = X_2$ axis (along Y_1) in the horizontal plane, $X_3 = 0$. Level sets of the Hamiltonian H in (1.5.2) are independent of the vertical coordinate. The axisymmetric invariants $\mathbf{X} \in \mathbb{R}^3$ evolve along the intersections of these level sets by $\dot{\mathbf{X}} = \nabla S^2 \times \nabla H$, as the vertical Hamiltonian knife $H = constant$ slices through the *hyperbolic onion* of level sets of S^2. In the coordinates, $Y_1 = (X_1 + X_2)/2\,, Y_2 = (X_2 - X_1)/2\,, Y_3 = X_3$, one has $S^2 = Y_1^2 - Y_2^2 - Y_3^2$. Being invariant under the flow of the Hamiltonian vector field $X_S = \{\,\cdot\,,\,S\}$, each point on any layer H^2 of the hyperbolic onion H^3 consists of an S^1 orbit in phase space under the diagonal rotation (1.3.21). This orbit is a circular rotation of both $\mathbf{q}$ and $\mathbf{p}$ on an image screen at position z by an angle ϕ about the optical axis.

1.5.1 Flows of Hamiltonian vector fields

Theorem 1.5.3 (Flows of Hamiltonian vector fields)
*Poisson brackets with the S^1-invariant phase space functions X_1, X_2 and X_3 generate linear homogeneous transformations of $(\mathbf{q},\,\mathbf{p}) \in T^*\mathbb{R}^2$, obtained by regarding the **Hamiltonian vector fields** obtained as in Definition 1.3.6 from the Poisson brackets as derivatives,*

$$\frac{d}{d\tau_1} := \{\,\cdot\,,\,X_1\}\,, \quad \frac{d}{d\tau_2} := \{\,\cdot\,,\,X_2\} \quad \textit{and} \quad \frac{d}{d\tau_3} := \{\,\cdot\,,\,X_3\}\,, \quad (1.5.3)$$

*in their **flow parameters** τ_1, τ_2 and τ_3, respectively.*

The flows themselves may be determined by integrating the characteristic equations of these Hamiltonian vector fields.

Proof.

- The sum $\frac{1}{2}(X_1 + X_2)$ is the harmonic-oscillator Hamiltonian. This Hamiltonian generates rotation of the $(\mathbf{q},\, \mathbf{p})$ phase space around its centre, by integrating the ***characteristic equations of its Hamiltonian vector field***,

$$\frac{d}{d\omega} = \Big\{ \cdot\, ,\, \frac{1}{2}(X_1 + X_2) \Big\} = \mathbf{p} \cdot \frac{\partial}{\partial \mathbf{q}} - \mathbf{q} \cdot \frac{\partial}{\partial \mathbf{p}}\,. \tag{1.5.4}$$

To see this, write the simultaneous equations

$$\frac{d}{d\omega} \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} = \left\{ \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} ,\, \frac{1}{2}(X_1 + X_2) \right\} ,$$

or in matrix form,[4]

$$\frac{d}{d\omega} \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} =: \frac{1}{2}(m_1 + m_2) \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} ,$$

for the 2×2 traceless matrices m_1 and m_2 defined by

$$m_1 = \begin{pmatrix} 0 & 0 \\ -2 & 0 \end{pmatrix} \quad \text{and} \quad m_2 = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} .$$

These ω-dynamics may be rewritten as a complex equation,

$$\frac{d}{d\omega}(\mathbf{q} + i\mathbf{p}) = -i(\mathbf{q} + i\mathbf{p})\,, \tag{1.5.5}$$

whose immediate solution is

$$\mathbf{q}(\omega) + i\mathbf{p}(\omega) = e^{-i\omega} \Big(\mathbf{q}(0) + i\mathbf{p}(0) \Big)\,.$$

[4]For rotational symmetry, it is sufficient to restrict attention to rays lying in a fixed azimuthal plane and, thus, we may write these actions using 2×2 matrices, rather than 4×4 matrices.

This solution may also be written in matrix form as

$$\begin{pmatrix} \mathbf{q}(\omega) \\ \mathbf{p}(\omega) \end{pmatrix} = \begin{pmatrix} \cos\omega & \sin\omega \\ -\sin\omega & \cos\omega \end{pmatrix} \begin{pmatrix} \mathbf{q}(0) \\ \mathbf{p}(0) \end{pmatrix}, \tag{1.5.6}$$

which is a ***diagonal clockwise rotation*** of $(\mathbf{q}, \mathbf{p})$. This solution sums the following exponential series

$$\begin{aligned} e^{\omega(m_1+m_2)/2} &= \sum_{n=0}^{\infty} \frac{\omega^n}{n!} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^n \\ &= \begin{pmatrix} \cos\omega & \sin\omega \\ -\sin\omega & \cos\omega \end{pmatrix}. \end{aligned} \tag{1.5.7}$$

This may also be verified by summing its even and odd powers separately.

Likewise, a nearly identical calculation yields

$$\frac{d}{d\gamma} \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} = \frac{1}{2}(m_2 - m_1) \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix},$$

for the dynamics of the Hamiltonian $H = (|\mathbf{p}|^2 - |\mathbf{q}|^2)/2$. The time, the solution is the hyperbolic rotation

$$\begin{pmatrix} \mathbf{q}(\gamma) \\ \mathbf{p}(\gamma) \end{pmatrix} = \begin{pmatrix} \cosh\gamma & \sinh\gamma \\ \sinh\gamma & \cosh\gamma \end{pmatrix} \begin{pmatrix} \mathbf{q}(0) \\ \mathbf{p}(0) \end{pmatrix}, \tag{1.5.8}$$

which, in turn, sums the exponential series

$$\begin{aligned} e^{\gamma(m_2-m_1)/2} &= \sum_{n=0}^{\infty} \frac{\gamma^n}{n!} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^n \\ &= \begin{pmatrix} \cosh\gamma & \sinh\gamma \\ \sinh\gamma & \cosh\gamma \end{pmatrix}. \end{aligned} \tag{1.5.9}$$

- In ray optics, the canonical Poisson bracket with the quadratic phase space function $X_1 = |\mathbf{q}|^2$ defines the action of the following *linear* Hamiltonian vector field:

$$\frac{d}{d\tau_1} = \Big\{ \cdot\,,\, X_1 \Big\} = -2\mathbf{q} \cdot \frac{\partial}{\partial \mathbf{p}}\,. \tag{1.5.10}$$

This action may be written equivalently in matrix form as

$$\frac{d}{d\tau_1}\begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ -2 & 0 \end{pmatrix}\begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} = m_1 \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix}.$$

Integration of this system of equations yields the finite transformation

$$\begin{aligned} \begin{pmatrix} \mathbf{q}(\tau_1) \\ \mathbf{p}(\tau_1) \end{pmatrix} &= e^{\tau_1 m_1} \begin{pmatrix} \mathbf{q}(0) \\ \mathbf{p}(0) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ -2\tau_1 & 1 \end{pmatrix}\begin{pmatrix} \mathbf{q}(0) \\ \mathbf{p}(0) \end{pmatrix} \\ &=: M_1(\tau_1)\begin{pmatrix} \mathbf{q}(0) \\ \mathbf{p}(0) \end{pmatrix}. \end{aligned} \tag{1.5.11}$$

This is an easy result, because the matrix m_1 is nilpotent. That is, $m_1^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, so the formal series representing the exponential of the matrix

$$e^{\tau_1 m_1} = \sum_{n=0}^{\infty} \frac{1}{n!}(\tau_1 m_1)^n \tag{1.5.12}$$

truncates at its second term. This solution may be interpreted as the ***action of a thin lens*** [Wo2004].

- Likewise, the canonical Poisson bracket with $X_2 = |\mathbf{p}|^2$ defines the linear Hamiltonian vector field,

$$\frac{d}{d\tau_2} = \Big\{ \cdot \,,\, X_2 \Big\} = 2\mathbf{p} \cdot \frac{\partial}{\partial \mathbf{q}}. \tag{1.5.13}$$

In matrix form, this is

$$\frac{d}{d\tau_2}\begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} = m_2 \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix},$$

in which the matrix m_2 is also nilpotent. Its integration generates the finite transformation

$$\begin{aligned}\begin{pmatrix}\mathbf{q}(\tau_2)\\ \mathbf{p}(\tau_2)\end{pmatrix} &= e^{\tau_2 m_2}\begin{pmatrix}\mathbf{q}(0)\\ \mathbf{p}(0)\end{pmatrix}\\ &= \begin{pmatrix}1 & 2\tau_2\\ 0 & 1\end{pmatrix}\begin{pmatrix}\mathbf{q}(0)\\ \mathbf{p}(0)\end{pmatrix}\\ &=: M_2(\tau_2)\begin{pmatrix}\mathbf{q}(0)\\ \mathbf{p}(0)\end{pmatrix}, \end{aligned}\tag{1.5.14}$$

corresponding to ***free propagation*** of light rays in a homogeneous medium.

- The transformation generated by $X_3 = \mathbf{q}\cdot\mathbf{p}$ compresses phase space along one coordinate and expands it along the other, while preserving skewness. Its Hamiltonian vector field is

$$\frac{d}{d\tau_3} = \Big\{\,\cdot\,,\,X_3\Big\} = \mathbf{q}\cdot\frac{\partial}{\partial\mathbf{q}} - \mathbf{p}\cdot\frac{\partial}{\partial\mathbf{p}}\,.$$

Being linear, this may be written in matrix form as

$$\frac{d}{d\tau_3}\begin{pmatrix}\mathbf{q}\\ \mathbf{p}\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}\mathbf{q}\\ \mathbf{p}\end{pmatrix} =: m_3\begin{pmatrix}\mathbf{q}\\ \mathbf{p}\end{pmatrix}.$$

The integration of this linear system generates the flow, or finite transformation,

$$\begin{aligned}\begin{pmatrix}\mathbf{q}(\tau_3)\\ \mathbf{p}(\tau_3)\end{pmatrix} &= e^{\tau_3 m_3}\begin{pmatrix}\mathbf{q}(0)\\ \mathbf{p}(0)\end{pmatrix}\\ &= \begin{pmatrix}e^{\tau_3} & 0\\ 0 & e^{-\tau_3}\end{pmatrix}\begin{pmatrix}\mathbf{q}(0)\\ \mathbf{p}(0)\end{pmatrix}\\ &=: M_3(\tau_3)\begin{pmatrix}\mathbf{q}(0)\\ \mathbf{p}(0)\end{pmatrix}, \end{aligned}\tag{1.5.15}$$

whose exponential series is easily summed, because m_3 is diagonal and constant. Thus, the quadratic quantity X_3 generates a transformation that takes one harmonic-oscillator Hamiltonian into another one corresponding to a different natural frequency. This transformation is called ***squeezing*** of light.

The proof of Theorem 1.5.3 is now finished. ■

1.6 Symplectic matrices

Remark 1.6.1 [Symplectic matrices]
Poisson brackets with the quadratic monomials on phase space X_1, X_2, X_3 correspond respectively to multiplication by the traceless constant matrices m_1, m_2, m_3. In turn, exponentiation of these traceless constant matrices leads to the corresponding matrices $M_1(\tau_1), M_2(\tau_2), M_3(\tau_3)$ in equations (1.5.11), (1.5.14) and (1.5.15). The latter are 2×2 ***symplectic matrices***. That is, these three matrices each satisfy

$$M_i(\tau_i) J M_i(\tau_i)^T = J \quad \text{(no sum on } i = 1, 2, 3\text{)}, \tag{1.6.1}$$

where

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} . \tag{1.6.2}$$

By their construction from the axisymmetric invariants X_1, X_2, X_3, each of the symplectic matrices $M_1(\tau_1), M_2(\tau_2), M_3(\tau_3)$ preserves the cross product $S = \mathbf{p} \times \mathbf{q}$.

Definition 1.6.2 (Lie transformation groups)

- *A **transformation** is a one-to-one mapping of a set onto itself.*
- *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 set of parameters.*

Theorem 1.6.3 (Symplectic group $Sp(2,\mathbb{R})$)
Under matrix multiplication, the set of 2×2 symplectic matrices forms a group.

> **Exercise.** Prove that the matrices $M_1(\tau_1), M_2(\tau_2), M_3(\tau_3)$ defined above all satisfy the defining relation (1.6.1) required to be symplectic. Prove that these matrices form a group under matrix multiplication. Conclude that they form a three-parameter Lie group. ★

Theorem 1.6.4 (Fundamental theorem of planar optics)
Any plane paraxial optical system, represented by a 2×2 symplectic matrix $M \in Sp(2,\mathbb{R})$ may be factored into subsystems consisting of products of three subgroups of the symplectic group, as

$$M = \begin{pmatrix} \cos\omega & \sin\omega \\ -\sin\omega & \cos\omega \end{pmatrix} \begin{pmatrix} e^{\tau_3} & 0 \\ 0 & e^{-\tau_3} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2\tau_1 & 1 \end{pmatrix} . \tag{1.6.3}$$

*This is a general result, called the **Iwasawa decomposition** of the symplectic matrix group, usually written as [Ge1961]*

$$Sp(2,\mathbb{R}) = \mathsf{KAN}\,. \tag{1.6.4}$$

*The rightmost matrix factor (nilpotent subgroup N) corresponds to a **thin lens**, whose parameter $2\tau_1$ is called its **Gaussian power** [Wo2004]. This factor does not affect the image at all, since $\mathbf{q}(\tau_1) = \mathbf{q}(0)$ from equation (1.5.11). However, the rightmost factor does change the direction of the rays that fall on each point of the screen. The middle factor (Abelian subgroup A) magnifies the image by the factor e^{τ_3}, while **squeezing** the light so that the product $\mathbf{q}\cdot\mathbf{p}$ remains the invariant as in equation (1.5.15). The leftmost factor (the maximal compact subgroup K) is a type of Fourier transform in angle $\omega \in S^1$ on a circle as in equation (1.5.6).*

For insightful discussions and references to the literature in the design and analysis of optical systems using the symplectic matrix approach, see, e.g., [Wo2004]. For many extensions of these ideas with applications to charged-particle beams, see [Dr2007].

Definition 1.6.5 (Hamiltonian matrices)
The traceless constant matrices

$$m_1 = \begin{pmatrix} 0 & 0 \\ -2 & 0 \end{pmatrix}, \; m_2 = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}, \; m_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad (1.6.5)$$

whose ***exponentiation*** *defines the $Sp(2,\mathbb{R})$ symplectic matrices*

$$e^{\tau_1 m_1} = M_1(\tau_1)\,, \quad e^{\tau_2 m_2} = M_2(\tau_2)\,, \quad e^{\tau_3 m_3} = M_3(\tau_3)\,, \qquad (1.6.6)$$

and which are the ***tangent vectors at their respective identity transformations****,*

$$\begin{aligned} m_1 &= \Big[M\,'_1(\tau_1) M_1^{-1}(\tau_1) \Big]_{\tau_1=0}\,, \\ m_2 &= \Big[M\,'_2(\tau_2) M_2^{-1}(\tau_2) \Big]_{\tau_2=0}\,, \\ m_3 &= \Big[M\,'_3(\tau_3) M_3^{-1}(\tau_3) \Big]_{\tau_3=0}\,, \end{aligned} \qquad (1.6.7)$$

are called ***Hamiltonian matrices****.*

Remark 1.6.6

- From their definitions, the Hamiltonian matrices m_i with $i = 1, 2, 3$, each satisfy

$$J m_i + m_i^T J = 0\,, \quad \text{where} \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \qquad (1.6.8)$$

  That is, $J m_i = (J m_i)^T$ is a symmetric matrix.

  > **Exercise.** Take the derivative of the definition of symplectic matrices (1.6.1) to prove statement (1.6.8) about Hamiltonian matrices. Verify that the Hamiltonian matrices in (1.6.5) satisfy (1.6.8). ★

- The respective actions of the symplectic matrices $M_1(\tau_1)$, $M_2(\tau_2)$, $M_3(\tau_3)$ in (1.6.6) on the phase space vector $(\mathbf{q},\, \mathbf{p})^T$ are the flows of the Hamiltonian vector fields $\{\,\cdot\,,\, X_1\}$, $\{\,\cdot\,,\, X_2\}$, and $\{\,\cdot\,,\, X_3\}$ corresponding to the axisymmetric invariants X_1, X_2 and X_3 in (1.4.5).

- The quadratic Hamiltonian,

$$\begin{aligned} H &= \frac{\omega}{2}(X_1 + X_2) + \frac{\gamma}{2}(X_2 - X_1) + \tau X_3 \qquad (1.6.9) \\ &= \frac{\omega}{2}(|\mathbf{p}|^2 + |\mathbf{q}|^2) + \frac{\gamma}{2}(|\mathbf{p}|^2 - |\mathbf{q}|^2) + \tau \mathbf{q}\cdot\mathbf{p}\,, \end{aligned}$$

is associated to the Hamiltonian matrix,

$$\begin{aligned} m(\omega,\gamma,\tau) &= \frac{\omega}{2}(m_1 + m_2) + \frac{\gamma}{2}(m_2 - m_1) + \tau m_3 \\ &= \begin{pmatrix} \tau & \gamma+\omega \\ \gamma-\omega & -\tau \end{pmatrix}. \qquad (1.6.10) \end{aligned}$$

The eigenvalues of the Hamiltonian matrix (1.6.10) are determined from

$$\lambda^2 + \Delta = 0\,, \quad \text{with} \quad \Delta = \det m = \omega^2 - \gamma^2 - \tau^2\,. \qquad (1.6.11)$$

Consequently, the eigenvalues come in pairs, given by

$$\lambda^{\pm} = \pm\sqrt{-\Delta} = \pm\sqrt{\tau^2 + \gamma^2 - \omega^2}\,. \qquad (1.6.12)$$

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

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

may alter the values of (ω,γ,τ) in (1.6.10). However, this action preserves eigenvalues, so it preserves the value of the determinant Δ. This means the orbits of the Hamiltonian flows lie of level sets of the determinant Δ.

The Hamiltonian flows corresponding to these eigenvalues change type, depending on whether $\Delta < 0$ (hyperbolic), $\Delta = 0$ (parabolic), or $\Delta > 0$ (elliptic), as illustrated in Figure 1.9 and summarised in the table below, cf. [Wo2004].

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

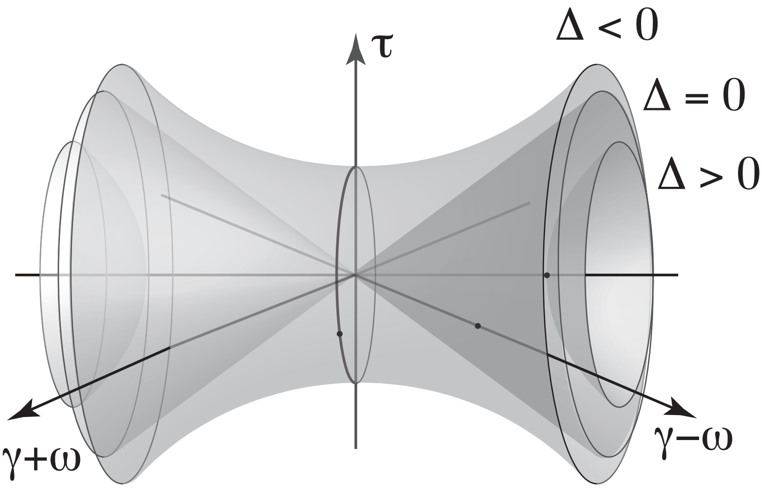


Figure 1.9: The action by matrix conjugation of a symplectic matrix on a Hamiltonian matrix changes its parameters $(\omega, \gamma, \tau) \in \mathbb{R}^3$, while preserving the value of the discriminant $\Delta = \omega^2 - \gamma^2 - \tau^2$. The flows corresponding to exponentiation of the Hamiltonian matrices with parameters $(\omega, \gamma, \tau) \in \mathbb{R}^3$ are divided into three families of orbits defined by the sign of Δ. These three families of orbits are hyperbolic ($\Delta < 0$), parabolic ($\Delta = 0$) and elliptic ($\Delta > 0$).

Remark 1.6.7 [Prelude to Lie algebras]

- In terms of the Hamiltonian matrices the KAN decomposition

(1.6.3) may be written as

$$M = e^{\omega(m_1+m_2)/2} e^{\tau_3 m_3} e^{\tau_1 m_1} . \tag{1.6.13}$$

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

$$[m_1, m_2] = 4m_3 , \quad [m_2, m_3] = -\, 2m_2 , \quad [m_3, m_1] = -\, 2m_1 .$$

The last observation (closure of the commutators) summons the definition of a Lie algebra. For this, we follow [Ol2000].

1.7 Lie algebras

1.7.1 Definitions

Definition 1.7.1 *A **Lie algebra** is a vector space $\mathfrak{g}$ together with a bilinear operation*

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

*called the **Lie bracket** for $\mathfrak{g}$, that satisfies the defining properties:*

(a) Bilinearity, e.g.,

$$[a\mathbf{u} + b\mathbf{v} \,,\, \mathbf{w}] = a[\mathbf{u} \,,\, \mathbf{w}] + b[\mathbf{v} \,,\, \mathbf{w}] ,$$

for constants $(a, b) \in \mathbb{R}$ and any vectors $(\mathbf{u}, \mathbf{v}, \mathbf{w}) \in \mathfrak{g}$;

(b) Skew symmetry

$$[\mathbf{u} \,,\, \mathbf{w}] = -[\mathbf{w} \,,\, \mathbf{u}] ;$$

(c) Jacobi identity

$$[\mathbf{u} \,,\, [\mathbf{v} \,,\, \mathbf{w}]\,] + [\mathbf{v} \,,\, [\mathbf{w} \,,\, \mathbf{u}]\,] + [\mathbf{w} \,,\, [\mathbf{u} \,,\, \mathbf{v}]\,] = 0 ,$$

for all $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in $\mathfrak{g}$.

1.7.2 Structure constants

Suppose $\mathfrak{g}$ is any finite dimensional Lie algebra. The Lie bracket for any choice of basis vectors $\{\mathbf{e}_1, \ldots, \mathbf{e}_r\}$ of $\mathfrak{g}$ must again lie in $\mathfrak{g}$. Thus, constants c_{ij}^k exist $i, j, k = 1, 2, \ldots, r$, called the ***structure constants*** of the Lie algebra $\mathfrak{g}$, such that

$$[\,\mathbf{e}_i\,,\,\mathbf{e}_j\,] = c_{ij}^k \mathbf{e}_k\,. \tag{1.7.1}$$

Since the $\{\mathbf{e}_1, \ldots, \mathbf{e}_r\}$ form a vector basis, the structure constants in (1.7.1) determine the Lie algebra $\mathfrak{g}$ from the bilinearity of the Lie bracket. The conditions of skew symmetry and the Jacobi identity place further constraints on the structure constants. These constraints are:

(i) Skew symmetry

$$c_{ji}^k = -\,c_{ij}^k\,, \tag{1.7.2}$$

and

(ii) Jacobi identity

$$c_{ij}^k c_{lk}^m + c_{li}^k c_{jk}^m + c_{jl}^k c_{ik}^m = 0\,. \tag{1.7.3}$$

Conversely, any set of constants c_{ij}^k that satisfy relations (1.7.2) and (1.7.3) defines a Lie algebra $\mathfrak{g}$.

Exercise. Prove that the Jacobi identity requires the relation (1.7.3). ★

Answer. The Jacobi identity involves summing three terms of the form,

$$[\,\mathbf{e}_l\,,\,[\,\mathbf{e}_i\,,\,\mathbf{e}_j\,]\,] = c_{ij}^k [\,\mathbf{e}_l\,,\,\mathbf{e}_k] = c_{ij}^k c_{lk}^m \mathbf{e}_m\,.$$

Summing over the three cyclic permutations of (l, i, j) of this expression yields the required relation (1.7.3) among the structure constants for the Jacobi identity to hold.

▲

1.7.3 Commutator tables

A convenient way to display the structure of a finite dimensional Lie algebra is to write its commutation relations in tabular form. If $\mathfrak{g}$ is an r-dimensional Lie algebra and $\{\mathbf{e}_1, \ldots, \mathbf{e}_r\}$ forms a basis of $\mathfrak{g}$, then its ***commutator table*** will be the $r \times r$ array whose (i, j)-th entry expresses the Lie bracket $[\,\mathbf{e}_i\,,\,\mathbf{e}_j\,]$. Commutator tables are always antisymmetric since $[\,\mathbf{e}_j\,,\,\mathbf{e}_i\,] = -[\,\mathbf{e}_i\,,\,\mathbf{e}_j\,]$. Hence, the diagonal entries all vanish. The structure constants may be easily read off the commutator table, since c^k_{ij} is the coefficient of $\mathbf{e}_k$ in the (i, j)-th entry of the table.

For example, the commutator table of the Hamiltonian matrices in equation (1.6.7) is given by

$$[m_i,\, m_j] = c^k_{ij} m_k = \begin{array}{|c|ccc|}\hline [\,\cdot\,,\,\cdot\,] & m_1 & m_2 & m_3 \\ \hline m_1 & 0 & 4m_3 & 2m_1 \\ m_2 & -4m_3 & 0 & -2m_2 \\ m_3 & -2m_1 & 2m_2 & 0 \\ \hline \end{array} \tag{1.7.4}$$

The structure constants are immediately read off the table as

$$c^3_{12} = 4 = -c^3_{21}\,, \quad c^1_{13} = c^2_{32} = 2 = -c^2_{23} = -c^1_{31}\,,$$

and all the other c^k_{ij}'s vanish.

Proposition 1.7.2 (Structure constants for $sp(2, \mathbb{R})$)
The commutation relations in (1.6.7) for the 2×2 Hamiltonian matrices define the structure constants for the ***symplectic Lie algebra*** *$sp(2, \mathbb{R})$.*

Proof. The exponentiation of the Hamiltonian matrices was shown in Theorem 1.5.3 of Section 1.5.1 to produce the symplectic Lie group, $Sp(2, \mathbb{R})$. Likewise, the tangent space at the identity of the symplectic Lie group $Sp(2, \mathbb{R})$ is the symplectic Lie algebra, $sp(2, \mathbb{R})$, a vector space whose basis may be chosen as the 2×2 Hamiltonian matrices. Thus, the commutation relations among these matrices yield the structure constants for $sp(2, \mathbb{R})$ in this basis. ∎

1.7.4 Poisson brackets among axisymmetric variables

Theorem 1.7.3 *The canonical Poisson brackets among the axisymmetric variables X_1, X_2 and X_3 in (1.4.5) close among themselves,*

$$\{X_1, X_2\} = 4X_3 \,, \quad \{X_2, X_3\} = -\,2X_2 \,, \quad \{X_3, X_1\} = -\,2X_1 \,.$$

In tabular form, this is

$$\{X_i,\, X_j\} = \begin{array}{|c|ccc|} \hline \{\,\cdot\,,\,\cdot\,\} & X_1 & X_2 & X_3 \\ \hline X_1 & 0 & 4X_3 & 2X_1 \\ X_2 & -4X_3 & 0 & -2X_2 \\ X_3 & -2X_1 & 2X_2 & 0 \\ \hline \end{array} \tag{1.7.5}$$

Proof. The proof is a direct verification using the chain rule for Poisson brackets,

$$\{X_i,\, X_j\} = \frac{\partial X_i}{\partial z_A}\{z_A,\, z_B\}\frac{\partial X_j}{\partial z_A}\,, \tag{1.7.6}$$

for the invariant quadratic monomials $X_i(z_A)$ in (1.4.5). Here one denotes $z_A = (q_A, p_A)$, with $A = 1, 2, 3$. ■

Remark 1.7.4 The closure in the Poisson commutator table (1.7.5) among the set of axisymmetric phase space functions $(X_1,\, X_2,\, X_3)$ is possible, because these functions are all quadratic monomials in the canonical variables. That is, the canonical Poisson bracket preserves the class of quadratic monomials in phase space.

Summary 1.7.5 *The 2×2 traceless matrices m_1, m_2 and m_3 in equation (1.6.7) provide a matrix commutator representation of the Poisson bracket relations in equation (1.7.5) for the quadratic monomials in phase space, X_1, X_2 and X_3, from which the matrices m_1, m_2 and m_3 were derived. Likewise, the 2×2 symplectic matrices $M_1(\tau_1)$, $M_2(\tau_2)$ and $M_3(\tau_3)$ provide a matrix representation of the transformations of the phase space vector $(\mathbf{q}, \mathbf{p})^T$. These transformations are generated by integrating the characteristic equations of the Hamiltonian vector fields $\{\,\cdot\,,\, X_1\}$, $\{\,\cdot\,,\, X_2\}$ and $\{\,\cdot\,,\, X_3\}$.*

1.7.5 Non-canonical $\mathbb{R}^3$ Poisson bracket for ray optics

The canonical Poisson bracket relations in (1.7.5) may be used to transform to another Poisson bracket expressed solely in terms of the variables $\mathbf{X} = (X_1, X_2, X_3) \in \mathbb{R}^3$ by using the chain rule again,

$$\frac{dF}{dt} = \{F, H\} = \frac{\partial F}{\partial X_i} \{X_i,\, X_j\} \frac{\partial H}{\partial X_j} \,. \tag{1.7.7}$$

Here, the quantities $\{X_i,\, X_j\}$, with $i, j = 1, 2, 3$, are obtained from Poisson commutator table in (1.7.5).

This chain rule calculation reveals that the Poisson bracket in the $\mathbb{R}^3$ variables (X_1, X_2, X_3) repeats the commutator table $[m_i,\, m_j] = c_{ij}^k m_k$ for the Lie algebra $sp(2, \mathbb{R})$ of Hamiltonian matrices in (1.7.4). Consequently, we may write this Poisson bracket equivalently as

$$\boxed{\{F, H\} = X_k \, c_{ij}^k \, \frac{\partial F}{\partial X_i} \frac{\partial H}{\partial X_j} \,.} \tag{1.7.8}$$

In particular, the Poisson bracket between two of these quadratic-monomial invariants is a linear function of them

$$\{X_i,\, X_j\} = c_{ij}^k X_k \,, \tag{1.7.9}$$

and we also have

$$\{X_l,\, \{X_i,\, X_j\}\} = c_{ij}^k \{X_l,\, X_k\} = c_{ij}^k c_{lk}^m X_m \,. \tag{1.7.10}$$

Hence, the Jacobi identity is satisfied for the Poisson bracket (1.7.7) as a consequence of

$$\begin{aligned}
&\{X_l,\, \{X_i,\, X_j\}\} + \{X_i,\, \{X_j,\, X_l\}\} + \{X_j,\, \{X_l,\, X_i\}\} \\
&= c_{ij}^k \{X_l,\, X_k\} + c_{jl}^k \{X_i,\, X_k\} + c_{li}^k \{X_j,\, X_k\} \\
&= \Big(c_{ij}^k c_{lk}^m + c_{jl}^k c_{ik}^m + c_{li}^k c_{jk}^m \Big) X_m = 0 \,,
\end{aligned}$$

followed by comparison with equation (1.7.3) for the Jacobi identity in terms of the structure constants.

Remark 1.7.6 This calculation for the Poisson bracket (1.7.8) provides an independent proof that it satisfies the Jacobi identity.

The chain rule calculation (1.7.7) also reveals the following.

Theorem 1.7.7 *Under the map*

$$T^*\mathbb{R}^2 \to \mathbb{R}^3 : (\mathbf{q}, \mathbf{p}) \to \mathbf{X} = (X_1, X_2, X_3)\,, \tag{1.7.11}$$

the Poisson bracket among the axisymmetric optical variables (1.4.5)

$$X_1 = |\mathbf{q}|^2 \geq 0\,, \quad X_2 = |\mathbf{p}|^2 \geq 0\,, \quad X_3 = \mathbf{p}\cdot\mathbf{q}\,,$$

may be expressed for $S^2 = X_1X_2 - X_3^2$ as

$$\begin{aligned} \frac{dF}{dt} = \{F, H\} &= \nabla F \cdot \nabla S^2 \times \nabla H \\ &= -\frac{\partial S^2}{\partial X_l}\epsilon_{ljk}\frac{\partial F}{\partial X_j}\,\frac{\partial H}{\partial X_k}\,. \end{aligned} \tag{1.7.12}$$

Proof. This is a direct verification using formula (1.7.7). For example,

$$2\epsilon_{123}\frac{\partial S^2}{\partial X_3} = -4X_3, \quad 2\epsilon_{132}\frac{\partial S^2}{\partial X_2} = 2X_1, \quad 2\epsilon_{231}\frac{\partial S^2}{\partial X_1} = -2X_2\,.$$

(The inessential factors of 2 may be absorbed into the definition of the independent variable, which here is the time, t.) ■

The standard symbol ϵ_{klm} used in the last relation in (1.7.12) to write the triple scalar product of vectors in index form is defined as follows.

Definition 1.7.8 (Antisymmetric symbol ϵ_{klm})
The symbol ϵ_{klm} with $\epsilon_{123} = 1$ is the totally antisymmetric tensor in three dimensions: it vanishes if any of its indices are repeated and it equals the parity of the permutations of the set $\{1, 2, 3\}$ when $\{k, l, m\}$ are all different. That is,

$$\epsilon_{kkm} = 0 \text{ (no sum)}$$

and

$$\epsilon_{klm} = +1 \text{ (resp. } -1\text{) for even (resp. odd) permutations of } \{1, 2, 3\}\,.$$

Remark 1.7.9 For three-dimensional vectors **A**, **B**, **C**, one has

$$(\mathbf{B}\times\mathbf{C})_k = \epsilon_{klm}B_lC_m \quad \text{and} \quad \big(\mathbf{A}\times(\mathbf{B}\times\mathbf{C})\big)_i = \epsilon_{ijk}A_j\epsilon_{klm}B_lC_m$$

Hence, the relation

$$\epsilon_{ijk}\epsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$$

verifies the familiar BAC minus CAB rule for the triple vector product. That is,

$$\mathbf{A}\times(\mathbf{B}\times\mathbf{C}) = \mathbf{B}(\mathbf{A}\cdot\mathbf{C}) - \mathbf{C}(\mathbf{A}\cdot\mathbf{B})\,.$$

Corollary 1.7.10 *The equations of Hamiltonian ray optics in axisymmetric translation-invariant media may be expressed with $H = H(X_1, X_2)$ as*

$$\dot{\mathbf{X}} = \nabla S^2 \times \nabla H\,, \quad \textit{with} \quad S^2 = X_1X_2 - X_3^2 \geq 0\,. \tag{1.7.13}$$

Thus, the flow preserves volume (that is, it satisfies div $\dot{\mathbf{X}} = 0$) and the evolution along the curve $\mathbf{X}(z) \in \mathbb{R}^3$ takes place on intersections of level surfaces of the axisymmetric media invariants S^2 and $H(X_1, X_2)$ in $\mathbb{R}^3$.

Remark 1.7.11 The Petzval invariant S^2 satisfies $\{S^2, H\} = 0$ with the bracket (1.7.12) for every Hamiltonian $H(X_1, X_2, X_3)$ expressed in these variables.

Definition 1.7.12 (Casimir, or distinguished function)
*A function that Poisson-commutes with all other functions on a certain space is the Poisson bracket's **Casimir**, or **distinguished function**.*

1.8 Equilibrium solutions

1.8.1 Energy-Casimir stability

Remark 1.8.1 [Critical energy plus Casimir equilibria]
A point of tangency of the level sets of Hamiltonian H and Casimir S^2 is an equilibrium solution of equation (1.7.13). This is because, at

such a point, the gradients of the Hamiltonian H and Casimir S^2 are collinear; so the right-hand side of (1.7.13) vanishes. At such points of tangency, the variation of the sum $H_\Phi = H + \Phi(S^2)$ vanishes, for some smooth function Φ. That is,

$$\begin{aligned} \delta H_\Phi(\mathbf{X}_e) &= DH_\Phi(\mathbf{X}_e) \cdot \delta\mathbf{X} \\ &= \Big[\nabla H + \Phi'(S^2)\nabla S^2\Big]_{\mathbf{X}_e} \cdot \delta\mathbf{X} = 0\,, \end{aligned}$$

when evaluated at equilibrium points $\mathbf{X}_e$ where the level sets of H and S^2 are tangent.

> **Exercise.** Show that a point $\mathbf{X}_e$ at which H_Φ has a critical point (i.e., $\delta H_\Phi = 0$) must be an equilibrium solution of equation (1.7.13). ★

Energy-Casimir stability of equilibria

The second variation of the sum $H_\Phi = H + \Phi(S^2)$ is a quadratic form in $\mathbb{R}^3$ given by

$$\delta^2 H_\Phi(\mathbf{X}_e) = \delta\mathbf{X} \cdot D^2 H_\Phi(\mathbf{X}_e) \cdot \delta\mathbf{X}\,.$$

Thus we have, by Taylor's theorem,

$$H_\Phi(\mathbf{X}_e + \delta\mathbf{X}) - H_\Phi(\mathbf{X}_e) = \frac{1}{2}\delta^2 H_\Phi(\mathbf{X}_e) + o(|\delta\mathbf{X}|^2)\,,$$

when evaluated at the critical point $\mathbf{X}_e$. Remarkably, the quadratic form $\delta^2 H_\Phi(\mathbf{X}_e)$ is the Hamiltonian for the dynamics linearised around the critical point. Consequently, the second variation $\delta^2 H_\Phi$ is preserved by the linearised dynamics in a neighbourhood of the equilibrium point.

> **Exercise.**
> • Linearise the dynamical equation (1.7.13) about an equilibrium $\mathbf{X}_e$ for which the quantity H_Φ has a critical point and show that the linearised dynamics conserves the quadratic form arising from the second variation.

• Show that the quadratic form is the Hamiltonian for the linearised dynamics.
• What is the corresponding Poisson bracket?
• Does this process provides a proper bracket for the linearised dynamics? Prove that it does. ★

The ***signature*** of the second variation provides a method for determining the stability of the critical point. This is the ***energy-Casimir stability method***. This method is based on the following.

Theorem 1.8.2 *A critical point $\mathbf{X}_e$ of $H_\Phi = H + \Phi(S^2)$ whose second variation is definite in sign is a stable equilibrium solution of equation (1.7.13).*

Proof. A critical point $\mathbf{X}_e$ of $H_\Phi = H + \Phi(S^2)$ is an equilibrium solution of equation (1.7.13). Sign definiteness of the second variation provides a norm $\|\delta\mathbf{X}\|^2 = |\delta^2 H_\Phi(\mathbf{X}_e)|$ for the perturbations around the equilibrium $\mathbf{X}_e$ that is conserved by the linearised dynamics. Being conserved by the dynamics linearised around the equilibrium, this sign-definite distance from $\mathbf{X}_e$ must remain constant. Therefore, in this case, the absolute value of sign-definite second variation $|\delta^2 H_\Phi(\mathbf{X}_e)|$ provides a distance from the equilibrium $\|\delta\mathbf{X}\|^2$ which is bounded in time under the linearised dynamics. Hence, the equilibrium solution is stable.

■

Remark 1.8.3 Even when the second variation is indefinite, it is still linearly conserved. However, an indefinite second variation does not provide a norm for the perturbations. Consequently, an indefinite second variation does not limit the growth of a perturbation away from its equilibrium.

Definition 1.8.4 (Geometrical nature of equilibria)

*An equilibrium whose second variation is sign-definite are called **elliptic**, because the level sets of the second variation in this case make closed, nearly elliptical contours in its Euclidean neighbourhood. Hence, the orbits on these closed level sets remain near the equilibria in the sense of the Euclidean norm on $\mathbb{R}^3$. (In $\mathbb{R}^3$ all norms are equivalent to the Euclidean norm.)*

*An equilibrium with sign-indefinite second variation is called **hyperbolic**, because the level sets of the second variation do not close locally in its Euclidean neighbourhood. Hence, in this case, an initial perturbation following a hyperbolic level set of the second variation may move out of the Euclidean neighbourhood of the equilibrium.*

1.9 Momentum maps

1.9.1 The action of $Sp(2,\mathbb{R})$ on $T^*\mathbb{R}^2 \simeq \mathbb{R}^2 \times \mathbb{R}^2$

The Lie group $Sp(2,\mathbb{R})$ of symplectic real matrices $M(s)$ acts diagonally on $\mathbf{z} = (\mathbf{q},\,\mathbf{p})^T \in T^*\mathbb{R}^2$ by matrix multiplication as

$$\mathbf{z}(s) = M(s)\mathbf{z}(0) = \exp(s\xi)\mathbf{z}(0)\,,$$

in which $M(s)JM^T(s) = J$ is a symplectic 2×2 matrix. The 2×2 matrix tangent to the symplectic matrix $M(s)$ at the identity $s=0$ is given by

$$\xi = \Big[M'(s)M^{-1}(s)\Big]_{s=0}\,.$$

This is a 2×2 Hamiltonian matrix in $sp(2,\mathbb{R})$, satisfying (1.6.8) as

$$J\xi + \xi^T J = 0 \quad \text{so that} \quad J\xi = (J\xi)^T\,. \tag{1.9.1}$$

That is, for $\xi \in sp(2,\mathbb{R})$, the matrix $J\xi$ is symmetric.

> **Exercise.** Verify (1.9.1), cf. (1.6.8). What is the corresponding formula for $\zeta = [M^{-1}(s)M'(s)]_{s=0}$?
>
> ★

The vector field $\xi_M(\mathbf{z}) \in T\mathbb{R}^2$ may be expressed as a derivative,

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

in which the diagonal action ($\xi\mathbf{z}$) of the Hamiltonian matrix (ξ) and the 2-component real multi-vector $\mathbf{z} = (\mathbf{q},\,\mathbf{p})^T$ has components given by $(\xi_{kl}q_l, \xi_{kl}p_l)^T$, with $k, l = 1, 2$. The matrix ξ is any linear combination of the traceless constant Hamiltonian matrices (1.6.5).

Definition 1.9.1 (Map $\mathcal{J} : T^*\mathbb{R}^2 \simeq \mathbb{R}^2 \times \mathbb{R}^2 \to sp(2,\mathbb{R})^*$)
The map, $\mathcal{J} : T^*\mathbb{R}^2 \simeq \mathbb{R}^2 \times \mathbb{R}^2 \to sp(2,\mathbb{R})^*$ is defined by

$$\begin{aligned}
\mathcal{J}^\xi(\mathbf{z}) &:= \Big\langle \mathcal{J}(\mathbf{z}),\,\xi\Big\rangle_{sp(2,\mathbb{R})^*\times sp(2,\mathbb{R})} \\
&= \Big(\mathbf{z},\,J\xi\mathbf{z}\Big)_{\mathbb{R}^2\times\mathbb{R}^2} \\
&:= z_k(J\xi)_{kl}z_l \\
&= \mathbf{z}^T\cdot J\xi\mathbf{z} \\
&= \mathrm{tr}\Big((\mathbf{z}\otimes\mathbf{z}^T J)\xi\Big)\,,
\end{aligned} \tag{1.9.2}$$

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

Remark 1.9.2 The map $\mathcal{J}(\mathbf{z})$ given in (1.9.2) by

$$\mathcal{J}(\mathbf{z}) = (\mathbf{z}\otimes\mathbf{z}^T J) \in sp(2,\mathbb{R})^*\,, \tag{1.9.3}$$

sends $\mathbf{z} = (\mathbf{q},\,\mathbf{p})^T \in \mathbb{R}^2 \times \mathbb{R}^2$ to $\mathcal{J}(\mathbf{z}) = (\mathbf{z}\otimes\mathbf{z}^T J)$, which is an element of $sp(2,\mathbb{R})^*$, the dual space to $sp(2,\mathbb{R})$. Under the pairing $\langle\,\cdot\,,\,\cdot\,\rangle : sp(2,\mathbb{R})^* \times sp(2,\mathbb{R}) \to \mathbb{R}$ given by the trace of the matrix product, one finds the Hamiltonian, or phase space function,

$$\Big\langle \mathcal{J}(\mathbf{z}),\,\xi\Big\rangle = \mathrm{tr}\left(\mathcal{J}(\mathbf{z})\,\xi\right)\,, \tag{1.9.4}$$

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

Remark 1.9.3 [Map to axisymmetric invariant variables]
The map, $\mathcal{J} : T^*\mathbb{R}^2 \simeq \mathbb{R}^2 \times \mathbb{R}^2 \to sp(2,\mathbb{R})^*$ in (1.9.2) for $Sp(2,\mathbb{R})$

acting diagonally on $\mathbb{R}^2 \times \mathbb{R}^2$ in equation (1.9.3) may be expressed in matrix form as

$$
\begin{aligned}
\mathcal{J} &= (\mathbf{z} \otimes \mathbf{z}^T J) \\
&= 2 \begin{pmatrix} \mathbf{p}\cdot\mathbf{q} & -|\mathbf{q}|^2 \\ |\mathbf{p}|^2 & -\mathbf{p}\cdot\mathbf{q} \end{pmatrix} \\
&= 2 \begin{pmatrix} X_3 & -X_1 \\ X_2 & -X_3 \end{pmatrix} .
\end{aligned} \tag{1.9.5}
$$

This is none other than matrix form of the map (1.7.11) to axisymmetric invariant variables.

$$
T^*\mathbb{R}^2 \to \mathbb{R}^3 : \ (\mathbf{q}, \mathbf{p})^T \to \mathbf{X} = (X_1, X_2, X_3) \,,
$$

defined as

$$
X_1 = |\mathbf{q}|^2 \geq 0 \,, \quad X_2 = |\mathbf{p}|^2 \geq 0 \,, \quad X_3 = \mathbf{p}\cdot\mathbf{q} \,. \tag{1.9.6}
$$

Applying the momentum map $\mathcal{J}$ to the vector of Hamiltonian matrices $\mathbf{m} = (m_1,\, m_2,\, m_3)$ in equation (1.6.5) yields the individual components,

$$
\mathcal{J}\cdot\mathbf{m} = 2\mathbf{X} \quad \Longleftrightarrow \quad \mathbf{X} = \frac{1}{2}\, z_k (J\mathbf{m})_{kl} z_l \,. \tag{1.9.7}
$$

Thus, the map, $\mathcal{J} : T^*\mathbb{R}^2 \simeq \mathbb{R}^2 \times \mathbb{R}^2 \to sp(2,\mathbb{R})^*$ recovers the components of the vector $\mathbf{X} = (X_1,\, X_2,\, X_3)$ that are related to the components of the Petzval invariant by $S^2 = X_1 X_2 - X_3^2$.

Exercise. Verify equation (1.9.7) explicitly by computing, for example,

$$
\begin{aligned}
X_1 &= \frac{1}{2} (\mathbf{q},\, \mathbf{p}) \cdot (J m_1) \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} \\
&= \frac{1}{2} (\mathbf{q},\, \mathbf{p}) \cdot \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ -2 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} \\
&= |\mathbf{q}|^2 \,.
\end{aligned}
$$

★

Remark 1.9.4 [Momentum maps for ray optics]
Our previous discussions have revealed that the axisymmetric variables (X_1, X_2, X_3) in (1.9.6) generate the Lie group of symplectic transformations (1.6.3) as flows of Hamiltonian vector fields. It turns out that this result is connected to the theory of ***momentum maps***. Momentum maps take phase space coordinates $(\mathbf{q},\, \mathbf{p})$ to the space of Hamiltonians whose flows are canonical transformations of phase space. An example of a momentum map already appeared in Definition 1.3.20.

The Hamiltonian functions for the one-parameter subgroups of the symplectic group $Sp(2,\mathbb{R})$ in the KAN decomposition (1.6.13) are given by

$$H_K = \frac{1}{2}(X_1 + X_2)\,, \quad H_A = X_3 \quad \text{and} \quad H_N = -\,X_1\,. \tag{1.9.8}$$

The three phase space functions,

$$H_K = \frac{1}{2}(|\mathbf{q}|^2 + |\mathbf{p}|^2)\,, \quad H_A = \mathbf{q}\cdot\mathbf{p}\,, \quad H_N = -\,|\mathbf{q}|^2\,, \tag{1.9.9}$$

map the phase space $(\mathbf{q},\, \mathbf{p})$ to these Hamiltonians whose corresponding Poisson brackets are the Hamiltonian vector fields for the corresponding one-parameter subgroups. These three Hamiltonians and, equally well, any other linear combinations of (X_1, X_2, X_3), arise from a single ***momentum map***, as we shall explain in Section 1.9.2.

Remark 1.9.5 ***Momentum maps are Poisson maps***. That is, they map Poisson brackets on phase space into Poisson brackets on the target space.

The corresponding Lie algebra product in $sp(2,\mathbb{R})$ was identified using Theorem 1.7.7 with the vector cross product in the space $\mathbb{R}^3$ by using the $\mathbb{R}^3$-bracket. The $\mathbb{R}^3$-brackets among the (X_1, X_2, X_3) closed among themselves. Therefore, as expected, the momentum map was found to be Poisson. In general, when the Poisson bracket relations are all linear, they will be Lie-Poisson brackets, defined below in Section 1.10.1.

1.9.2 Summary: Properties of momentum maps

A momentum map takes phase space coordinates $(\mathbf{q},\, \mathbf{p})$ to the space of Hamiltonians, whose flows are canonical transformations of phase space. The ingredients of the momentum map are: (i) a representation of the infinitesimal action of the Lie algebra of the transformation group on the coordinate space; and (ii) an appropriate pairing with the conjugate momentum space. For example, one may construct a momentum map by using the familiar pairing $\langle\!\langle\, \cdot\,,\,\cdot\,\rangle\!\rangle$ between momentum in phase space and the velocity in the tangent space of the configuration manifold as it appears in the Legendre transformation. For this pairing, the momentum map is derived from the cotangent lift of the infinitesimal action $\xi_M(\mathbf{q})$ of the Lie algebra of the transformation group on the configuration manifold to its action on the canonical momentum. In this case, the formula for the momentum map $\mathcal{J}(\mathbf{q},\, \mathbf{p})$ is

$$\mathcal{J}^\xi(\mathbf{q},\, \mathbf{p}) = \Big\langle \mathcal{J}(\mathbf{q},\, \mathbf{p}),\, \xi \Big\rangle = \Big\langle\!\!\Big\langle \mathbf{p},\, \xi_M(\mathbf{q}) \Big\rangle\!\!\Big\rangle\,, \tag{1.9.10}$$

in which the other pairing $\langle\, \cdot\,,\,\cdot\,\rangle$ is between the Lie algebra and its dual. This means the momentum map $\mathcal{J}$ for the Hamiltonian $\mathcal{J}^\xi$ lives in the dual space of the Lie algebra belonging to the Lie symmetry. The flow of its vector field $X_{\mathcal{J}^\xi} = \{\,\cdot\,,\, \mathcal{J}^\xi\}$ is the transformation of phase space by the cotangent lift of a Lie group symmetry infinitesimally generated for configuration space by $\xi_M(\mathbf{q})$. The computation of the Lagrange invariant S in (1.3.25) is an example of this type of momentum map.

Not all momentum maps arise as cotangent lifts. Momentum maps may also arise from the infinitesimal action of the Lie algebra on the phase space manifold $\xi_{T^*M}(\mathbf{z})$ with $\mathbf{z} = (\mathbf{q},\, \mathbf{p})$ by using the pairing with the symplectic form. The formula for the momentum map is then

$$\mathcal{J}^{\xi}(\mathbf{z}) = \Big\langle \mathcal{J}(\mathbf{z}),\, \xi \Big\rangle = \Big(\mathbf{z},\, J\xi_{T^*M}(\mathbf{z})\Big)\,, \tag{1.9.11}$$

where J is the symplectic form and $(\,\cdot\,,\,\cdot\,)$ is the inner product on phase space $T^*\mathbb{R} \simeq \mathbb{R}\times\mathbb{R}$ for n degrees of freedom. The transformation to axisymmetric variables in (1.9.5) is an example of a momentum map obtained from the symplectic pairing. Both of these approaches are useful and we have seen that both types of momentum maps are summoned when reduction by S^1 axisymmetry is applied in ray optics. The present chapter explores the consequences of S^1 symmetry and the reductions of phase space associated with the momentum maps for this symmetry.

The level sets of the momentum maps provide the geometrical setting for dynamics with symmetry. The components of the momentum map live on the dual of the Lie symmetry algebra, which is a linear space. The level sets of the components of the momentum map provide the natural coordinates for the reduced dynamics. Thus, the motion takes place in a ***reduced space*** whose coordinates are invariant under the original S^1 symmetry. The motion in the reduced space lies on a level set of the momentum map for the S^1 symmetry. It also lies on a level set of the Hamiltonian. Hence, the dynamics in the reduced space of coordinates that are invariant under the S^1 symmetry is confined to the intersections in the reduced space of the level sets of the Hamiltonian and the momentum map associated with that symmetry. Moreover, in most cases, restriction to either level set results in symplectic (canonical) dynamics.

After the solution for this S^1-reduced motion is determined, one must reconstruct the phase associated with the S^1 symmetry, which decouples from the dynamics of the rest of system through the process of reduction. Thus, each point on the manifolds defined by the level sets of the Hamiltonian and the momentum map in the reduced space is associated with an orbit of the phase on S^1. This S^1 phase must be reconstructed from the solution on the reduced space of S^1-invariant functions. The reconstruction of the phase is of interest in its own right, because it contains both geometric and dynamic components, as discussed in Section 1.12.2.

One advantage of this geometric setting is that it readily reveals how ***bifurcations*** arise under changes of parameters in the system, for example, under changes in parameters in the Hamiltonian. In this setting, bifurcations are topological transitions in the intersections of level surfaces of ***orbit manifolds*** of the Hamiltonian and momentum map. The motion proceeds along these intersections in the reduced space whose points are defined by S^1-invariant coordinates. These topological changes in the intersections of the orbit manifolds accompany qualitative changes in the solution behaviour, such as the change of stability of an equilibrium, or the creation or destruction of equilibria. The display of these changes of topology in the reduced space of S^1-invariant functions also allows a visual classification of potential bifurcations. That is, it affords an opportunity to organise the ***choreography of bifurcations*** that are available to the system as its parameters are varied. For an example of this type of geometric bifurcation analysis, see Section 4.5.5.

Remark 1.9.6 The two results:
(1) that the action of a Lie group G with Lie algebra $\mathfrak{g}$ on a symplectic manifold P should be accompanied by a momentum map $J : P \to \mathfrak{g}^*$; and
(2) that the orbits of this action are themselves symplectic manifolds, both occur already in [Lie1890]. See [We1983] for an interesting discussion of Lie's contributions to the theory of momentum maps.

The reader should consult [MaRa1994, OrRa2004] for more discussions and additional examples of momentum maps.

1.10 Lie-Poisson brackets

1.10.1 The $\mathbb{R}^3$-bracket for ray optics is Lie-Poisson

The Casimir invariant $S^2 = X_1X_2 - X_3^2$ for the $\mathbb{R}^3$-bracket (1.7.12) is quadratic. In such cases, one may write the Poisson bracket on $\mathbb{R}^3$ in the suggestive form with a pairing $\langle \cdot\,,\,\cdot \rangle$,

$$\{F, H\} = -\, X_k\, c^k_{ij} \frac{\partial F}{\partial X_i} \frac{\partial H}{\partial X_j} =: -\left\langle \mathbf{X}\,,\, \left[\frac{\partial F}{\partial \mathbf{X}}\,,\, \frac{\partial H}{\partial \mathbf{X}} \right] \right\rangle, \qquad (1.10.1)$$

where c_{ij}^k with $i, j, k = 1, 2, 3$ are the structure constants of a three-dimensional Lie algebra operation denoted as $[\,\cdot\,,\,\cdot\,]$. In the particular case of ray optics, $c_{12}^3 = 4$, $c_{23}^2 = 2$, $c_{31}^1 = 2$ and the rest of the structure constants either vanish, or are obtained from antisymmetry of c_{ij}^k under exchange of any pair of its indices. These values are the structure constants of the 2×2 Hamiltonian matrices (1.6.5), which represent any of the Lie algebras $sp(2, \mathbb{R})$, $so(2, 1)$, $su(1, 1)$, or $sl(2, \mathbb{R})$. Thus, the reduced description of Hamiltonian ray optics in terms of axisymmetric $\mathbb{R}^3$ variables may be defined on the dual space of any of these Lie algebras, say, $sp(2, \mathbb{R})^*$ for definiteness, where duality is defined by pairing $\langle\,\cdot\,,\,\cdot\,\rangle$ in $\mathbb{R}^3$ (contraction of indices). Since $\mathbb{R}^3$ is dual to itself under this pairing, upper and lower indices are equivalent.

Definition 1.10.1 (Lie-Poisson bracket)
*A **Lie-Poisson bracket** is a bracket operation defined as a linear functional of a Lie algebra bracket by a real-valued pairing between a Lie algebra and its dual space.*

Remark 1.10.2 Equation (1.10.1) defines a Lie-Poisson bracket. Being a linear functional of an operation (the Lie bracket $[\,\cdot\,,\,\cdot\,]$) which satisfies the Jacobi identity, any Lie-Poisson bracket also satisfies the Jacobi identity.

1.10.2 Lie-Poisson brackets with quadratic Casimirs

An interesting class of Lie-Poisson brackets emerges from the $\mathbb{R}^3$ Poisson bracket,

$$\{F, H\}_C := -\nabla C \cdot \nabla F \times \nabla H\,, \tag{1.10.2}$$

when its ***Casimir gradient is the linear form on $\mathbb{R}^3$ given by $\nabla C = \mathsf{K}\mathbf{X}$*** associated with the 3×3 symmetric matrix $\mathsf{K}^T = \mathsf{K}$. This bracket may be written equivalently in various notations, including index form, $\mathbb{R}^3$ vector form, and Lie-Poisson form, as

$$\begin{aligned}\{F,\,H\}_{\mathsf{K}} &= -\nabla C\cdot\nabla F\times\nabla H\\ &= -X_l\mathsf{K}^{li}\epsilon_{ijk}\frac{\partial F}{\partial X_j}\frac{\partial H}{\partial X_k}\\ &= -\mathbf{X}\cdot\mathsf{K}\left(\frac{\partial F}{\partial\mathbf{X}}\times\frac{\partial H}{\partial\mathbf{X}}\right)\\ &=: -\left\langle\mathbf{X}\,,\,\left[\frac{\partial F}{\partial\mathbf{X}}\,,\,\frac{\partial H}{\partial\mathbf{X}}\right]_{\mathsf{K}}\right\rangle .\end{aligned}\tag{1.10.3}$$

Remark 1.10.3 The triple scalar product of gradients in the $\mathbb{R}^3$-bracket (1.10.2) is the determinant of the Jacobian matrix for the transformation $(X_1, X_2, X_3)\to(C, F, H)$, which is known to satisfy the Jacobi identity. Being a special case, the Poisson bracket $\{F,\,H\}_{\mathsf{K}}$ also satisfies the Jacobi identity.

In terms of the $\mathbb{R}^3$ components, the Poisson bracket (1.10.3) yields

$$\{X_j,\,X_k\}_{\mathsf{K}} = -X_l\mathsf{K}^{li}\epsilon_{ijk}\,.\tag{1.10.4}$$

The Lie-Poisson form in (1.10.3) associates the $\mathbb{R}^3$ bracket to a Lie algebra with structure constants given in the dual vector basis by

$$[\mathbf{e}_j,\,\mathbf{e}_k]_{\mathsf{K}} = \mathbf{e}_l\mathsf{K}^{li}\epsilon_{ijk} =: \mathbf{e}_l c^{\,l}_{jk}\,.\tag{1.10.5}$$

The Lie group belonging to this Lie algebra is the invariance group of the quadratic Casimir. Namely, it is the transformation group G_{K} with elements $O(s)\in G_{\mathsf{K}}$ with $O(t)|_{t=0}=Id$ whose action from the left on $\mathbb{R}^3$ is given by $\mathbf{X}\to O\mathbf{X}$, such that

$$O^T(t)\mathsf{K}O(t) = \mathsf{K}\tag{1.10.6}$$

or, equivalently,

$$\mathsf{K}^{-1}O^T(t)\mathsf{K} = O^{-1}(t)\,,\tag{1.10.7}$$

for the 3×3 symmetric matrix $\mathsf{K}^T=\mathsf{K}$.

Definition 1.10.4 *An $n \times n$ **orthogonal** matrix $O(t)$ satisfies*

$$O^T(t)\mathrm{Id}\, O(t) = \mathrm{Id} \qquad (1.10.8)$$

in which Id *is the $n \times n$ identity matrix. These matrices represent the orthogonal Lie group in n dimensions, denoted $O(n)$.*

Definition 1.10.5 (Quasi-orthogonal matrices)
*A matrix $O(t)$ satisfying (1.10.6) is called **quasi-orthogonal** with respect to the symmetric matrix K.*

The quasi-orthogonal transformations $\mathbf{X} \to O\mathbf{X}$ are *not* orthogonal, unless $\mathsf{K} = \mathrm{Id}$. However, they do form a Lie group under matrix multiplication, since for any two of them O_1 and O_2, we have

$$(O_1O_2)^T\mathsf{K}(O_1O_2) = O_2^T(O_1^T\mathsf{K}O_1)O_2 = O_2^T\mathsf{K}O_2 = \mathsf{K}\,. \qquad (1.10.9)$$

The corresponding Lie algebra $\mathfrak{g}_{\mathsf{K}}$ is the derivative of the defining condition of the Lie group (1.10.6), evaluated at the identity. This yields,

$$0 = \big[\dot{O}^TO^{-T}\big]_{t=0}\mathsf{K} + \mathsf{K}\big[O^{-1}\dot{O}\big]_{t=0}\,.$$

Consequently, if $\widehat{X} = [O^{-1}\dot{O}]_{t=0} \in \mathfrak{g}_{\mathsf{K}}$, the quantity $\mathsf{K}\widehat{X}$ is skew. That is,

$$(\mathsf{K}\widehat{X})^T = -\,\mathsf{K}\widehat{X}\,.$$

A vector representation of this skew matrix is provided by the following ***hat map*** from the Lie algebra $\mathfrak{g}_{\mathsf{K}}$ to vectors in $\mathbb{R}^3$,

$$\widehat{\ }: \mathfrak{g}_{\mathsf{K}} \to \mathbb{R}^3 \quad \text{defined by} \quad (\mathsf{K}\widehat{X})_{jk} = -\,X_l\mathsf{K}^{li}\epsilon_{ijk}\,. \qquad (1.10.10)$$

When K is invertible, the hat map $(\,\widehat{\cdot}\,)$ in (1.10.10) is a linear isomorphism. For any vectors $\mathbf{u}$, $\mathbf{v} \in \mathbb{R}^3$ with components u^j, v^k, where $j, k = 1, 2, 3$, one computes

$$\begin{aligned} u^j(\mathsf{K}\widehat{X})_{jk}v^k &= -\,\mathbf{X}\cdot\mathsf{K}(\mathbf{u}\times\mathbf{v}) \\ &=: -\,\mathbf{X}\cdot[\mathbf{u}\,,\,\mathbf{v}]_{\mathsf{K}}\,. \end{aligned}$$

This is the Lie-Poisson bracket for the Lie algebra structure represented on $\mathbb{R}^3$ by the vector product,

$$[\mathbf{u}, \mathbf{v}]_{\mathsf{K}} = \mathsf{K}(\mathbf{u} \times \mathbf{v}) \,. \tag{1.10.11}$$

Thus, the Lie algebra of the Lie group of transformations of $\mathbb{R}^3$ leaving invariant the quadratic form $\frac{1}{2}\mathbf{X}^T \cdot \mathsf{K}\mathbf{X}$ may be identified with the cross product of vectors in $\mathbb{R}^3$ by using the K-pairing instead of the usual dot product. For example, in the case of the Petzval invariant we have

$$\nabla S^2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix} \mathbf{X} \,.$$

Consequently,

$$\mathsf{K} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix},$$

for ray optics, with $\mathbf{X} = (X_1,\, X_2,\, X_3)^T$.

> **Exercise.** Verify that inserting this formula for K into formula (1.10.4) recovers the Lie-Poisson bracket relations (1.7.5) for ray optics (up to an inessential constant). ★

Hence, we have proved the following theorem.

Theorem 1.10.6 *Consider the $\mathbb{R}^3$ bracket in equation (1.10.3)*

$$\{F, H\}_{\mathsf{K}} := -\nabla C_{\mathsf{K}} \cdot \nabla F \times \nabla H \quad \textit{with} \quad C_{\mathsf{K}} = \frac{1}{2}\mathbf{X} \cdot \mathsf{K}\mathbf{X} \,, \tag{1.10.12}$$

in which $\mathsf{K}^T = \mathsf{K}$ is a 3×3 real symmetric matrix and $\mathbf{X} \in \mathbb{R}^3$. The quadratic form C_{K} is the Casimir function for the Lie-Poisson bracket given by

$$\{F,\, H\}_{\mathsf{K}} \;=\; -\,\mathbf{X} \cdot \mathsf{K}\left(\frac{\partial F}{\partial \mathbf{X}} \times \frac{\partial H}{\partial \mathbf{X}}\right), \tag{1.10.13}$$

$$=: \; -\left\langle \mathbf{X}\,,\, \left[\frac{\partial F}{\partial \mathbf{X}}\,,\, \frac{\partial H}{\partial \mathbf{X}}\right]_{\mathsf{K}} \right\rangle, \tag{1.10.14}$$

defined on the dual of the three-dimensional Lie algebra $\mathfrak{g}_{\mathsf{K}}$, whose bracket has the following vector product representation for $\mathbf{u},\ \mathbf{v}\in\mathbb{R}^3$,

$$[\,\mathbf{u},\mathbf{v}\,]_{\mathsf{K}} = \mathsf{K}(\mathbf{u}\times\mathbf{v})\,. \tag{1.10.15}$$

This is the Lie algebra bracket for the Lie group G_{K} of transformations of $\mathbb{R}^3$ given by action from the left $\mathbf{X}\to O\mathbf{X}$, such that $O^T\mathsf{K}O=\mathsf{K}$, thereby leaving the quadratic form C_{K} invariant.

Definition 1.10.7 (The ad **and** ad^* **operations)**
*The adjoint (*ad*) and coadjoint (*ad^**) operations are defined for the Lie-Poisson bracket (1.10.14) with quadratic Casimir, $C_{\mathsf{K}}=\frac{1}{2}\,\mathbf{X}\cdot\mathsf{K}\mathbf{X}$, as follows.*

$$\begin{aligned}\langle\mathbf{X}\,,\,[\,\mathbf{u},\mathbf{v}\,]_{\mathsf{K}}\rangle &= \langle\mathbf{X}\,,\,\mathrm{ad}_{\mathbf{u}}\mathbf{v}\rangle = \langle\mathrm{ad}^*_{\mathbf{u}}\mathbf{X}\,,\,\mathbf{v}\rangle \\ &= \mathsf{K}\mathbf{X}\cdot(\mathbf{u}\times\mathbf{v}) = (\mathsf{K}\mathbf{X}\times\mathbf{u})\cdot\mathbf{v}\,.\end{aligned} \tag{1.10.16}$$

Thus, we have explicitly,

$$\mathrm{ad}_{\mathbf{u}}\mathbf{v} = \mathsf{K}(\mathbf{u}\times\mathbf{v}) \quad \textit{and} \quad \mathrm{ad}^*_{\mathbf{u}}\mathbf{X} = -\,\mathbf{u}\times\mathsf{K}\mathbf{X}\,. \tag{1.10.17}$$

These definitions of the ad and ad^* operations yield the following theorem for the dynamics.

Theorem 1.10.8 (Lie-Poisson dynamics)
The Lie-Poisson dynamics (1.10.13) - (1.10.14) is expressed in terms of the ad *and* ad^* *operations by*

$$\begin{aligned}\frac{dF}{dt} = \{F,\,H\}_{\mathsf{K}} &= \left\langle\mathbf{X}\,,\,\mathrm{ad}_{\partial H/\partial\mathbf{X}}\,\frac{\partial F}{\partial\mathbf{X}}\right\rangle \\ &= \left\langle\mathrm{ad}^*_{\partial H/\partial\mathbf{X}}\mathbf{X}\,,\,\frac{\partial F}{\partial\mathbf{X}}\right\rangle,\end{aligned} \tag{1.10.18}$$

so that the Lie-Poisson dynamics expresses itself as coadjoint motion,

$$\frac{d\mathbf{X}}{dt} = \{\mathbf{X},\,H\}_{\mathsf{K}} = \mathrm{ad}^*_{\partial H/\partial\mathbf{X}}\mathbf{X} = -\,\frac{\partial H}{\partial\mathbf{X}}\times\mathsf{K}\mathbf{X}\,. \tag{1.10.19}$$

By construction, this equation conserves the quadratic Casimir, $C_{\mathsf{K}}=\frac{1}{2}\,\mathbf{X}\cdot\mathsf{K}\mathbf{X}$.

Exercise. Write the equations of coadjoint motion (1.10.19) for $\mathsf{K} = \mathrm{diag}(1, 1, 1)$ and $H = X_1^2 - X_3^2/2$. ★

1.11 Divergenceless vector fields

1.11.1 Jacobi identity

One may verify directly that the $\mathbb{R}^3$ bracket in (1.7.12) and in the class of brackets (1.10.12) does indeed satisfy the defining properties of a Poisson bracket. Clearly, it is a bilinear, skew-symmetric form. To show that it is also a Leibnitz operator that satisfies the Jacobi identity, we identify the bracket in (1.7.12) with the following ***divergenceless vector field*** on $\mathbb{R}^3$ defined by

$$X_H = \{\,\cdot\,, H\} = \nabla S^2 \times \nabla H \cdot \nabla \in \mathfrak{X}\,. \tag{1.11.1}$$

This isomorphism identifies the bracket in (1.11.1) acting on functions on $\mathbb{R}^3$ with the action of the divergenceless vector fields $\mathfrak{X}$. It remains to verify the Jacobi identity explicitly, by using the properties of the ***commutator of divergenceless vector fields***.

Definition 1.11.1 (Jacobi-Lie bracket)
*The **commutator** of two divergenceless vector fields $u, v \in \mathfrak{X}$ is defined to be*

$$\big[v,\, w\big] = \big[\mathbf{v}\cdot\nabla,\, \mathbf{w}\cdot\nabla\big] = \Big((\mathbf{v}\cdot\nabla)\mathbf{w} - (\mathbf{w}\cdot\nabla)\mathbf{v}\Big)\cdot\nabla\,. \tag{1.11.2}$$

*The coefficient of the commutator of vector fields is called the **Jacobi-Lie bracket**. It may be written without risk of confusion in the same notation as*

$$\big[v,\, w\big] = (\mathbf{v}\cdot\nabla)\mathbf{w} - (\mathbf{w}\cdot\nabla)\mathbf{v}\,. \tag{1.11.3}$$

In Euclidean vector components, the Jacobi-Lie bracket (1.11.3) is expressed as

$$\big[v,\, w\big]_i = w_{i,j}v_j - v_{i,j}w_j\,. \tag{1.11.4}$$

*Here, a **subscript comma denotes partial derivative**, e.g., $v_{i,j} = \partial v_i/\partial x_j$ and one **sums repeated indices** over their range; for example, $i, j = 1, 2, 3$, in three dimensions.*

Exercise. Show that $[v, w]_{i,i} = 0$ for the expression in (1.11.4); so the commutator of two divergenceless vector fields yields another one. ★

Remark 1.11.2 [Interpreting commutators of vector fields]

We may interpret a smooth vector field in $\mathbb{R}^3$ as the tangent at the identity ($\epsilon = 0$) of a one-parameter flow ϕ_ϵ in $\mathbb{R}^3$ parameterised by $\epsilon \in \mathbb{R}$ and given by integrating

$$\frac{d}{d\epsilon}\bigg|_{\epsilon=0} = v_i(\mathbf{x})\frac{\partial}{\partial x_i}\,. \tag{1.11.5}$$

The characteristic equations of this flow are

$$\frac{dx_i}{d\epsilon} = v_i(\mathbf{x}(\epsilon))\,, \quad \text{so that} \quad \frac{dx_i}{d\epsilon}\bigg|_{\epsilon=0} = v_i(\mathbf{x})\,, \quad i = 1, 2, 3\,. \tag{1.11.6}$$

Integration of the characteristic equations (1.11.6) yields the solution for the ***flow*** $\mathbf{x}(\epsilon) = \phi_\epsilon \mathbf{x}$ of the vector field defined by (1.11.5), whose initial condition starts from $\mathbf{x} = \mathbf{x}(0)$. Suppose the characteristic equations for two such flows parameterised by $(\epsilon, \sigma) \in \mathbb{R}$ are given respectively by,

$$\frac{dx_i}{d\epsilon} = v_i(\mathbf{x}(\epsilon)) \quad \text{and} \quad \frac{dx_i}{d\sigma} = w_i(\mathbf{x}(\sigma))\,.$$

The difference of their cross derivatives evaluated at the identity yields the Jacobi-Lie bracket,

$$\begin{aligned}
\frac{d}{d\epsilon}\bigg|_{\epsilon=0}\frac{dx_i}{d\sigma}\bigg|_{\sigma=0} - \frac{d}{d\sigma}\bigg|_{\sigma=0}\frac{dx_i}{d\epsilon}\bigg|_{\epsilon=0} &= \frac{d}{d\epsilon}w_i(\mathbf{x}(\epsilon))\bigg|_{\epsilon=0} - \frac{d}{d\sigma}v_i(\mathbf{x}(\sigma))\bigg|_{\sigma=0} \\
&= \frac{\partial w_i}{\partial x_j}\frac{dx_i}{d\epsilon}\bigg|_{\epsilon=0} - \frac{\partial v_i}{\partial x_j}\frac{dx_i}{d\sigma}\bigg|_{\sigma=0} \\
&= w_{i,j}v_j - v_{i,j}w_j \\
&= [v, w]_i\,.
\end{aligned}$$

Thus, the Jacobi-Lie bracket of vector fields is the difference between the cross-derivatives with respect to their corresponding characteristic equations, evaluated at the identity. Of course, this difference of cross derivatives would vanish if each derivative were not evaluated ***before*** taking the next one.

The composition of Jacobi-Lie brackets for three divergenceless vector fields $u, v, w \in \mathfrak{X}$ has components given by

$$\begin{aligned}\big[u,[v,w]\big]_i &= u_k v_j w_{i,kj} + u_k v_{j,k} w_{i,j} - u_k w_{j,k} v_{i,j} \\ &\qquad - u_k w_j v_{i,jk} - v_j w_{k,j} u_{i,k} + w_j v_{k,j} u_{i,k} \,. \end{aligned} \tag{1.11.7}$$

Equivalently, in vector form,

$$\begin{aligned}\big[u,[v,w]\big] &= \mathbf{uv} : \nabla\nabla\mathbf{w} + \mathbf{u}\cdot\nabla\mathbf{v}^T\cdot\nabla\mathbf{w}^T - \mathbf{u}\cdot\nabla\mathbf{w}^T\cdot\nabla\mathbf{v}^T \\ &\qquad - \mathbf{uw} : \nabla\nabla\mathbf{v} - \mathbf{v}\cdot\nabla\mathbf{w}^T\cdot\nabla\mathbf{u}^T + \mathbf{w}\cdot\nabla\mathbf{v}^T\cdot\nabla\mathbf{u}^T \,.\end{aligned}$$

Theorem 1.11.3 *The Jacobi-Lie bracket of divergenceless vector fields satisfies the **Jacobi identity**,*

$$[u,[v,w]\,] + [v,[w,u]\,] + [w,[u,v]\,] = 0\,. \tag{1.11.8}$$

Proof. Direct verification using (1.11.7) and summing over cyclic permutations. ∎

> **Exercise.** Prove Theorem 1.11.3 in streamlined notation obtained by writing
>
> $$[v,\, w] = v(w) - w(v)\,,$$
>
> and using bilinearity of the Jacobi-Lie bracket. ★

Lemma 1.11.4 *The $\mathbb{R}^3$-bracket (1.7.12) may be identified with the divergenceless vector fields in (1.11.1) by*

$$[X_G, X_H] = -X_{\{G,H\}}\,, \tag{1.11.9}$$

where $[X_G, X_H]$ is the Jacobi-Lie bracket of vector fields X_G and X_H.

Proof. Equation (1.11.9) may be verified by a direct calculation,

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

∎

Remark 1.11.5 The last step in the proof of Lemma 1.11.4 uses the Jacobi identity for the $\mathbb{R}^3$-bracket, which follows from the Jacobi identity for divergenceless vector fields, since

$$X_F X_G X_H = -\{F,\,\{G,\,\{H,\,\cdot\,\}\}\}$$

1.11.2 Geometric forms of Poisson brackets

Determinant & wedge-product forms of the canonical bracket

For one degree of freedom, the canonical Poisson bracket $\{F, H\}$ is the same as the determinant for a change of variables $(q, p) \to (F(q, p), H(q, p))$,

$$\{F, H\} = \frac{\partial F}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial H}{\partial q}\frac{\partial F}{\partial p} = \det\frac{\partial(F, H)}{\partial(q, p)}\,. \tag{1.11.10}$$

This may be written in terms of the differentials of the functions $(F(q, p), H(q, p))$ defined as

$$dF = \frac{\partial F}{\partial q}\,dq + \frac{\partial F}{\partial p}\,dp \quad \text{and} \quad dH = \frac{\partial H}{\partial q}\,dq + \frac{\partial H}{\partial p}\,dp\,, \tag{1.11.11}$$

by writing the canonical Poisson bracket $\{F, H\}$ as a phase space density

$$dF \wedge dH = \det\frac{\partial(F, H)}{\partial(q, p)} dq \wedge dp = \{F, H\} dq \wedge dp\,. \tag{1.11.12}$$

Here the wedge product $\wedge$ in $dF \wedge dH = -\,dH \wedge dF$ is introduced to impose the antisymmetry of the Jacobian determinant under interchange of its columns.

Definition 1.11.6 (Wedge product of differentials)
The wedge product of differentials (dF, dG, dH) of any smooth functions (F, G, H) is defined by its following three properties.

(i) $\wedge$ is anticommutative: $dF \wedge dG = -dG \wedge dF$;

(ii) $\wedge$ *is bilinear:* $(adF + bdG) \wedge dH = a(dF \wedge dH) + b(dG \wedge dH)$*;*

(iii) $\wedge$ *is associative:* $dF \wedge (dG \wedge dH) = (dF \wedge dG) \wedge dH$*.*

Remark 1.11.7 These are the usual properties of area elements and volume elements in integral calculus. These properties also apply in computing changes of variables.

> **Exercise.** Verify formula (1.11.12) from equation (1.11.11) and the linearity and antisymmetry of the wedge product, so that, e.g., $dq \wedge dp = -dp \wedge dq$ and $dq \wedge dq = 0$.
>
> ★

Determinant & wedge-product forms of the $\mathbb{R}^3$-bracket

The $\mathbb{R}^3$-bracket in equation (1.7.12) may also be rewritten equivalently as a Jacobian determinant, namely,

$$\{F,\, H\} = -\,\nabla S^2 \cdot \nabla F \times \nabla H = -\,\frac{\partial(S^2, F, H)}{\partial(X_1, X_2, X_3)}\,, \qquad (1.11.13)$$

where

$$\frac{\partial(F_1, F_2, F_3)}{\partial(X_1, X_2, X_3)} = \det\left(\frac{\partial \mathbf{F}}{\partial \mathbf{X}}\right)\,. \qquad (1.11.14)$$

The determinant in three dimensions may be defined using the antisymmetric tensor symbol ϵ_{ijk} as

$$\epsilon_{ijk} \det\left(\frac{\partial \mathbf{F}}{\partial \mathbf{X}}\right) = \epsilon_{abc}\, \frac{\partial F_a}{\partial X_i}\frac{\partial F_b}{\partial X_j}\frac{\partial F_c}{\partial X_k}\,, \qquad (1.11.15)$$

where, as mentioned earlier, we sum on repeated indices over their range. We shall keep track of the antisymmetry of the determinant in three dimensions by using the ***wedge product*** ($\wedge$)

$$\det\left(\frac{\partial \mathbf{F}}{\partial \mathbf{X}}\right) dX_1 \wedge dX_2 \wedge dX_3 = dF_1 \wedge dF_2 \wedge dF_3\,. \qquad (1.11.16)$$

Thus, the $\mathbb{R}^3$-bracket in equation (1.7.12) may be rewritten equivalently in wedge-product form as

$$
\begin{aligned}
\{F,\, H\}\, dX_1 \wedge dX_2 \wedge dX_3 &= -(\nabla S^2 \cdot \nabla F \times \nabla H)\, dX_1 \wedge dX_2 \wedge dX_3 \\
&= -\, dS^2 \wedge dF \wedge dH\,.
\end{aligned}
$$

Poisson brackets of this type are called ***Nambu brackets***, since [Na1973] introduced them in three dimensions. They can be generalised to any dimension, but this requires additional compatibility conditions [Ta1994].

1.11.3 Nambu brackets

Theorem 1.11.8 (Nambu brackets [Na1973])
*For any smooth functions $F\, H \in \mathcal{F}(\mathbb{R}^3)$ of coordinates $\mathbf{X} \in \mathbb{R}^3$ with volume element $d^{\,3}X$, the **Nambu bracket***

$$\{F,\, H\} : \mathcal{F}(\mathbb{R}^3) \times \mathcal{F}(\mathbb{R}^3) \to \mathcal{F}(\mathbb{R}^3)$$

defined by

$$
\begin{aligned}
\{F,\, H\}\, d^{\,3}X &= -\nabla C \cdot \nabla F \times \nabla H d^{\,3}X \\
&= -\, dC \wedge dF \wedge dH\,,
\end{aligned}
\tag{1.11.17}
$$

possesses the properties (1.3.4) required of a Poisson bracket for any choice of distinguished smooth function $C : \mathbb{R}^3 \to \mathbb{R}$.

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

$$X_H = \{\,\cdot\,,\, h\} = \nabla C \times \nabla H \cdot \nabla\,,$$

in which

$$\operatorname{div}(\nabla C \times \nabla H) = 0\,.$$

Divergenceless vector fields are derivative operators that satisfy the Leibnitz product rule and the Jacobi identity. These properties hold in this case for any choice of smooth functions $C,\, H \in \mathcal{F}(\mathbb{R}^3)$. The other two properties – bilinearity and skew symmetry – hold as properties of the wedge product. Hence, the Nambu $\mathbb{R}^3$ bracket in (1.11.17) satisfies all the properties required of a Poisson bracket specified in Definition 1.3.4. ∎

1.12 Geometry of solution behaviour

1.12.1 Restricting axisymmetric ray optics to level sets

Having realised that the $\mathbb{R}^3$-bracket in equation (1.7.12) is associated to Jacobian determinants for changes of variables, it is natural to transform the dynamics of the axisymmetric optical variables (1.4.5) from three dimensions $(X_1, X_2, X_3) \in \mathbb{R}^3$ to one of its level sets $S^2 > 0$. For convenience, we first make a linear change of Cartesian coordinates in $\mathbb{R}^3$ that explicitly displays the axisymmetry of the level sets of S^2 under rotations, namely,

$$S^2 = X_1X_2 - X_3^2 = Y_1^2 - Y_2^2 - Y_3^2\,, \tag{1.12.1}$$

with

$$Y_1 = \frac{1}{2}(X_1 + X_2)\,, \quad Y_2 = \frac{1}{2}(X_2 - X_1)\,, \quad Y_3 = X_3\,.$$

In these new Cartesian coordinates $(Y_1, Y_2, Y_3) \in \mathbb{R}^3$, the level sets of S^2 are manifestly invariant under rotations about the Y_1-axis.

Exercise. Show that this linear change of Cartesian coordinates preserves the orientation of volume elements, but scales them by a constant factor of one-half. That is, show

$$\{F,\, H\}\, dY_1 \wedge dY_2 \wedge dY_3 = \frac{1}{2}\,\{F,\, H\}\, dX_1 \wedge dX_2 \wedge dX_3\,.$$

The overall constant factor of one-half here is unimportant, because it may be simply absorbed into the units of axial distance in the dynamics induced by the $\mathbb{R}^3$-bracket for axisymmetric ray optics in the Y-variables. ★

Each of the family of hyperboloids of revolution in (1.12.1) comprises a layer in the "hyperbolic onion" preserved by axisymmetric ray optics. We use hyperbolic polar coordinates on these layers in analogy to spherical coordinates,

$$Y_1 = S\cosh u\,, \quad Y_2 = S\sinh u\cos\psi\,, \quad Y_3 = S\sinh u\sin\psi\,. \tag{1.12.2}$$

The $\mathbb{R}^3$-bracket (1.7.12) thereby transforms into hyperbolic coordinates (1.12.2) as

$$\{F,\,H\}\,dY_1\wedge dY_2\wedge dY_3 = -\,\{F,\,H\}_{hyperb}\,S^2\,dS\wedge d\psi\wedge d\cosh u\,. \tag{1.12.3}$$

Note that the oriented quantity

$$S^2 d\cosh u\wedge d\psi = -S^2 d\psi\wedge d\cosh u\,,$$

is the ***area element on the hyperboloid*** corresponding to the constant S^2.

On a constant level surface of S^2 the function $\{F,H\}_{hyperb}$ only depends on $(\cosh u,\psi)$ so the Poisson bracket for optical motion on any *particular* hyperboloid is then

$$\begin{aligned}\{F,\,H\}\,d^{\,3}Y &= -\,S^2 dS\wedge dF\wedge dH \qquad (1.12.4)\\ &= -\,S^2 dS\wedge\{F,\,H\}_{hyperb}\,d\psi\wedge d\cosh u\,,\end{aligned}$$

with

$$\{F,\,H\}_{hyperb} = \left(\frac{\partial F}{\partial\psi}\frac{\partial H}{\partial\cosh u} - \frac{\partial H}{\partial\cosh u}\frac{\partial F}{\partial\psi}\right)\,.$$

Being a constant of the motion, the value of S^2 may be absorbed by a choice of units for any given initial condition and the Poisson bracket for the optical motion thereby becomes ***canonical on each hyperboloid,***

$$\frac{d\psi}{dz} = \{\psi,\,H\}_{hyperb} = \frac{\partial H}{\partial\cosh u}\,, \tag{1.12.5}$$

$$\frac{d\cosh u}{dz} = \{\cosh u,\,H\}_{hyperb} = -\frac{\partial H}{\partial\psi}\,. \tag{1.12.6}$$

In the Cartesian variables $(Y_1,Y_2,Y_3)\in\mathbb{R}^3$, one has $\cosh u = Y_1/S$ and $\psi = \tan^{-1}(Y_3/Y_2)$. In the original variables,

$$\cosh u = \frac{X_1+X_2}{2S} \quad\text{and}\quad \psi = \tan^{-1}\frac{2X_3}{X_2-X_1}\,.$$

Example 1.12.1 *For a paraxial harmonic guide, whose Hamiltonian is,*

$$H = \frac{1}{2}(|\mathbf{p}|^2 + |\mathbf{q}|^2) = \frac{1}{2}(X_1 + X_2) = Y_1 \,, \tag{1.12.7}$$

the ray paths consist of circles cut by the intersections of level sets of the planes $Y_1 = const$ with the hyperboloids of revolution about the Y_1-axis, given by $S^2 = const$.

The dynamics for $\mathbf{Y} \in \mathbb{R}^3$ is given by

$$\dot{\mathbf{Y}} = \big\{\mathbf{Y},\, H\big\} = \nabla_{\mathbf{Y}} S^2 \times \widehat{\mathbf{Y}}_{\mathbf{1}} = 2\widehat{\mathbf{Y}}_{\mathbf{1}} \times \mathbf{Y} \,, \tag{1.12.8}$$

on using the (1.12.1) to transform the $\mathbb{R}^3$ bracket in (1.7.12). Thus, for the paraxial harmonic guide, the rays spiral down the optical axis following circular helices whose radius is determined by their initial conditions.

Exercise. Verify that equation (1.12.3) transforms the $\mathbb{R}^3$-bracket from Cartesian to hyperboloidal coordinates, by using the definitions in equations (1.12.2). ★

Exercise. Reduce $\{F, H\}_{hyperb}$ to the conical level set $S = 0$. ★

Exercise. Reduce the $\mathbb{R}^3$ dynamics of (1.7.12) to level sets of the Hamiltonian

$$H = aX_1 + bX_2 + cX_3 \,,$$

for constants (a, b, c). Explain how this reduction simplifies the equations of motion. ★

1.12.2 Geometric phase on level sets of $S^2 = p_\phi^2$

In polar coordinates, the axisymmetric invariants are

$$\begin{aligned}
Y_1 &= \frac{1}{2}\Big(p_r^2 + p_\phi^2/r^2 + r^2\Big)\,,\\
Y_2 &= \frac{1}{2}\Big(p_r^2 + p_\phi^2/r^2 - r^2\Big)\,,\\
Y_3 &= rp_r\,.
\end{aligned}$$

Hence, the corresponding volume elements are found to be

$$\begin{aligned}
d^{\,3}Y &=: dY_1\wedge dY_2\wedge dY_3\\
&= d\frac{S^3}{3}\wedge d\cosh u\wedge d\psi\\
&= dp_\phi^2\wedge dp_r\wedge dr\,.
\end{aligned}\tag{1.12.9}$$

On a level set of $S = p_\phi$ this implies

$$S\,d\cosh u\wedge d\psi = 2\,dp_r\wedge dr\,,\tag{1.12.10}$$

so the transformation of variables $(\cosh u,\,\psi)\to(p_r,\,r)$ is ***canonical*** on level sets of $S = p_\phi$.

One recalls Stokes Theorem on phase space

$$\iint_A dp_j\wedge dq_j = \oint_{\partial A} p_j dq_j\,,\tag{1.12.11}$$

where the boundary of the phase space area ∂A is taken around a loop on a closed orbit. Either in polar coordinates or on an invariant hyperboloid $S = p_\phi$ this loop integral becomes

$$\begin{aligned}
\oint \mathbf{p}\cdot d\mathbf{q} := \oint p_j dq_j &= \oint\Big(p_\phi d\phi + p_r dr\Big)\\
&= \oint\Big(\frac{S^3}{3}d\phi + \cosh u\,d\psi\Big)\,.
\end{aligned}$$

Thus we may compute the total phase change around a closed periodic orbit on the level set of hyperboloid S from

$$\begin{aligned}\oint \frac{S^3}{3} d\phi &= \frac{S^3}{3}\Delta\phi \\ &= \underbrace{-\oint \cosh u \, d\psi}_{\text{Geometric } \Delta\phi} + \underbrace{\oint \mathbf{p}\cdot d\mathbf{q}}_{\text{Dynamic } \Delta\phi} \, . \end{aligned} \tag{1.12.12}$$

Evidently, one may denote the total change in phase as the sum

$$\Delta\phi = \Delta\phi_{geom} + \Delta\phi_{dyn} \, ,$$

by identifying the corresponding terms in the previous formula. By the Stokes theorem (1.12.11), one sees that the geometric phase associated with a periodic motion on a particular hyperboloid is given by the hyperbolic solid angle enclosed by the orbit, times a constant factor depending on the level set value $S = p_\phi$. Thus, the name: ***geometric phase***.

1.13 Geometric ray optics in anisotropic media

> Every ray of light has therefore two opposite sides.... And since the crystal by this disposition or virtue does not act upon the rays except when one of their sides of unusual refraction looks toward that coast, this argues a virtue or disposition in those sides of the rays which answers to and sympathises with that virtue or disposition of the crystal, as the poles of two magnets answer to one another....
>
> – Newton, *Optiks* 1704

Some media have directional properties that are exhibited by differences in the transmission of light in different directions. This effect is seen, for example, in certain crystals. Fermat's principle for such media still conceives light rays as lines in space (i.e., no polarisation vectors, yet), but the refractive index along the paths of the rays in the medium is allowed to depend on both position and *direction*. In this case, Theorem 1.1.15 adapts easily to yield the expected 3D eikonal equation (1.1.7).

1.13.1 Fermat and Huygens principles for anisotropic media

Let us investigate how Fermat's principle (1.1.8) changes, when Huygens equation (3.2.20) allows the velocity of the wave fronts to depend on direction because of material anisotropy.

We consider a medium in which the tangent vector of a light ray does not point along the normal to its wave-front, $\nabla S(\mathbf{r})$, but instead satisfies a matrix relation depending on the spatial location along the ray path, $\mathbf{r}(s)$, as

$$\frac{d\mathbf{r}}{ds} = \mathcal{D}^{-1}(\mathbf{r})\nabla S\,, \tag{1.13.1}$$

with with $\dot{\mathbf{r}}(s) := d\mathbf{r}/ds$ and $|\dot{\mathbf{r}}| = 1$ for an invertible matrix function $\mathcal{D}$ that characterises the medium. In components, this ***anisotropic Huygens equation*** is written

$$\frac{\partial S}{\partial r^i} = \mathcal{D}_{ij}(\mathbf{r})\frac{dr^j}{ds}\,. \tag{1.13.2}$$

We shall write the Euler-Lagrange equation for $\mathbf{r}(s)$ that arises from Fermat's principle for this anisotropic version of Huygens equation and derive its corresponding Snell's Law.

We begin by taking the square on both sides of the vector equation (1.13.2), which produces the anisotropic version of the scalar eikonal equation (1.1.19), now written in the form

$$|\nabla S|^2 = \frac{d\mathbf{r}}{ds}^T(\mathcal{D}^T\mathcal{D})\frac{d\mathbf{r}}{ds} = \frac{dr^i}{ds}(\mathcal{D}_{ik}\mathcal{D}_{kj})\frac{dr^j}{ds} = n^2(\mathbf{r},\dot{\mathbf{r}}). \tag{1.13.3}$$

Substituting this expression into Fermat's principle yields

$$\begin{aligned} 0 = \delta \mathsf{A} &= \delta\int_A^B n(\mathbf{r}(s),\dot{\mathbf{r}}(s))\,ds \qquad (1.13.4)\\ &= \delta\int_A^B \sqrt{\dot{\mathbf{r}}(s)\cdot(\mathcal{D}^T\mathcal{D})(\mathbf{r})\cdot\dot{\mathbf{r}}(s)}\,ds. \end{aligned}$$

This is the variational principle for the curve that leaves the distance between two points A and B stationary under variations in a space with ***Riemannian metric*** given by $\mathcal{G} := \mathcal{D}^T\mathcal{D}$, whose element of length is defined by

$$n^2(\mathbf{r}, \dot{\mathbf{r}})ds^2 = d\mathbf{r} \cdot \mathcal{G}(\mathbf{r}) \cdot d\mathbf{r}. \tag{1.13.5}$$

When $\mathcal{G}_{ij} = n^2(\mathbf{r})\,\delta_{ij}$, one recovers the isotropic case discussed in §1.1.1.

Remark 1.13.1 *By construction, the quantity under the integral in (1.13.4), now rewritten as*

$$\mathsf{A} = \int_A^B \sqrt{\dot{\mathbf{r}} \cdot \mathcal{G}(\mathbf{r}) \cdot \dot{\mathbf{r}}}\; ds = \int_A^B n(\mathbf{r}(s), \dot{\mathbf{r}}(s))\, ds \tag{1.13.6}$$

is homogeneous of degree 1 and thus is invariant under reparameterising the ray path $\mathbf{r}(s) \to \mathbf{r}(\sigma)$.

Continuing the variational computation leads to

$$\begin{aligned} 0 = \delta\mathsf{A} &= \delta \int_A^B \sqrt{\dot{\mathbf{r}}(s) \cdot \mathcal{G}(\mathbf{r}) \cdot \dot{\mathbf{r}}(s)}\, ds \\ &= \int_A^B \left[\frac{\partial n(\mathbf{r}, \dot{\mathbf{r}})}{\partial \mathbf{r}} - \frac{d}{ds}\left(\frac{\mathcal{G}(\mathbf{r}) \cdot \dot{\mathbf{r}}}{n(\mathbf{r}, \dot{\mathbf{r}})} \right) \right] \cdot \delta\mathbf{r}\, ds \end{aligned}$$

upon substituting $\sqrt{\dot{\mathbf{r}}(s) \cdot \mathcal{G}(\mathbf{r}) \cdot \dot{\mathbf{r}}(s)} = n(\mathbf{r}, \dot{\mathbf{r}})$.

Stationarity, $\delta\mathsf{A} = 0$ now yields the Euler-Lagrange equation

$$\dot{\mathbf{p}} = \frac{\partial n(\mathbf{r}, \dot{\mathbf{r}})}{\partial \mathbf{r}}\,, \quad \text{with} \quad \mathbf{p} := \frac{\mathcal{G}(\mathbf{r}) \cdot \dot{\mathbf{r}}}{n(\mathbf{r}, \dot{\mathbf{r}})} = \frac{\partial n(\mathbf{r}, \dot{\mathbf{r}})}{\partial \dot{\mathbf{r}}}\,, \tag{1.13.7}$$

where the vector $\mathbf{p}$ is the ***optical momentum***. This is the ***eikonal equation for an anisotropic medium*** whose Huygens equation is (1.13.1).

Note the elegant formula

$$\mathbf{p} \cdot d\mathbf{r} = n(\mathbf{r}, \dot{\mathbf{r}})\, ds\,, \quad \text{or equivalently} \quad \mathbf{p} \cdot \dot{\mathbf{r}} = n(\mathbf{r}, \dot{\mathbf{r}})\,,$$

which follows, as in the isotropic case, from Euler's relation (1.1.14) for homogeneous functions of degree 1, and in particular for the Lagrangian in (1.13.6).

Exercise. Transform the arc-length variable in Fermat's principle (1.13.4) to $d\sigma = n(\mathbf{r}, \dot{\mathbf{r}})ds$ so that

$$d\sigma^2 = d\mathbf{r} \cdot \mathcal{G}(\mathbf{r}) \cdot d\mathbf{r} \tag{1.13.8}$$

and recalculate its Euler-Lagrange equation. ★

Answer. Denoting $\mathbf{r}'(\sigma) = d\mathbf{r}/d\sigma$ in Fermat's principle (1.13.4) and recalculating the stationary condition yields

$$0 = \delta \mathsf{A} \quad = \quad \int_A^B \left[\frac{d\sigma}{2\, nds} r'^k \frac{\partial \mathcal{G}_{kl}}{\partial \mathbf{r}} r'^l - \frac{d}{d\sigma} \left(\frac{d\sigma}{nds} \mathcal{G} \cdot \mathbf{r}' \right) \right] \cdot \delta \mathbf{r} \, d\sigma .$$

Applying $n(\mathbf{r}, \dot{\mathbf{r}})ds = d\sigma$ leads to the Euler-Lagrange equation

$$\frac{1}{2} r'^k \frac{\partial \mathcal{G}_{kl}}{\partial \mathbf{r}} r'^l - \frac{d}{d\sigma} \big(\mathcal{G}(\mathbf{r}) \cdot \mathbf{r}'(\sigma) \big) = 0 \, . \tag{1.13.9}$$

A curve $\mathbf{r}(\sigma)$ satisfying this equation leaves the length between points A and B stationary under variations and is called a ***geodesic*** with respect to the metric $\mathcal{G}(\mathbf{r})$, whose length element is defined in equation (1.13.8).

▲

Exercise. Does the variational principle

$$0 = \delta \mathsf{A} = \delta \int_A^B \tfrac{1}{2} \mathbf{r}'(\sigma) \cdot \mathcal{G}(\mathbf{r}) \cdot \mathbf{r}'(\sigma) \, d\sigma$$

also imply the Euler-Lagrange equation (1.13.9)? Prove it.

Hint: take a look at equation (1.1.12) and the remark after it about Finsler geometry and singular Lagrangians. Does this calculation reveal a general principle about the stationary conditions of Fermat's principle for singular Lagrangians of degree 1 and their associated induced Lagrangians of degree 2?

★

Exercise. Show that the Euler-Lagrange equation (1.13.9) may be written as

$$\mathbf{r}''(\sigma) = -\mathbf{G}(\mathbf{r}, \mathbf{r}') \,, \quad \text{with} \quad G^i := \Gamma^i_{jk}(\mathbf{r}) r'^j r'^k \tag{1.13.10}$$

and

$$\Gamma^i_{jk}(\mathbf{r}) = \frac{1}{2}\mathcal{G}^{il}\left[\frac{\partial \mathcal{G}_{lk}(\mathbf{r})}{\partial r^j} + \frac{\partial \mathcal{G}_{lj}(\mathbf{r})}{\partial r^k} - \frac{\partial \mathcal{G}_{jk}(\mathbf{r})}{\partial r^l}\right], \tag{1.13.11}$$

where $\mathcal{G}_{ki}\mathcal{G}^{il} = \delta^l_k$. Equation (1.13.11) identifies $\Gamma^i_{jk}(\mathbf{r})$ as the ***Christoffel coefficients*** of the Riemannian metric $\mathcal{G}_{ij}(\mathbf{r})$ in equation (1.13.8). Note that the Christoffel coefficients are symmetric under exchange of their lower indices, $\Gamma^i_{jk}(\mathbf{r}) = \Gamma^i_{kj}(\mathbf{r})$.

The vector $\mathbf{G}$ in equation (1.13.10) is called the ***geodesic spray*** of the Riemannian metric $\mathcal{G}_{ij}(\mathbf{r})$. Its analytical properties (e.g., smoothness) govern the behaviour of the solutions $\mathbf{r}(\sigma)$ for the geodesic paths. ★

1.13.2 Ibn Sahl-Snell law for anisotropic media

The statement of the Ibn Sahl-Snell law relation at discontinuities of the refractive index for isotropic media is a bit more involved for an anisotropic medium with Huygens equation (1.13.1).

A break in the direction $\widehat{\mathbf{s}}$ of the ray vector is still expected for anisotropic media at any interface of finite discontinuity in the material properties (refractive index and metric) that may be encountered along the ray path $\mathbf{r}(s)$. According to the eikonal equation for anisotropic media (1.13.7) the jump in 3D optical momentum across a material interface, $\Delta\mathbf{p} := \mathbf{p} - \bar{\mathbf{p}}$, must satisfy the relation

$$\Delta\mathbf{p} \times \frac{\partial n}{\partial \mathbf{r}} = \Delta\left(\frac{\mathcal{G}(\mathbf{r})\cdot\dot{\mathbf{r}}}{n(\mathbf{r}, \dot{\mathbf{r}})}\right) \times \frac{\partial n}{\partial \mathbf{r}} = 0\,. \tag{1.13.12}$$

This means the jump in optical momentum $\Delta\mathbf{p}$ at a material interface can only occur in the direction normal to the interface. Hence, the transverse projection of the 3D optical momenta onto the tangent plane will be invariant across the interface.

Let ψ and ψ' denote the angles of incident and transmitted ***momentum directions***, measured from the normal $\hat{\mathbf{z}}$ through the interface. Preservation of the transverse components of the momentum vector

$$\big(\mathbf{p} - \bar{\mathbf{p}}\big) \times \hat{\mathbf{z}} = 0\,, \tag{1.13.13}$$

means that these momentum vectors must lie in the same plane and the angles $\psi = \cos^{-1}(\mathbf{p}\cdot\hat{\mathbf{z}}/|\mathbf{p}|)$ and $\bar{\psi} = \cos^{-1}(\bar{\mathbf{p}}\cdot\hat{\mathbf{z}}/|\bar{\mathbf{p}}|)$ must satisfy

$$|\mathbf{p}|\sin\psi = |\bar{\mathbf{p}}|\sin\bar{\psi}\,. \tag{1.13.14}$$

Relation (1.13.14) determines the angles of incidence and transmission of the *optical momentum directions*. However, the transverse components of the optical momentum alone are not enough to determine the ray directions, in general, because the inversion using equation (1.13.7) involves all three of the momentum and velocity components.

1.14 Ten geometrical features of ray optics

1. The design of axisymmetric planar optical systems reduces to multiplication of ***symplectic matrices*** corresponding to each element of the system, as in Theorem 1.6.4.

2. Hamiltonian evolution occurs by ***canonical transformations***. Such transformations may be obtained by integrating the characteristic equations of Hamiltonian vector fields, which are defined by Poisson-bracket operations with smooth functions on phase space, as in the proof of Theorem 1.5.3.

3. The Poisson bracket is associated geometrically with the Jacobian for canonical transformations in Section 1.11.2. Canonical transformations are generated by Poisson-bracket operations and these transformations preserve the Jacobian.

4. A one-parameter symmetry, that is, an invariance under canonical transformations that are generated by a Hamiltonian vector field $X_{p_\phi} = \{\,\cdot\,, p_\phi\,\}$, separates out an angle, ϕ, whose canonically conjugate momentum p_ϕ is conserved. As discussed in Section 1.3.2, the conserved quantity p_ϕ may be an important bifurcation parameter for the remaining reduced system. The dynamics of the angle ϕ decouples from the reduced system and can be determined as a quadrature after solving the reduced system.

5. Given a symmetry of the Hamiltonian, it may be wise to transform from phase space coordinates to invariant variables as in (1.4.5). This transformation defines the quotient map, which quotients out the angle(s) conjugate to the symmetry generator. The image of the quotient map produces the orbit manifold, a reduced manifold whose points are orbits under the symmetry transformation. The corresponding transformation of the Poisson bracket is done using the chain rule as in (1.7.6). Closure of the Poisson brackets of the invariant variables amongst themselves is a necessary condition for the quotient map to be a momentum map, as discussed in Section 1.9.2.

6. Closure of the Poisson brackets among an odd number of invariant variables is no cause for regret. It only means that this Poisson bracket among the invariant variables is not canonical (symplectic). For example, the Nambu $\mathbb{R}^3$ bracket (1.11.17) arises this way.

7. The bracket resulting from transforming to invariant variables could also be Lie-Poisson. This will happen when the new invariant variables are quadratic in the phase space variables, as occurs for the Poisson brackets among the axisymmetric variables X_1, X_2 and X_3 in (1.4.5). Then the Poisson brackets among them are *linear* in the new variables with constant coefficients. Those constant coefficients are dual to the structure constants of a Lie algebra. In that case, the brackets will take the Lie-Poisson form (1.10.1) and the transformation to invariant variables will be the momentum map associated as

in Remark 1.9.4 with the action of the symmetry group on the phase space.

8. The orbits of the solutions in the space of axisymmetric invariant variables in ray optics lie on the intersections of level sets of the Hamiltonian and the Casimir for the noncanonical bracket. The Petzval invariant S^2 in (1.7.13) is the Casimir for the Nambu bracket in $\mathbf{R}^3$, which for axisymmetric, translation-invariant ray optics is also a Lie-Poisson bracket. In this case, the ray paths are revealed when the Hamiltonian knife slices through the level sets of the Petzval invariant. These level sets are the layers of the *hyperbolic onion* shown in Figure 1.8. When restricted to a level set of the Petzval invariant, the dynamics becomes symplectic.

9. The *phases* associated with reconstructing the solution from the reduced space of invariant variables by going back to the original space of canonical coordinates and momenta naturally divide into their geometric and dynamic parts as in equation (1.12.12). In ray optics as governed by Fermat's principle, the geometric phase is related to the area enclosed by a periodic solution on a symplectic level set of the Petzval invariant S^2. This is no surprise, because the Poisson bracket on the level set is determined from the Jacobian using that area element.

10. The Lagrangian in Fermat's principle is typically homogeneous of degree 1 and thus is singular; that is, its Hessian determinant vanishes, as occurs in the example of Fermat's principle in anisotropic media discussed in Section 1.13. In this case, the ray directions cannot be determined from the optical momentum and the coordinate along the ray. In particular, Snell's Law of refraction at an interface determines momentum directions, but not ray directions. Even so, the Euler-Lagrange equations resulting from Fermat's principle may be regularised by using as induced Lagrangian that is homogeneous of degree 2, in the sense that reparameterised solutions for the ray paths may be obtained from the resulting ordinary

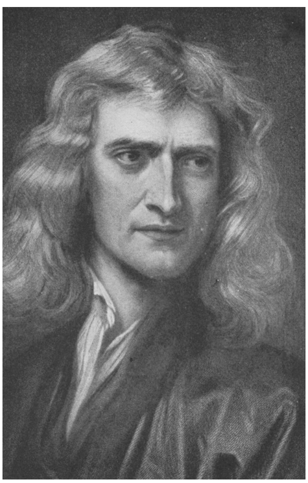
differential equations.

The Euler-Lagrange equation (1.13.9) for geometric optics may be written equivalently in the form (1.13.10), which shows that ray paths follow geodesic motion through a Riemannian space whose metric is determined from optical material parameters whose physical meaning is derived from Huygens equation relating the ray path to the motion of the wave front.

Chapter 2

Newton, Lagrange, Hamilton & the rigid body

2.1 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 2.1.1 (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* $(\mathbf{r}, \dot{\mathbf{r}}, \ddot{\mathbf{r}})$.

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

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

Definition 2.1.2 (Galilean transformations)
Transformations of reference location, origin of time, orientation or state of uniform translation at constant velocity are called ***Galilean transformations****.*

Definition 2.1.3 (Uniform rectilinear motion)
Coordinate systems related by Galilean transformations are said to be in ***uniform rectilinear motion*** *relative to each other.*

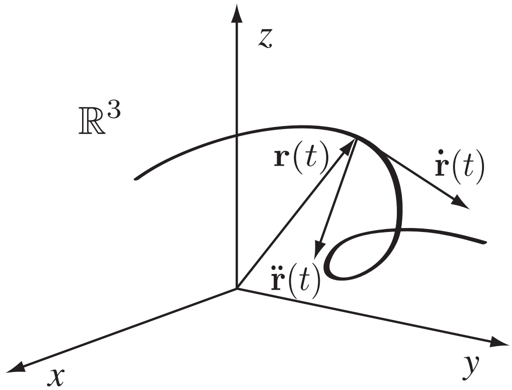


Figure 2.1: Position $\mathbf{r}(t) \in \mathbb{R}^3 \times \mathbb{R}$, velocity $\dot{\mathbf{r}}(t) \in T\mathbb{R}^3 \times \mathbb{R}$ and acceleration $\ddot{\mathbf{r}}(t) \in TT\mathbb{R}^3 \times \mathbb{R}$ along a trajectory of motion governed by Newton's Second Law, $m\ddot{\mathbf{r}} = \mathbf{F}(\mathbf{r}, \dot{\mathbf{r}})$.

Proposition 2.1.4 (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 2.1.5 (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.*

2.1.1 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 2.1.6 *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 2.1.7 (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*

$$\mathbf{F}_j = \sum_{k \neq j} \mathbf{F}_{jk} \quad \textit{where} \quad \mathbf{F}_{jk} = -\mathbf{F}_{kj}\,. \tag{2.1.1}$$

Remark 2.1.8 *Newton's law of gravitational force applies to closed systems, since*

$$\mathbf{F}_{jk} = \frac{\gamma\, m_j m_k}{|\mathbf{r}_{jk}|^3}\mathbf{r}_{jk}\,, \quad \textit{where} \quad \mathbf{r}_{jk} = \mathbf{r}_j - \mathbf{r}_k\,, \tag{2.1.2}$$

with gravitational constant γ.

Exercise. Prove that Newton's law of motion

$$m_j \ddot{\mathbf{r}}_j = \sum_{k\neq j} \mathbf{F}_{jk}\,, \tag{2.1.3}$$

with gravitational forces $\mathbf{F}_{jk}$ in (2.1.2) is Galilean invariant. ★

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

$$\mathbf{F}_{jk} = -\mathbf{F}_{kj}\,. \tag{2.1.4}$$

Corollary 2.1.9 (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, $\dot{\mathbf{p}}_1 + \dot{\mathbf{p}}_2 = m_1\dot{\mathbf{v}}_1 + m_2\dot{\mathbf{v}}_2 = \mathbf{F}_{12} + \mathbf{F}_{21} = 0\,.$ ■

2.1.2 Dynamical quantities

Definition 2.1.10 (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 2.1.11 (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 2.1.12 (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 2.1.13 (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 2.1.14 (Conservative forces)
The forces $\mathbf{F}_j(\mathbf{r}_1, \dots \mathbf{r}_N)$ for a closed system of N particles are ***conservative****, if*

$$\begin{aligned}\sum_j \Big\langle \mathbf{F}_j \,,\, d\mathbf{r}_j \Big\rangle &= -\, dV(\mathbf{r}_1, \dots, \mathbf{r}_N) \\ &:= \sum_j \Big\langle -\, \frac{\partial V}{\partial \mathbf{r}_j} \,,\, d\mathbf{r}_j \Big\rangle \,, \end{aligned} \tag{2.1.5}$$

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

Remark 2.1.15 *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 2.1.16 (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

$$\frac{dK}{dt} = \sum_j \Big\langle \mathbf{F}_j \,,\, \frac{d\mathbf{r}_j}{dt} \Big\rangle = -\, \frac{dV}{dt} \,, \tag{2.1.6}$$

for conservative forces. ∎

Proposition 2.1.17 (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 2.1.18 (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 2.1.19 (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 2.1.20 (Conserving total angular momentum)
In particular, Proposition 2.1.19 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 $(\mathbf{r}_i - \mathbf{r}_j)$.

2.1.3 Newtonian form of free rigid rotation

Definition 2.1.21 *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

$$\mathbf{e}_a(t) = O(t)\mathbf{e}_a(0)\,, \quad a = 1, 2, 3, \tag{2.1.7}$$

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

$$\begin{aligned} \mathbf{r}(t) &= r_0^A(t)\mathbf{e}_A(0)\,, \quad \text{in the fixed basis,} && (2.1.8)\\ &= r^a\mathbf{e}_a(t)\,, \quad \text{in the rotating basis,} && (2.1.9) \end{aligned}$$

and the components r^a relative to the rotating basis satisfy $r^a = \delta^a_A r_0^A(0)$ 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 is called the ***spatial frame*** and the rotating basis 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 (2.1.7) at any time.

Lemma 2.1.22 *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

$$\dot{\mathbf{r}}(t) = r^a\dot{\mathbf{e}}_a(t) = r^a\dot{O}(t)\mathbf{e}_a(0) =: r^a\dot{O}O^{-1}(t)\mathbf{e}_a(t) =: \widehat{\omega}(t)\mathbf{r}\,, \tag{2.1.10}$$

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 2.1.23 (Skew symmetry)
The matrix $\widehat{\omega}(t) = \dot{O}O^{-1}(t)$ in (2.1.10) is skew symmetric. That is,

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

Definition 2.1.24 (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*

$$(\dot{O}O^{-1})_{ab}(t) = \widehat{\omega}_{ab}(t) = -\,\epsilon_{abc}\,\omega_c(t) . \tag{2.1.11}$$

*Equation (2.1.11) 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

$$\widehat{\omega} = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} , \tag{2.1.12}$$

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

$$\dot{\mathbf{r}}(t) =: \widehat{\omega}(t)\mathbf{r} =: \boldsymbol{\omega}(t) \times \mathbf{r} . \tag{2.1.13}$$

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. ★

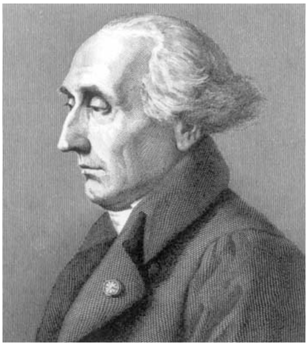
Definition 2.1.25 (Angular momentum of rigid rotation)
The ***angular momentum*** *is defined as the derivative of the kinetic energy with respect to angular velocity. In the present case, this definition produces the linear relation,*

$$\begin{aligned}
\mathbf{J} = \frac{\partial K}{\partial\boldsymbol{\omega}} &= -\sum_{j=1}^{N} m_j\mathbf{r}_j\times\Big(\mathbf{r}_j\times\boldsymbol{\omega}\Big) \\
&= \sum_{j=1}^{N} m_j\Big(|\mathbf{r}_j|^2 Id - \mathbf{r}_j\otimes\mathbf{r}_j\Big)\Big)\boldsymbol{\omega} \\
&=: \mathbb{I}\,\boldsymbol{\omega}\,,
\end{aligned} \tag{2.1.14}$$

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

Exercise. Show that this definition of angular momentum recovers the previous one for a single rotating particle. ★

2.2 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.

2.2.1 Basic definitions for manifolds

Definition 2.2.1 (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 $\phi_\alpha\left(U_\alpha \cap U_\beta\right)$ 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* $\phi_\alpha\left(U_\alpha \cap U_\beta\right)$.

Remark 2.2.2 *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 2.2.3 (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

$$\begin{aligned}
&\phi_N : U_N \to \xi_N + i\eta_N = \frac{x+iy}{1-z} = e^{i\phi}\cot(\theta/2), \\
\textit{and}\qquad &\phi_S : U_S \to \xi_S + i\eta_S = \frac{x-iy}{1+z} = e^{-i\phi}\tan(\theta/2).
\end{aligned}$$

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 2.2.1 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. ★

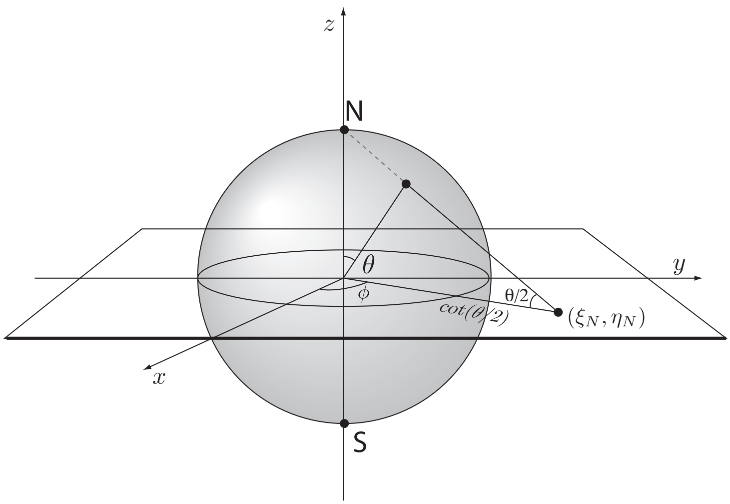


Figure 2.2: In the stereographic projection of the Riemann sphere onto the complex plane from the *North* pole, complex numbers lying outside (resp., inside) the unit circle are projected from points in the upper (resp., lower) hemisphere.

Definition 2.2.4 (Submersion)

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

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

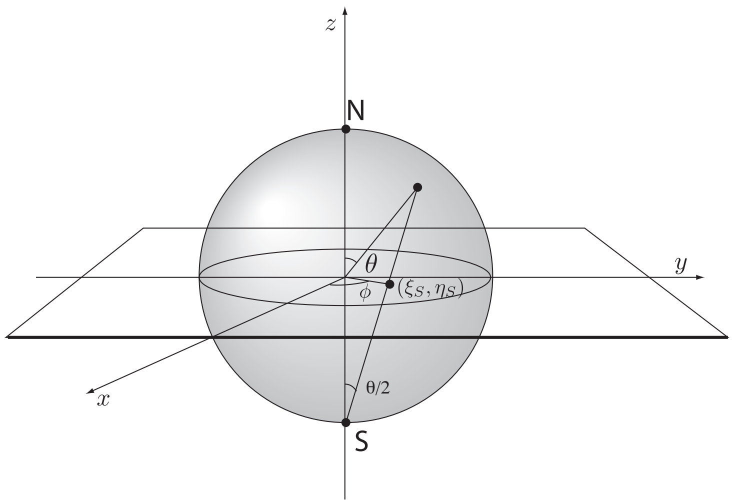


Figure 2.3: In the stereographic projection of the Riemann sphere onto the complex plane from the *South* pole, complex numbers lying outside (resp., inside) the unit circle are also projected from points in the upper (resp., lower) hemisphere.

with $\det(\partial f_i/\partial x^a) \neq 0$, $a = 1, 2, \ldots, n$, *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 2.2.5 *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 2.2.6 (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 2.2.7 (Tangent space to level sets)
Suppose the set

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

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

Example 2.2.9 (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 2.2.10 (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 2.2.11 (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 2.2.12 (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 2.2.13 (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 2.4.

Definition 2.2.14 (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 2.2.15 *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 2.2.16 (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 $t \in]t_1, t_2[$.

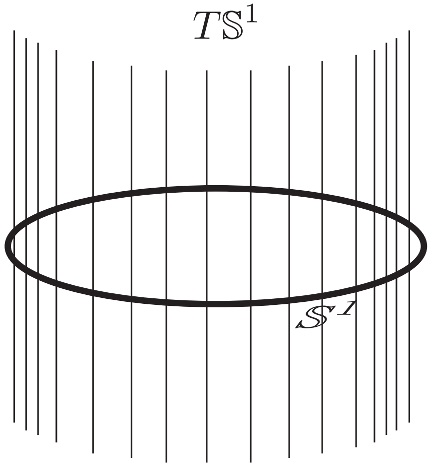


Figure 2.4: The tangent bundle TS^1 of the circle S^1 with coordinates $(\theta, \dot{\theta})$ is the union of the circle with a one-dimensional vector space of line vectors (the angular velocities).

Remark 2.2.17 *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 2.2.18 (Vector basis)
As in (2.2.4) a vector field $\dot{q}$ is defined by the components of its directional derivatives in the chosen coordinate basis, so that, for example,

$$\dot{q} = \dot{q}^a \frac{\partial}{\partial q^a} \quad (\textbf{\textit{vector basis}}). \tag{2.2.1}$$

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

Definition 2.2.19 (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 2.2.20 (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 2.2.21 (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*

$$df(q) = \frac{\partial f}{\partial q^b}\, dq^b \quad \textit{(\textbf{dual basis})}. \tag{2.2.2}$$

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 2.2.22 (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 2.2.23 (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 2.2.24 (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 2.2.25 (Cotangent bundle of a manifold)
The disjoint union of cotangent spaces to M at the points $q \in M$ given by

$$T^*M = \bigcup_{q \in M} T_q^*M \tag{2.2.3}$$

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

Remark 2.2.26 (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*

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

while the components in a new dual basis are

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

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 $T(q)$. ★

2.2.2 Euler-Lagrange equation on a manifold

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

$$\dot{\mathbf{r}}_j(t) = \sum_{a=1}^{K} \frac{\partial \mathbf{r}_j}{\partial q^a}\, \dot{q}^a(t) \quad \text{for} \quad j = 1, 2, \dots N\,. \tag{2.2.4}$$

Remark 2.2.27 *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 2.2.28 (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 2.2.29 (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 2.2.30 (Constrained acceleration formula)
The induced kinetic energy $T(q,\dot{q})$ on the manifold M satisfies

$$\sum_{j=1}^{N}\Big\langle m_j\ddot{\mathbf{r}}_j\,,\,d\mathbf{r}_j(q)\Big\rangle\Big|_{\mathbf{r}(q)} = \sum_{a=1}^{K}\Big\langle \frac{d}{dt}\frac{\partial T}{\partial \dot{q}^a} - \frac{\partial T}{\partial q^a}\,,\,dq^a\Big\rangle\,. \qquad (2.2.5)$$

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

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

■

Lemma 2.2.31 (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. ■

The proof of Theorem 2.2.29 now proceeds by assembling the formulas for constrained acceleration and work rate in the previous two lemmas, as a direct calculation.

Proof.

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

■

Corollary 2.2.32 (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****,*

$$\big[\,L\,\big]_{q^a} := \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^a} - \frac{\partial L}{\partial q^a} = 0\,, \tag{2.2.6}$$

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

$$L(q,\dot{q}) := T(q,\dot{q}) - V\big(\mathbf{r}(q)\big)\,. \tag{2.2.7}$$

Proof. This corollary of Theorem 2.2.29 follows by independence of the differential basis elements dq^a in the final line of its proof. ∎

Theorem 2.2.33 (Hamilton's principle of stationary action)
*The **Euler-Lagrange equation**,*

$$\big[\, L\,\big]_{q^a} := \frac{d}{dt}\frac{\partial L}{\partial \dot q^a} - \frac{\partial L}{\partial q^a} = 0\,, \tag{2.2.8}$$

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

$$S = \int_{t_1}^{t_2} L(q,\dot q)\,dt\,. \tag{2.2.9}$$

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

$$\delta S = 0\,, \tag{2.2.10}$$

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

Remark 2.2.34 (Variational derivatives)
The variational derivative has already appeared when introducing Fermat's principle in equations (1.1.27) - (1.1.32). Its definition will be repeated below for convenience.

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

$$\delta S = \delta \int_{t_1}^{t_2} L(q(t),\dot q(t))dt := \frac{d}{ds}\bigg|_{s=0} \int_{t_1}^{t_2} L(q(t,s),\dot q(t,s))dt\,. \tag{2.2.11}$$

Differentiating under the integral sign, denoting

$$\delta q(t) := \frac{d}{ds}\bigg|_{s=0} q(t,s)\,, \tag{2.2.12}$$

and integrating by parts produces

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

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 (2.2.6) hold, that is, provided $[\,L\,]_q = 0$. ∎

2.2.3 Geodesic motion on Riemannian manifolds

The kinetic energy in Theorem 2.2.29 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 $a, b = 1, 2, \ldots, K$, where K is the dimension of the manifold M.

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

$$L(q,\,\dot{q}) = \frac{1}{2}\dot{q}^b g_{bc}(q)\dot{q}^c =: \frac{1}{2}\|\dot{q}\|^2 \ \geq 0\,. \tag{2.2.13}$$

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

$$\ddot{q}^{\,c} + \Gamma^c_{be}(q)\dot{q}^b\dot{q}^e = 0\,, \tag{2.2.14}$$

with

$$\Gamma^c_{be}(q) = \frac{1}{2}g^{ca}\left[\frac{\partial g_{ae}(q)}{\partial q^b} + \frac{\partial g_{ab}(q)}{\partial q^e} - \frac{\partial g_{be}(q)}{\partial q^a}\right], \tag{2.2.15}$$

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 (1.1.11) for Fermat's principle; namely

$$L(\mathbf{q},\,\dot{\mathbf{q}}) = \frac{1}{2}n^2(\mathbf{q})\,\dot{q}^b\delta_{bc}\dot{q}^c\,, \tag{2.2.16}$$

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

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 $O \in SO(3)$,

$$O^TO = Id\,, \quad \text{so that} \quad O^{-1} = O^T \quad \text{and} \quad \det O = 1\,. \tag{2.2.17}$$

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 $\widehat{\Omega}$,

$$\widehat{\Omega}(t) = O^{-1}(t)\dot{O}(t)\,, \tag{2.2.18}$$

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 $(O^T(t)O(t))^{\cdot} = 0$, so that

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

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

$$\delta\widehat{\Omega} = \widehat{\Xi}^{\,\boldsymbol{\cdot}} + \widehat{\Omega}\widehat{\Xi} - \widehat{\Xi}\widehat{\Omega}\,, \tag{2.2.19}$$

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

Proof. The variational formula (2.2.19) follows by subtracting the time derivative $\widehat{\Xi}^{\,\boldsymbol{\cdot}} = (O^{-1}\delta O)^{\boldsymbol{\cdot}}$ from the variational derivative $\delta\widehat{\Omega} = \delta(O^{-1}\dot{O})$ in the relations

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

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

Theorem 2.2.36 (Geodesic motion on $SO(3)$)
The Euler-Lagrange equation for Hamilton's principle

$$\delta S = 0 \quad with \quad S = \int L(\widehat{\Omega})\,dt\,, \tag{2.2.20}$$

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

$$L(\widehat{\Omega}) = -\frac{1}{2}\mathrm{tr}(\widehat{\Omega}\mathbb{A}\widehat{\Omega})\,, \tag{2.2.21}$$

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

$$\frac{d\widehat{\Pi}}{dt} = -\,[\,\widehat{\Omega}\,,\,\widehat{\Pi}\,] \quad with \quad \widehat{\Pi} = \mathbb{A}\widehat{\Omega} + \widehat{\Omega}\mathbb{A} = \frac{\delta L}{\delta\widehat{\Omega}} = -\,\widehat{\Pi}^T\,. \tag{2.2.22}$$

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

$$\begin{aligned}
\delta S &=: \int_a^b \Big\langle \frac{\delta L}{\delta \widehat{\Omega}} \,,\, \delta\widehat{\Omega} \Big\rangle \, dt \\
&= -\frac{1}{2}\int_a^b \mathrm{tr}\left(\delta\widehat{\Omega}\, \frac{\delta L}{\delta \widehat{\Omega}} \right) dt \\
&= -\frac{1}{2}\int_a^b \mathrm{tr}\left(\delta\widehat{\Omega}\, (\mathbb{A}\widehat{\Omega} + \widehat{\Omega}\mathbb{A}) \right) dt \\
&= -\frac{1}{2}\int_a^b \mathrm{tr}\left(\delta\widehat{\Omega}\, \widehat{\Pi} \right) dt \\
&= \int_a^b \Big\langle \widehat{\Pi} \,,\, \delta\widehat{\Omega} \Big\rangle \, dt \,. \qquad (2.2.23)
\end{aligned}$$

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.,

$$\Big\langle M \,,\, N \Big\rangle =: \frac{1}{2}\mathrm{tr}\Big(M^T N \Big) = \frac{1}{2}\mathrm{tr}\Big(N^T M \Big) \,. \qquad (2.2.24)$$

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 (2.2.19) for $\delta\widehat{\Omega}$ into the variation of the action (2.2.23) leads to

$$\begin{aligned}
\delta S &= -\frac{1}{2}\int_a^b \mathrm{tr}\left(\delta\widehat{\Omega}\widehat{\Pi} \right) dt \\
&= -\frac{1}{2}\int_a^b \mathrm{tr}\left((\widehat{\Xi}^{\,\cdot} + \widehat{\Omega}\widehat{\Xi} - \widehat{\Xi}\widehat{\Omega})\, \widehat{\Pi} \right) dt \,. \qquad (2.2.25)
\end{aligned}$$

Permuting cyclically under the trace again yields

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

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

$$\delta S = -\frac{1}{2}\int_a^b \operatorname{tr}\left(\widehat{\Xi}\left(-\widehat{\Pi}^{\,\cdot} + \widehat{\Pi}\widehat{\Omega} - \widehat{\Omega}\widehat{\Pi}\right)\right) dt\,. \tag{2.2.26}$$

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 (2.2.22). ∎

Remark 2.2.37 (Interpretation of Theorem 2.2.36)
Equation (2.2.22) 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

$$\dot{O}(t) = O(t)\widehat{\Omega}(t)\,. \tag{2.2.27}$$

This is the ***reconstruction formula****, obtained from the definition (2.2.18) of the angular velocity matrix. In the classical literature, such an integration is called a* ***quadrature****.*

Corollary 2.2.38 *Formula (2.2.22) for the evolution of $\widehat{\Pi}(t)$ is equivalent to the conservation law*

$$\frac{d}{dt}\widehat{\pi}(t) = 0\,, \quad \textit{where} \quad \widehat{\pi}(t) := O(t)\widehat{\Pi}(t)O^{-1}(t)\,. \tag{2.2.28}$$

Proof. This may be verified by a direct computation that uses the reconstruction formula in (2.2.27). ∎

Remark 2.2.39 *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 1.1.16) 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 $\mathrm{tr}(\widehat{\Xi}\widehat{\Pi})|_a^b$ arising in the variation δS in (2.2.26) and invoke left-invariance of the Lagrangian (2.2.21) under $O \to U_\epsilon O$ with $U_\epsilon \in SO(3)$. For this symmetry transformation, $\delta O = \widehat{\Gamma} O$ with $\widehat{\Gamma} = \frac{d}{d\epsilon}|_{\epsilon=0} U_\epsilon$, so $\widehat{\Xi} = O^{-1}\widehat{\Gamma} O$. ★

2.2.4 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 (2.2.22) turns out to be the matrix version of Euler's equations for rigid body motion.

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 (2.1.11) to convert the skew-symmetric matrix dynamics (2.2.22) 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

$$\boldsymbol{\Omega} = \Omega_1 \mathbf{e}_1 + \Omega_2 \mathbf{e}_2 + \Omega_3 \mathbf{e}_3 \,, \tag{2.2.29}$$

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

$$\widehat{\Omega}_{ij} = -\,\epsilon_{ijk}\Omega_k \,, \tag{2.2.30}$$

which takes the skew-symmetric matrix form of (2.1.13)

$$\widehat{\Omega} = \begin{pmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{pmatrix} . \tag{2.2.31}$$

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

$$\widehat{\Pi} = \begin{pmatrix} 0 & -I_3\Omega_3 & I_2\Omega_2 \\ I_3\Omega_3 & 0 & -I_1\Omega_1 \\ -I_2\Omega_2 & I_1\Omega_1 & 0 \end{pmatrix} , \tag{2.2.32}$$

with

$$I_1 = a_2 + a_3 \,, \quad I_2 = a_1 + a_3 \,, \quad I_3 = a_1 + a_2 \,. \tag{2.2.33}$$

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

$$\mathbf{\Pi} = \Pi_1 \mathbf{e}_1 + \Pi_2 \mathbf{e}_2 + \Pi_3 \mathbf{e}_3 \,, \tag{2.2.34}$$

and the Lagrangian (2.2.21) in these vector components is expressed as

$$L = \frac{1}{2}\Big(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_1^3 \Big) \,, \tag{2.2.35}$$

with body angular momentum components,

$$\Pi_i = \frac{\delta L}{\delta \Omega_i} = I_i\Omega_i \,, \quad i = 1, 2, 3, \quad \text{(no sum)} \,. \tag{2.2.36}$$

In this vector representation, the matrix Euler-Lagrange equation (2.2.22) becomes Euler's equation for rigid body motion.

In vector form, Euler's equations are,

$$\dot{\mathbf{\Pi}} = -\,\mathbf{\Omega} \times \mathbf{\Pi} \,, \tag{2.2.37}$$

whose vector components are expressed as

$$\begin{aligned} I_1\dot{\Omega}_1 &= (I_2 - I_3)\Omega_2\Omega_3 = -(a_2 - a_3)\Omega_2\Omega_3 \,, \\ I_2\dot{\Omega}_2 &= (I_3 - I_1)\Omega_3\Omega_1 = -(a_3 - a_1)\Omega_3\Omega_1 \,, \\ I_3\dot{\Omega}_3 &= (I_1 - I_2)\Omega_1\Omega_2 = -(a_1 - a_2)\Omega_1\Omega_2 \,. \end{aligned} \tag{2.2.38}$$

Corollary. Equation (2.2.37) implies as in Corollary 2.2.38 that

$$\dot{\boldsymbol{\pi}}(t) = 0\,, \quad \text{for} \quad \boldsymbol{\pi}(t) = O(t)\boldsymbol{\Pi}(t)\,, \tag{2.2.39}$$

on using the hat map $O^{-1}\dot{O}(t) = \widehat{\Omega}(t) = \boldsymbol{\Omega}(t)\times$, as in (2.2.30).

Remark 2.2.40 (Two interpretations of Euler's equations)

1. *Euler's equations describe conservation of spatial angular momentum $\boldsymbol{\pi}(t) = O(t)\boldsymbol{\Pi}(t)$ 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 (2.2.33). 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.

2.3 Hamilton

Hamilton

Hamilton's approach to geometric optics sketched in Chapter 1 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.

2.3.1 Legendre transform

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

Definition 2.3.1 (Legendre transform and fibre derivative)
*The **Legendre transformation** is defined by using the **fibre derivative** of the Lagrangian,*

$$p = \frac{\partial L}{\partial \dot{q}} \,. \tag{2.3.1}$$

*The name **fibre derivative** refers to Definition 2.2.19 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 2.3.2 *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 2.3.3 (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****, cf. equation (1.2.8)*

$$H(p,q) = \langle\, p\,,\,\dot q\,\rangle - L(q,\dot q)\,. \tag{2.3.2}$$

Remark 2.3.4 *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.,*

$$\det\frac{\partial^2 L}{\partial\dot q\partial\dot q} = \det\frac{\partial p(q,\dot q)}{\partial\dot q} \neq 0 \quad \textit{(suppressing indices)}\,. \tag{2.3.3}$$

Definition 2.3.5 (Non-degenerate Lagrangian system)
A Lagrangian system (M,L) is said to be ***non-degenerate*** *if the* ***Hessian matrix***

$$H_L(q,\dot q) = \frac{\partial^2 L}{\partial\dot q\partial\dot q} \quad \textit{(again suppressing indices)} \tag{2.3.4}$$

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

Remark 2.3.6 *In Chapter 1, Remark 1.1.5, we saw 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 appeared in Fermat's principle for ray paths. That Lagrangian was 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. However, as we saw in Chapter 1, that case may be regularised by transforming to a related Riemannian description, in which the Lagrangian is quadratic in the tangent vector.

2.3.2 Hamilton's canonical equations

Theorem 2.3.7 (Hamiltonian equations) *When the Lagrangian is **non-degenerate** (hyperregular), the Euler-Lagrange equations*

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

*in (2.2.8) are equivalent to **Hamilton's canonical equations**,*

$$\dot{q} = \frac{\partial H}{\partial p}\,, \qquad \dot{p} = -\,\frac{\partial H}{\partial q}\,, \tag{2.3.5}$$

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 (2.3.2) 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 2.3.8 *The Euler-Lagrange equations are second order and they determine **curves in configuration space** $q \in M$. In contrast, Hamilton's equations are first order and they determine **curves in phase space** $(q,p) \in T^*M$, a space whose dimension is twice the dimension of the configuration space M.*

Definition 2.3.9 (Number of degrees of freedom)
The dimension of the configuration space is called the ***number of degrees of freedom****.*

Remark 2.3.10 *Each degree of freedom has its own coordinate and momentum in phase space.*

Remark 2.3.11 (Momentum vs position in phase space)
As discussed in Definition 3.3.11, 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_q^*M*. This basis for 1-forms in* T_q^*M *is dual to the vector basis* $\partial/\partial q^1, \dots, \partial/\partial q^n$ *for the tangent bundle* T_qM *at* $q = (q^1, \dots, q^n)$*.*

2.3.3 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

$$S = \int_{t_a}^{t_b} L(q, \dot{q}) + p\Big(\frac{dq}{dt} - \dot{q}\Big)\, dt\,, \tag{2.3.6}$$

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

$$\begin{aligned}\delta S &= \int_{t_a}^{t_b} \left(\frac{\partial L}{\partial q} - \frac{dp}{dt}\right)\delta q + \left(\frac{\partial L}{\partial \dot{q}} - p\right)\delta \dot{q} + \Big(\frac{dq}{dt} - \dot{q}\Big)\delta p\, dt \\ &\quad + \Big[p\,\delta q\Big]_{t_a}^{t_b}. \end{aligned} \tag{2.3.7}$$

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

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

- 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 (2.3.2) into the constrained action (2.3.6) yields the ***phase space action***

$$
S = \int_{t_a}^{t_b} \Big(p\frac{dq}{dt} - H(q,p) \Big)\, dt\,. \tag{2.3.8}
$$

Varying the phase space action in (2.3.8) 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 (2.3.8) recover Hamilton's canonical equations (2.3.5).

Hamiltonian evolution along a curve $(q(t), p(t)) \in T^*M$ satisfying equations (2.3.5) 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

$$
\begin{aligned}
\frac{dF}{dt} &= \frac{\partial F}{\partial q}\frac{dq}{dt} + \frac{\partial H}{\partial q}\frac{dp}{dt} \\
&= \frac{\partial F}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial H}{\partial q}\frac{\partial F}{\partial p} =: \{F\,,\,H\} && (2.3.9) \\
&= \left(\frac{\partial H}{\partial p}\frac{\partial}{\partial q} - \frac{\partial H}{\partial q}\frac{\partial}{\partial p} \right) F =: X_H F\,. && (2.3.10)
\end{aligned}
$$

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*** $X_H = \{\,\cdot\,,\,H\,\}$.

2.3.4 Poisson brackets

Definition 2.3.12 (Canonical Poisson bracket)
*Hamilton's canonical equations are associated to the **canonical Poisson bracket** for functions on phase space, defined by*

$$\dot{p} = \{p\,,\,H\}\,,\quad \dot{q} = \{q\,,\,H\}\,. \tag{2.3.11}$$

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

$$\dot{F}(q,p) = \{F\,,\,H\} = \frac{\partial F}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial H}{\partial q}\frac{\partial F}{\partial p}\,. \tag{2.3.12}$$

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

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

namely,

$$dF \wedge dH = \det\frac{\partial(F,H)}{\partial(q,p)}dq \wedge dp = \{F,H\}dq \wedge dp\,. \tag{2.3.13}$$

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 2.3.14 (The canonical Poisson bracket)
*The definition of the **canonical Poisson bracket** in (2.3.12) 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*.*

2.3.5 Canonical transformations

Definition 2.3.15 (Transformation)
A ***transformation*** *is a one-to-one mapping of a set onto itself.*

Example 2.3.16 *For example, under a change of variables*

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

in phase space T^*M*, the Poisson bracket in (2.3.13) transforms via the Jacobian determinant, as*

$$\begin{aligned} dF \wedge dH &= \{F, H\} dq \wedge dp \\ &= \{F, H\} \det \frac{\partial(q,p)}{\partial(Q,P)} dQ \wedge dP\,. \end{aligned} \tag{2.3.14}$$

Definition 2.3.17 (Canonical transformations)
When the Jacobian determinant is equal to unity, that is, when

$$\det \frac{\partial(q,p)}{\partial(Q,P)} = 1\,, \quad \textit{so that} \quad dq \wedge dp = dQ \wedge dP\,, \tag{2.3.15}$$

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*

$$\dot{P} = \{P\,,\,H\}\,,\quad \dot{Q} = \{Q\,,\,H\}\,. \tag{2.3.16}$$

Remark 2.3.18 *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*

$$\det \frac{\partial(q,p)}{\partial(Q,P)} = \text{constant}\,, \tag{2.3.17}$$

may still be said to be canonical.

Definition 2.3.19 (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 2.3.20 *The canonical transformations form a group.*

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

Remark 2.3.21 *The smooth parameter dependence needed to show that the canonical transformations actually form a Lie group will arise from their definition in terms of the Poisson bracket.*

2.3.6 Flows of Hamiltonian vector fields

The Leibnitz property (product rule) in Proposition 2.3.14 suggests the canonical Poisson bracket is a type of derivative. This derivation property of the Poisson bracket allows its use in the definition of a ***Hamiltonian vector field***.

Definition 2.3.22 (Hamiltonian vector field)
The Poisson bracket expression

$$X_H = \{\cdot\,,\,H\} = \frac{\partial H}{\partial p}\frac{\partial}{\partial q} - \frac{\partial H}{\partial q}\frac{\partial}{\partial p}\,, \tag{2.3.18}$$

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

Proposition 2.3.23 *Solutions of Hamilton's canonical equations $q(t)$ and $p(t)$ are the **characteristic paths** of the first order linear partial differential operator X_H. That is, X_H corresponds to the time derivative along these characteristic paths.*

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

Definition 2.3.24 (Hamiltonian flow)
*The union of the characteristic paths of the Hamiltonian vector field X_H in phase space T^*M is called the **flow** of the Hamiltonian vector field X_H. That is, the flow of X_H is the collection of maps $\phi_t : T^*M \to T^*M$ satisfying*

$$\frac{d\phi_t}{dt} = X_H\big(\phi_t(q,p)\big) = \{\phi_t,\,H\}\,, \tag{2.3.19}$$

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

Theorem 2.3.25 *Canonical transformations result from the smooth flows of Hamiltonian vector fields. That is, Poisson brackets generate canonical transformations.*

Proof. According to Definition 2.3.17, a transformation

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

which depends smoothly on a parameter ϵ is canonical, provided it preserves area in phase space (up to a constant factor that defines the units of area). That is, provided it satisfies the condition in equation (2.3.15), discussed further in Chapter 3, namely

$$dq(\epsilon) \wedge dp(\epsilon) = dq(0) \wedge dp(0)\,. \tag{2.3.20}$$

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

$$\begin{aligned} \frac{d}{d\epsilon}\Big(dq(\epsilon) \wedge dp(\epsilon)\Big) &= d(X_F q) \wedge dp + dq \wedge d(X_F p) \\ &= d(F_{,p}) \wedge dp + dq \wedge d(F_{,q}) \\ &= (F_{,pq}dq + F_{,pp}dp) \wedge dp + dq \wedge (F_{,qq}dq + F_{,qp}dp) \\ &= (F_{,pq} - F_{,qp})dq \wedge dp \\ &= 0\,, \end{aligned} \tag{2.3.21}$$

by equality of cross derivatives of F and asymmetry of the wedge product. Therefore, condition (2.3.20) holds and the transformation is canonical. ■

Corollary 2.3.26 *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 (2.3.8) 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 (2.3.8) satisfies $\delta S = 0$ for $\delta q = \xi_M(q) \in TM$.

What does Noether's theorem (Corollary 1.1.16) imply for this phase space action principle? ★

Answer. Noether's theorem implies conservation of the quantity

$$J^\xi = \big\langle p,\, \xi_M(q) \big\rangle_{T^*M \times TM} \in \mathbb{R}\,, \tag{2.3.22}$$

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,

$$\frac{dJ^\xi}{dt} = \{J^\xi,\, H\} = 0\,. \tag{2.3.23}$$

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

$$\begin{aligned} 0 = X_{J^\xi} H &= \frac{\partial J^\xi}{\partial p}\frac{\partial H}{\partial q} - \frac{\partial J^\xi}{\partial q}\frac{\partial H}{\partial p} \\ &= \xi_M(q)\partial_q H - p\xi_M'(q)\partial_p H \\ &= \frac{d}{d\epsilon}\Big|_{\epsilon=0} H\big(q(\epsilon),\, p(\epsilon)\big)\,. \end{aligned} \tag{2.3.24}$$

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 (2.3.22) will be a constant of the canonical motion under H.

▲

Definition 2.3.27 (Cotangent lift momentum map)
On introducing a pairing $\langle\,\cdot\,,\,\cdot\,\rangle : \mathfrak{g}^* \times \mathfrak{g} \to \mathbb{R}$ one may define a map $J : T^*M \to \mathfrak{g}^*$ in terms of this pairing and the Noether endpoint quantity in (2.3.22 as

$$J^\xi = \big\langle p,\, \xi_M(q)\big\rangle_{T^*M\times TM} =: \big\langle J(q,p),\, \xi\big\rangle_{\mathfrak{g}^*\times\mathfrak{g}}\,, \tag{2.3.25}$$

*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.** Compare Definition 2.3.27 with Definition 1.3.20, by identifying corresponding terms. ★

> **Exercise. (Cotangent lift momentum maps are Poisson)** Show that cotangent lift momentum maps are Poisson. That is, show that, for smooth functions F and H,
>
> $$\big\{F\circ J\,,\, H\circ J\big\} = \big\{F\,,\, H\big\}\circ J\,. \tag{2.3.26}$$
>
> This relation defines a ***Lie-Poisson bracket*** on $\mathfrak{g}^*$ that inherits the properties in Proposition 2.3.14 of the canonical Poisson bracket.
>
> ★

2.3.7 Properties of Hamiltonian vector fields

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

Definition 2.3.28 (Hamiltonian vector field commutator)
*The **commutator** of the Hamiltonian vector fields X_F and X_H is defined as*

$$[X_F\,,\, X_H] = X_F X_H - X_H X_F\,, \tag{2.3.27}$$

which is again a Hamiltonian vector field.

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

Lemma 2.3.29 *Hamiltonian vector fields satisfy the Jacobi identity,*

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

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

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

Summation over cyclic permutations then yields the result. ∎

Lemma 2.3.30 *The Jacobi identity holds for the canonical Poisson bracket* $\{\cdot\,,\,\cdot\}$*,*

$$\{F,\,\{G,\,H\}\}+\{G,\,\{H,\,F\}\}+\{H,\,\{F,\,G\}\}=0\,. \tag{2.3.28}$$

Proof. Formula (2.3.28) may be proved by direct computation, as in Proposition 2.3.14. This identity may also be verified formally by the same calculation as in the proof of the previous Lemma, by writing $\{G,\,H\}=G(H)-H(G)$ symbolically. ∎

Remark 2.3.31 (Lie algebra of Hamiltonian vector fields)
The Jacobi identity defines the Lie algebra property of Hamiltonian vector fields, which form a Lie subalgebra of all vector fields on phase space.

Theorem 2.3.32 (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*

$$X_{\{F\,,\,H\}}=-\,[X_F\,,\,X_H]\,. \tag{2.3.29}$$

Proof. One recalls the remark after Lemma 1.11.4,

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

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 (2.3.28) and skew symmetry. ■

2.4 Rigid-body motion

2.4.1 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 2.2.36 for geodesic motion on $SO(3)$ – interpreted also as rigid body dynamics – then its simple, beautiful and well-known Hamiltonian formulation emerges.

Definition 2.4.1 *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 (2.2.35),

$$L(\mathbf{\Omega}) = \frac{1}{2}\,\mathbf{\Omega}\cdot\mathbb{I}\mathbf{\Omega}\,, \tag{2.4.1}$$

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

$$\mathbf{\Pi} = \frac{\delta L}{\delta\mathbf{\Omega}} = \mathbb{I}\mathbf{\Omega}\,, \tag{2.4.2}$$

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

$$\Pi_i = I_i\Omega_i = \frac{\partial L}{\partial\Omega_i}\,, \quad i = 1, 2, 3, \tag{2.4.3}$$

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 2.1.25 in the Newtonian setting.

2.4.2 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 (2.4.1), one finds the expected expression for the rigid body Hamiltonian,

$$H(\mathbf{\Pi}) = \frac{1}{2}\,\mathbf{\Pi}\cdot\mathbb{I}^{-1}\mathbf{\Pi} := \frac{\Pi_1^2}{2I_1} + \frac{\Pi_2^2}{2I_2} + \frac{\Pi_3^2}{2I_3}\,. \tag{2.4.4}$$

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:

$$\frac{d}{dt}(\mathbb{I}\boldsymbol{\Omega}) = \mathbb{I}\boldsymbol{\Omega} \times \boldsymbol{\Omega} \Longleftrightarrow \frac{d\boldsymbol{\Pi}}{dt} = \boldsymbol{\Pi} \times \frac{\partial H}{\partial \boldsymbol{\Pi}} := \{\boldsymbol{\Pi}, H\} \,. \tag{2.4.5}$$

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

$$\{F, H\}(\boldsymbol{\Pi}) := -\,\boldsymbol{\Pi} \cdot \left(\frac{\partial F}{\partial \boldsymbol{\Pi}} \times \frac{\partial H}{\partial \boldsymbol{\Pi}} \right), \tag{2.4.6}$$

or, in components,

$$\{\Pi_j, \Pi_k\} = -\,\Pi_i \epsilon_{ijk} \,. \tag{2.4.7}$$

For the Hamiltonian (2.4.4), one checks that the Euler equations in terms of the rigid body angular momenta,

$$\begin{aligned} \frac{d\Pi_1}{dt} &= -\left(\frac{1}{I_2} - \frac{1}{I_3}\right)\Pi_2\Pi_3 \,, \\ \frac{d\Pi_2}{dt} &= -\left(\frac{1}{I_3} - \frac{1}{I_1}\right)\Pi_3\Pi_1 \,, \\ \frac{d\Pi_3}{dt} &= -\left(\frac{1}{I_1} - \frac{1}{I_2}\right)\Pi_1\Pi_2 \,, \end{aligned} \tag{2.4.8}$$

that is, the equations in vector form,

$$\frac{d\boldsymbol{\Pi}}{dt} = -\,\boldsymbol{\Omega} \times \boldsymbol{\Pi} \,, \tag{2.4.9}$$

are equivalent to

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

The Poisson bracket proposed in (2.4.6) may be rewritten in terms of coordinates $\mathbf{\Pi} \in \mathbb{R}^3$ as

$$\{F, H\} = -\nabla \frac{|\mathbf{\Pi}|^2}{2} \cdot \nabla F \times \nabla H \,, \tag{2.4.10}$$

where ∇ denotes $\partial/\partial\mathbf{\Pi}$. This is an example of the Nambu $\mathbb{R}^3$ bracket in (1.11.17), which we learned in Chapter 1 satisfies the defining relations to be a Poisson bracket. 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 (2.4.9) rewritten as

$$\frac{d\mathbf{\Pi}}{dt} = \Big\{\mathbf{\Pi}, H\Big\} = \nabla\frac{|\mathbf{\Pi}|^2}{2} \times \nabla H \tag{2.4.11}$$

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 (2.4.4).

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

Euler's equations (2.4.11) are expressible in vector form as

$$\frac{d}{dt}\mathbf{\Pi} = \nabla L \times \nabla H \,, \tag{2.4.12}$$

where H is the rotational kinetic energy

$$H = \frac{\Pi_1^2}{2I_1} + \frac{\Pi_2^2}{2I_2} + \frac{\Pi_3^2}{2I_3} \,, \tag{2.4.13}$$

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

$$L = \frac{1}{2}(\Pi_1^2 + \Pi_2^2 + \Pi_3^2) \,, \tag{2.4.14}$$

with gradient

$$\nabla L = (\Pi_1,\, \Pi_2,\, \Pi_3) \,. \tag{2.4.15}$$

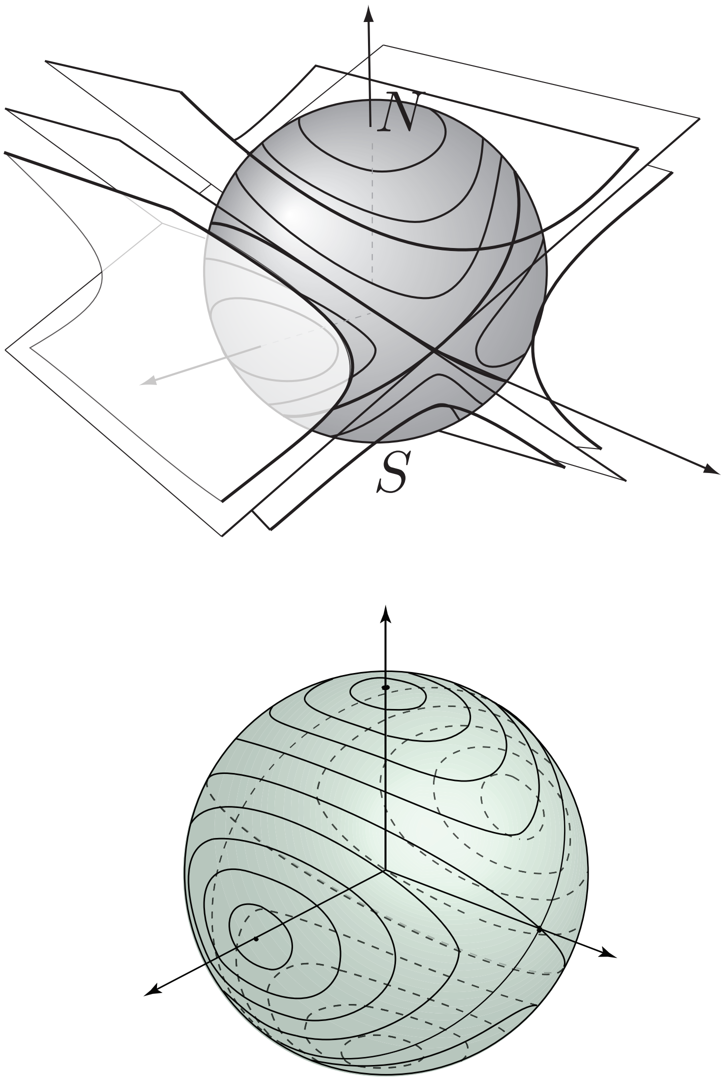


Figure 2.5: The dynamics of a rotating rigid body may be represented as a divergenceless flow along the intersections in $\mathbb{R}^3$ of the level sets of two conserved quantities: the angular momentum sphere $|\mathbf{\Pi}|^2 = const$ and the hyperbolic cylinders $G = const$ in equation (2.4.18).

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 2.4.2 *Euler's equations for the rigid body (2.4.12) may be written equivalently as*

$$\frac{d}{dt}\mathbf{\Pi} = \nabla L \times \nabla G \,, \quad \textit{where} \quad G = H - \frac{L}{I_2}\,, \tag{2.4.16}$$

or, explicitly, L and G are given by

$$L = \frac{1}{2}(\Pi_1^2 + \Pi_2^2 + \Pi_3^2) \tag{2.4.17}$$

and

$$G = \Pi_1^2\left(\frac{1}{2I_1} - \frac{1}{2I_2}\right) - \Pi_3^2\left(\frac{1}{2I_2} - \frac{1}{2I_3}\right). \tag{2.4.18}$$

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

$$\frac{d}{dt}\mathbf{\Pi} = \nabla L \times \nabla H = \nabla L \times \nabla\left(H - \frac{L}{I_2}\right) = \nabla L \times \nabla G\,. \tag{2.4.19}$$

■

Remark 2.4.3 *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 2.5.*

2.4.4 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 [HoMa1991] 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 (2.4.10).

Theorem 2.4.4 *Euler's equations for the rigid body (2.4.12) may be written equivalently as*

$$\frac{d}{dt}\mathbf{\Pi} = \nabla A \times \nabla B\,, \tag{2.4.20}$$

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

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{bmatrix} a & b \\ c & e \end{bmatrix} \begin{pmatrix} H \\ L \end{pmatrix}, \tag{2.4.21}$$

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

Proof. Recall from equation (2.4.6) that

$$\begin{aligned} \{F, H\}d^3\Pi \;&:=\; dF \wedge dL \wedge dH \\ &= \frac{1}{ae - bc}\, dF \wedge d(aH + bL) \wedge d(cH + eL)\,, \end{aligned} \tag{2.4.22}$$

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 2.4.5 (Equilibria)
For a general choice for the linear combination of A and B in (2.4.21), 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 2.4.6 *Euler's equations for the rigid body (2.4.12) may be written equivalently as*

$$\frac{d}{dt}\mathbf{\Pi} = \nabla N \times \nabla K \,, \tag{2.4.23}$$

where K and N are

$$K = \frac{\Pi_1^2}{2k_1^2} + \frac{\Pi_2^2}{2k_2^2} \qquad \textit{and} \qquad N = \frac{\Pi_2^2}{2k_3^2} + \frac{1}{2}\Pi_3^2 \,, \tag{2.4.24}$$

for

$$\frac{1}{k_1^2} = \frac{1}{I_1} - \frac{1}{I_3}, \qquad \frac{1}{k_2^2} = \frac{1}{I_2} - \frac{1}{I_3}, \qquad \frac{1}{k_3^2} = \frac{I_3(I_2 - I_1)}{I_2(I_3 - I_1)} \,. \tag{2.4.25}$$

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

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

yields

$$\{F, H\} d^3\Pi := dF \wedge dL \wedge dH = dF \wedge dN \wedge dK \,, \tag{2.4.26}$$

from which equations (2.4.23) - (2.4.25) 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

$$H = eN - bK = \frac{1}{I_3} N + \frac{I_3}{I_3 - I_1} K \,. \tag{2.4.27}$$

Consequently, we may write

$$dF \wedge dL \wedge dH = dF \wedge dN \wedge dK = -I_3\, dF \wedge dK \wedge dH\,. \qquad (2.4.28)$$

■

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.

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

$$d^{\,3}\Pi := d\Pi_1 \wedge d\Pi_2 \wedge d\Pi_3 = k_1 k_2\, dK \wedge d\theta \wedge dp\,.$$

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 (2.4.26) as

$$\begin{aligned} \{F,\, H\}\, d^{\,3}\Pi &= -\, dL \wedge dF \wedge dH \\ &= k_1 k_2\, dK \wedge \{F,\, H\}_{\text{EllipCyl}}\, d\theta \wedge dp\,. \end{aligned} \tag{2.4.29}$$

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

$$\{F, H\}_{\text{EllipCyl}} = \frac{1}{k_1 k_2}\left(\frac{\partial F}{\partial p}\frac{\partial H}{\partial \theta} - \frac{\partial F}{\partial \theta}\frac{\partial H}{\partial p}\right),$$

which is *symplectic*. In particular, it satisfies

$$\{p, \theta\}_{\text{EllipCyl}} = \frac{1}{k_1 k_2}\,. \tag{2.4.30}$$

The restriction of the Hamiltonian H to the symplectic level set of the elliptic cylinder $K = constant$ is by (2.4.13)

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

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

$$\begin{aligned} \frac{d\theta}{dt} &= \left\{\theta, \frac{N}{I_3}\right\}_{\text{EllipCyl}} = \frac{1}{k_1 k_2 I_3}\frac{\partial N}{\partial p} = -\frac{1}{k_1 k_2 I_3} p\,, \\ \frac{dp}{dt} &= \left\{p, \frac{N}{I_3}\right\}_{\text{EllipCyl}} = \frac{1}{k_1 k_2 I_3}\frac{\partial N}{\partial \theta} = \frac{1}{k_1 k_2 I_3}\frac{k_2^2}{k_3^2} r^2 \sin\theta\cos\theta\,. \end{aligned}$$

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

$$\frac{d^2\theta}{dt^2} = -\frac{r^2}{k_1 k_2 I_3}\sin 2\theta\,.$$

In terms of the original rigid body parameters, this becomes

$$\frac{d^2\theta}{dt^2} = -\frac{K}{I_3^2}\left(\frac{1}{I_1} - \frac{1}{I_2}\right)\sin 2\theta\,. \tag{2.4.31}$$

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

Proposition 2.4.7 *Rigid body motion reduces to pendulum motion on level surfaces of K.*

Corollary 2.4.8 *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 2.6.*

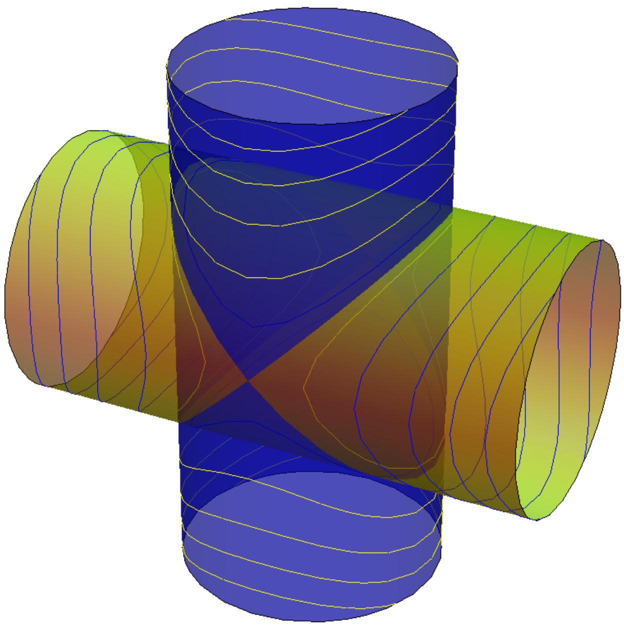


Figure 2.6: The dynamics of the rigid body in three-dimensional body angular momentum space is recovered by taking the union in $\mathbb{R}^3$ of the intersections of level surfaces of two orthogonal families of concentric cylinders. (Only one member of each family is shown in the figure here, although 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 (2.4.31).

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 $\theta = 0; \pi/2,\ \pi$ 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.

2.5 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 is treated as an enhanced coursework example by using spherical polar coordinates and the traditional methods of Newton, Lagrange and Hamilton in Appendix A. 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 $S^2 \in \mathbb{R}^3$:

$$TS^2 = \big\{(\mathbf{x}, \dot{\mathbf{x}}) \in T\mathbb{R}^3 \simeq \mathbb{R}^6 \;\big|\; 1 - |\mathbf{x}|^2 = 0, \mathbf{x} \cdot \dot{\mathbf{x}} = 0\big\}. \tag{2.5.1}$$

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.

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

$$L(\mathbf{x}, \dot{\mathbf{x}}) = \tfrac{1}{2}|\dot{\mathbf{x}}|^2 - g\hat{\mathbf{e}}_3 \cdot \mathbf{x} - \tfrac{1}{2}\mu(1 - |\mathbf{x}|^2), \tag{2.5.2}$$

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

$$\ddot{\mathbf{x}} = -g\hat{\mathbf{e}}_3 + \mu\mathbf{x}\,. \tag{2.5.3}$$

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

$$\mu = g\hat{\mathbf{e}}_3 \cdot \mathbf{x} - |\dot{\mathbf{x}}|^2\,. \tag{2.5.4}$$

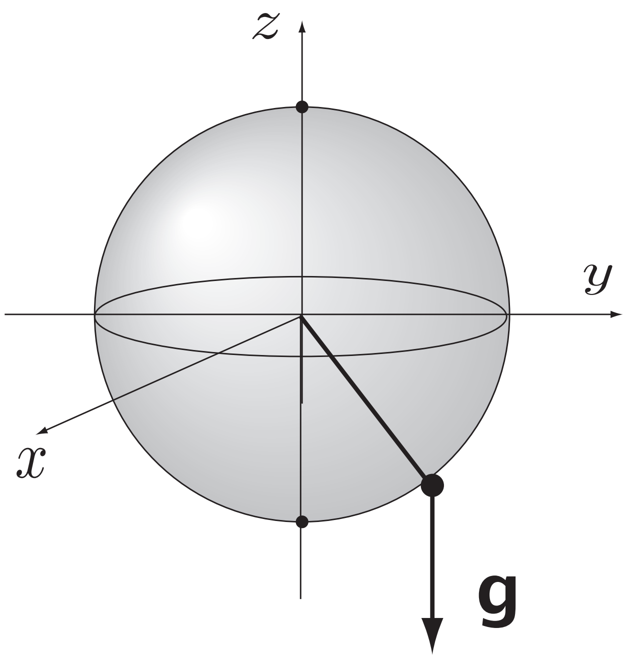


Figure 2.7: Spherical pendulum moving under gravity on TS^2 in $\mathbb{R}^3$.

Remark 2.5.1 *In Newtonian mechanics, the motion equation obtained by substituting (2.5.4) into (2.5.3) 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 (2.5.2) 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 1.1.16) 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 (2.5.3) conserve

$$J_3(\mathbf{x}, \dot{\mathbf{x}}) = \dot{\mathbf{x}} \cdot \delta \mathbf{x} = \mathbf{x} \times \dot{\mathbf{x}} \cdot \hat{\mathbf{e}}_3 \,, \tag{2.5.5}$$

which is the angular momentum about the vertical axis.

Legendre transform and canonical equations The fibre derivative of the Lagrangian L in (2.5.2) is

$$\mathbf{y} = \frac{\partial L}{\partial \dot{\mathbf{x}}} = \dot{\mathbf{x}} \,. \tag{2.5.6}$$

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

$$H(\mathbf{x}, \mathbf{y}) = \tfrac{1}{2}|\mathbf{y}|^2 + g\hat{\mathbf{e}}_3 \cdot \mathbf{x} + \tfrac{1}{2}(g\hat{\mathbf{e}}_3 \cdot \mathbf{x} - |\mathbf{y}|^2)(1 - |\mathbf{x}|^2) \,, \tag{2.5.7}$$

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

$$\dot{\mathbf{x}} = \mathbf{y} \quad \text{and} \quad \dot{\mathbf{y}} = -g\hat{\mathbf{e}}_3 + (g\hat{\mathbf{e}}_3 \cdot \mathbf{x} - |\mathbf{y}|^2)\mathbf{x} \,. \tag{2.5.8}$$

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 (2.5.1). On TS^2, equations (2.5.8) conserve the energy

$$E(\mathbf{x}, \mathbf{y}) = \tfrac{1}{2}|\mathbf{y}|^2 + g\hat{\mathbf{e}}_3 \cdot \mathbf{x} \,, \tag{2.5.9}$$

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

$$\begin{aligned} X_{J_3} = \{\,\cdot\,,\, J_3\} &= \frac{\partial J_3}{\partial \mathbf{y}} \cdot \frac{\partial}{\partial \mathbf{x}} - \frac{\partial J_3}{\partial \mathbf{x}} \cdot \frac{\partial}{\partial \mathbf{y}} \\ &= \hat{\mathbf{e}}_3 \times \mathbf{x} \cdot \frac{\partial}{\partial \mathbf{x}} + \hat{\mathbf{e}}_3 \times \mathbf{y} \cdot \frac{\partial}{\partial \mathbf{y}} \,, \end{aligned} \tag{2.5.10}$$

for infinitesimal rotations about the vertical axis $\hat{\mathbf{e}}_3$. Because of the S^1 symmetry of the Hamiltonian in (2.5.7) under these rotations, we have the conservation law,

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

2.5.1 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]

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

Quotient map The transformation defined by

$$
\pi : \ (\mathbf{x}, \mathbf{y}) \to \{\sigma_j(\mathbf{x}, \mathbf{y}), \ j = 1, \dots, 6\} \tag{2.5.11}
$$

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 (2.5.10).

The six S^1-invariants that define the quotient map in (2.5.11) for the spherical pendulum satisfy the cubic algebraic relation

$$
\sigma_5^2 + \sigma_6^2 = \sigma_4(\sigma_3 - \sigma_2^2) \, . \tag{2.5.12}
$$

They also satisfy the positivity conditions

$$
\sigma_4 \geq 0, \quad \sigma_3 \geq \sigma_2^2. \tag{2.5.13}
$$

In these variables, the defining relations (2.5.1) for TS^2 become

$$
\sigma_4 + \sigma_1^2 = 1 \quad \text{and} \quad \sigma_5 + \sigma_1 \sigma_2 = 0 \, . \tag{2.5.14}
$$

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 (2.5.12) – (2.5.14) 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 (2.5.14) 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

$$\sigma_2^2+\sigma_6^2=\sigma_3(1-\sigma_1^2)\,, \tag{2.5.15}$$

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

$$C(\sigma_1,\sigma_2,\sigma_3;J_3^2)=\sigma_3(1-\sigma_1^2)-\sigma_2^2-J_3^2=0\,. \tag{2.5.16}$$

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,

$$\sigma_3=\frac{\sigma_2^2+J_3^2}{1-\sigma_1^2}\,. \tag{2.5.17}$$

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

$$H=\tfrac{1}{2}\sigma_3+g\sigma_1\,, \tag{2.5.18}$$

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 (2.5.16). Consequently, in $\mathbb{R}^3$-vector form, the motion is governed by the cross-product formula

$$\dot{\boldsymbol{\sigma}}=\frac{\partial C}{\partial\boldsymbol{\sigma}}\times\frac{\partial H}{\partial\boldsymbol{\sigma}}\,. \tag{2.5.19}$$

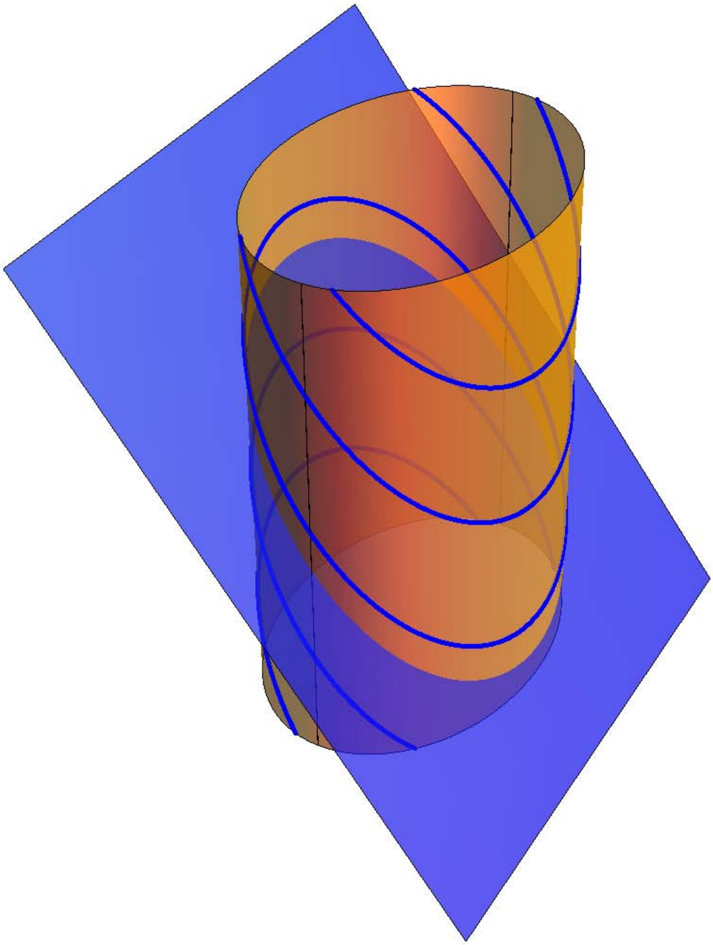


Figure 2.8: The dynamics of the spherical pendulum in the space of S^1 invariants $(\sigma_1, \sigma_2, \sigma_3)$ is recovered by taking the union in $\mathbb{R}^3$ of the intersections of level sets of two families of surfaces. These surfaces are the roughly cylindrical level sets of angular momentum about the vertical axis given in (2.5.17) and the (planar) level sets of the Hamiltonian in (2.5.18). (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 (2.5.24).

In components, this evolution is expressed as

$$\dot{\sigma}_i = \{\sigma_i, H\} = \epsilon_{ijk} \frac{\partial C}{\partial \sigma_j} \frac{\partial H}{\partial \sigma_k} \quad \text{with} \quad i, j, k = 1, 2, 3. \tag{2.5.20}$$

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

$$\{F, H\} = -\frac{\partial C}{\partial \boldsymbol{\sigma}} \cdot \frac{\partial F}{\partial \boldsymbol{\sigma}} \times \frac{\partial H}{\partial \boldsymbol{\sigma}} = -\,\epsilon_{ijk} \frac{\partial C}{\partial \sigma_i} \frac{\partial F}{\partial \sigma_j} \frac{\partial H}{\partial \sigma_k} \,. \tag{2.5.21}$$

This is another example of the Nambu $\mathbb{R}^3$ bracket in (1.11.17) introduced in [Na1973], which we learned in Chapter 1 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 (2.5.16) 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 (2.5.21) 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 (2.5.21) ensures that the orbit map for the spherical pendulum $\pi : T\mathbb{R}^3 \to \mathbb{R}^6$ in (B.7.1) is a ***Poisson map***. That is, the subspace obtained by using the relations (2.5.14) 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,

$$\{\sigma_i, \sigma_j\} = \epsilon_{ijk} \frac{\partial C}{\partial \sigma_k} \,, \tag{2.5.22}$$

with C given in (2.5.16). 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 (2.5.10).

Remark 2.5.2 The proof that the Nambu bracket in (2.5.21) satisfies the defining properties in Proposition 2.3.14 that are required for it to be a genuine Poisson bracket is given in Section 1.11.3. Another example of the Nambu bracket will be discussed in Section 4.4.4.

Reduced motion: Restriction in $\mathbb{R}^3$ to Hamiltonian planes The individual components of the equations of motion may be obtained from (2.5.20) as

$$\dot{\sigma}_1 = -\sigma_2 \,, \qquad \dot{\sigma}_2 = \sigma_1\sigma_3 + g(1-\sigma_1^2) \,, \qquad \dot{\sigma}_3 = 2g\sigma_2 \,. \quad (2.5.23)$$

Substituting $\sigma_3 = 2(H - g\sigma_1)$ from equation (2.5.18) and setting the acceleration of gravity to be unity $g = 1$ yields

$$\ddot{\sigma}_1 = 3\sigma_1^2 - 2H\sigma_1 - 1 \quad (2.5.24)$$

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

$$\tfrac{1}{2}\dot{\sigma}_1^2 + V(\sigma_1) = E \quad (2.5.25)$$

with the potential $V(\sigma_1)$ parameterised by H in (2.5.18) and given by

$$V(\sigma_1) = -\,\sigma_1^3 + H\sigma_1^2 + \sigma_1 \quad (2.5.26)$$

Equation (2.5.25) 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 2.9.

Each curve in the lower panel of Figure 2.9 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 (2.5.18) with a level set of J_3^2 given by $C = 0$ in equation (2.5.16). 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 (2.5.25) starts and ends with zero velocity at the North pole of the unit sphere. Refer to Section A.6 for a discussion of the properties of motion in a cubic potential and the details of how to compute its homoclinic trajectory.

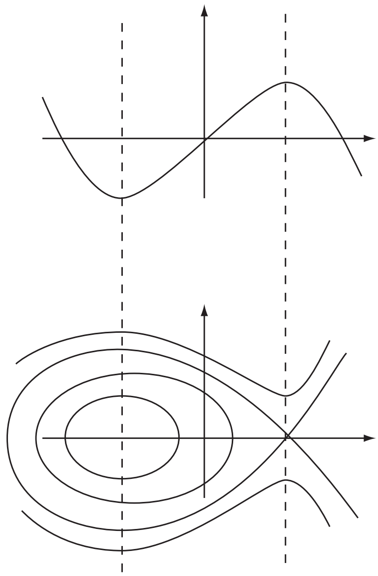


Figure 2.9: The upper panel shows a sketch of the cubic potential $V(\sigma_1)$ in equation (2.5.26) for the case $H = 0$. For $H = 0$, the potential has three zeros located at $\sigma_1 = 0, \pm 1$ and two critical points (relative equilibria) at $\sigma_1 = -\sqrt{3}/3$ (centre) and $\sigma_1 = +\sqrt{3}/3$ (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 (2.5.25) for $H = 0$.

2.5.2 Geometric phase for the spherical pendulum

We write the Nambu bracket (2.5.21) for the spherical pendulum as a differential form in $\mathbb{R}^3$

$$\{F, H\}\, d^3\sigma = dC \wedge dF \wedge dH \,, \tag{2.5.27}$$

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

$$\{f, h\} d\sigma_1 \wedge d\sigma_2 = df \wedge dh = \left(\frac{\partial f}{\partial \sigma_1} \frac{\partial h}{\partial \sigma_2} - \frac{\partial f}{\partial \sigma_2} \frac{\partial h}{\partial \sigma_1} \right) d\sigma_1 \wedge d\sigma_2 \tag{2.5.28}$$

and we recover equation (2.5.24) in canonical form with Hamiltonian

$$h(\sigma_1, \sigma_2) = -\Big(\tfrac{1}{2}\sigma_2^2 - \sigma_1^3 + H\sigma_1^2 + \sigma_1\Big) = -\Big(\tfrac{1}{2}\sigma_2^2 + V(\sigma_1)\Big), \quad (2.5.29)$$

which, not unexpectedly, is also the conserved energy integral in (2.5.25) for motion on level sets of H.

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

$$p_i dq_i = \sigma_2\, d\sigma_1 + H d\psi\,, \quad (2.5.30)$$

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 (4.3.1) gives

$$H d\psi = -\,\sigma_2\, d\sigma_1 + p_i dq_i\,. \quad (2.5.31)$$

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:

$$\oint H\, d\psi = H\,\Delta\psi = \underbrace{-\oint \sigma_2\, d\sigma_1}_{\text{geometric}} + \underbrace{\oint p_i dq_i}_{\text{dynamic}}\,. \quad (2.5.32)$$

On writing this decomposition of the phase as

$$\Delta\psi = \Delta\psi_{geom} + \Delta\psi_{dyn}\,, \quad (2.5.33)$$

one sees from (2.5.23) that

$$H\Delta\psi_{geom} = \oint \sigma_2^2\, dt = \iint d\sigma_1 \wedge d\sigma_2 \quad (2.5.34)$$

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

$$H\Delta\psi_{dyn} = \oint p_i\, dq_i = \int_0^T (-\sigma_2\dot{\sigma}_1 + H\dot{\psi})\, dt\,. \tag{2.5.35}$$

Hence, from the canonical equations $\dot{\sigma}_1 = \partial h/\partial\sigma_2$ and $\dot{\psi} = \partial h/\partial H$ with Hamiltonian h in (2.5.29), we have

$$\begin{aligned} \Delta\psi_{dyn} &= \frac{1}{H}\int_0^T \left(\sigma_2\frac{\partial h}{\partial\sigma_2} + H\frac{\partial h}{\partial H}\right) dt \\ &= \frac{2T}{H}\Big(h + \big\langle V(\sigma_1)\big\rangle - \tfrac{1}{2}H\big\langle\sigma_1^2\big\rangle\Big) \\ &= \frac{2T}{H}\Big(h + \big\langle V(\sigma_1)\big\rangle\Big) - T\big\langle\sigma_1^2\big\rangle\,, \end{aligned} \tag{2.5.36}$$

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 (2.5.33) 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. The remaining homoclinic trajectory is determined as in Section A.6.

Chapter 3

Lie, Poincaré, Cartan: Differential forms

3.1 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 this chapter are, as follows.

- Define differential forms using exterior product (wedge product) in a local basis.
- Define the push-forward and pull-back of a differential form under a smooth invertible map.

- Define the operation of contraction, or substitution of a vector field into a differential form.
- Define the exterior derivative of a differential form.
- 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.
- Derive the various identities for Lie derivatives acting on differential forms and illustrate them using steady incompressible fluid flows as an example.
- Explain Nambu's bracket for divergenceless vector fields in $\mathbb{R}^3$ in the language of differential forms.
- Define the Hodge star operation and illustrate its application in Maxwell's equations.
- 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 3.1.1 (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,*

$$\frac{dq}{dt} = \frac{\partial H}{\partial p}\,, \qquad \frac{dp}{dt} = -\,\frac{\partial H}{\partial q}\,, \tag{3.1.1}$$

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 (3.1.1) is called a ***Hamiltonian flow****.*

Definition 3.1.2 (Symplectic 2-form)
The oriented area in phase space

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

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

Definition 3.1.3 (Sym·plec·tic)
From the Greek for plaiting, braiding or joining together.

Definition 3.1.4 (Symplectic flows)
*Flows that preserve area in phase space are said to be **symplectic**.*

Remark 3.1.5 *There is a close analogy here between oriented phase space area measured by the symplectic two-form $\omega = dq \wedge dp$ and the corresponding symplectic matrix transformations (1.6.1) that preserve the phase space area $S = \mathbf{q} \times \mathbf{p}$ in axisymmetric screen optics. The wedge product in the symplectic two-form $\omega = dq \wedge dp$ is the analog of the cross product of vectors in the skewness $S = \mathbf{q} \times \mathbf{p}$, which is invariant in axisymmetric ray optics.*

Theorem 3.1.6 (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 (2.3.20) in Theorem 2.3.25. 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. ∎

3.2 Preliminaries for exterior calculus

3.2.1 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 3.2.1 (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 3.2.2 (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 3.2.3 (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*

3.3.14 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.

3.2.2 Contraction

Definition 3.2.4 (Contraction)
*In exterior calculus, the operation of **contraction** denoted as $\rfloor$ 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**,*

$$\partial_q \rfloor dq = 1 = \partial_p \rfloor dp \,, \quad \textit{and} \quad \partial_q \rfloor dp = 0 = \partial_p \rfloor dq \,, \qquad (3.2.1)$$

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

$$X_H = \dot{q}\frac{\partial}{\partial q} + \dot{p}\frac{\partial}{\partial p} = H_p\partial_q - H_q\partial_p = \{\,\cdot\,,\,H\,\}\,, \qquad (3.2.2)$$

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

$$X_H \rfloor dq = H_p \quad \textit{and} \quad X_H \rfloor dp = -\,H_q\,. \qquad (3.2.3)$$

Definition 3.2.5 (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*

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

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 (3.2.1).

This calculation proves the following.

Theorem 3.2.6 (Hamiltonian vector field) *The Hamiltonian vector field* $X_H = \{\,\cdot\,,\,H\,\}$ *satisfies*

$$X_H \,\lrcorner\, \omega \quad = \quad dH \quad \textit{with} \quad \omega = dq\,\wedge\,dp\,. \tag{3.2.4}$$

Remark 3.2.7 *The purely geometric nature of relation (3.2.4) argues for it to be taken as the* ***definition*** *of a Hamiltonian vector field.*

Lemma 3.2.8 ($d^2 = 0$ 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 (3.2.4) also implies the following.

Corollary 3.2.9 *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 (3.2.4). Along $(dq/dt, dp/dt) = (\dot{q}, \dot{p}) = (H_p, -H_q)$, for a solution of Hamilton's equations, we have

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

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 3.2.10 (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 3.2.9 may be simply re-stated as

Corollary 3.2.11 (Poincaré's theorem)
The flow of a Hamiltonian vector field is symplectic.

Definition 3.2.12 (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 3.2.13 *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 3.3.18.*

Remark 3.2.14 (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 3.2.15 *By Corollary 3.2.11 of Poincaré's Theorem 3.1.6, the Hamiltonian phase flow g_t is a one-parameter group of canonical transformations.*

Theorem 3.2.16 (Preservation of Hamiltonian form)
Canonical transformations preserve Hamiltonian form.

Proof. The coordinate-free relation (3.2.4) 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 3.2.17 (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

$$\frac{d}{dt}\Big(d\,\frac{\partial L}{\partial \dot q}\,\wedge\,dq\Big) = 0\,, \tag{3.2.5}$$

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,

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

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

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

Proposition 3.2.18 *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 $a, b = 1, \dots, K$,

$$d\theta_L = \frac{\partial^2 L}{\partial \dot{q}^a \partial \dot{q}^b} d\dot{q}^b \wedge dq^a + \frac{\partial^2 L}{\partial \dot{q}^a \partial q^b} dq^b \wedge dq^a \,, \qquad (3.2.6)$$

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 (2.3.3) is non-degenerate. ■

Definition 3.2.19 (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 3.2.20 (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.*

3.2.3 Hamilton-Jacobi equation

Definition 3.2.21 (Steady generating functions)
A sufficient condition for a transformation $(Q, P) = (Q(p, q), P(p, q))$ to be canonical is that

$$P \cdot dQ - p \cdot dq = dF\,. \tag{3.2.7}$$

*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 3.2.22 (Time-dependent generating functions)

Generating functions based on the phase space action in (2.3.8) 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 (2.3.8) transforms as

$$p \cdot dq - H(q, p)dt = P \cdot dQ - K(Q, P)dt + dS\,, \tag{3.2.8}$$

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 (3.2.8), so that

$$\begin{aligned} dS &= \frac{\partial S}{\partial q} \cdot dq + \frac{\partial S}{\partial t}\, dt + \frac{\partial S}{\partial Q} \cdot dQ \\ &= p \cdot dq - H(q, p)dt - P \cdot dQ\,. \end{aligned} \tag{3.2.9}$$

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

$$\frac{\partial S}{\partial q} = p\,, \quad \frac{\partial S}{\partial Q} = -P\,, \quad \frac{\partial S}{\partial t} + H(q,p) = 0\,. \tag{3.2.10}$$

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

$$\frac{\partial S}{\partial t}(q,t,Q) + H\Big(q, \frac{\partial S}{\partial q}\Big) = 0\,. \tag{3.2.11}$$

Thus, the Hamilton-Jacobi equation (3.2.11) 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,

$$S(q,t,Q) = \int_{t_a}^{t} dS = \int_{t_a}^{t} p \cdot dq - H(q,p)\,dt\,. \tag{3.2.12}$$

In geometrical optics, the solution S of the Hamilton-Jacobi equation (3.2.11) is called ***Hamilton's characteristic function***.

Remark 3.2.23 (Hamilton's characteristic function in optics)

Hamilton's characteristic function S in (3.2.12) has an interesting interpretation in terms of geometric optics. As we saw in Chapter 1, 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]

$$\phi = \int \mathbf{k} \cdot d\mathbf{r} - \omega(\mathbf{k}, \mathbf{r})\,dt\,. \tag{3.2.13}$$

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 1.1 is given by $\int \omega dt$.

On comparing the phase relation in (3.2.13) to the Hamilton-Jacobi solution in equation (3.2.12), 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

$$\begin{aligned} \frac{d\mathbf{r}}{dt} &= \frac{\partial \omega}{\partial \mathbf{k}} = \frac{c}{n}\frac{\mathbf{k}}{k} = \frac{c^2}{n^2\omega}\mathbf{k}\,, && (3.2.14)\\ \frac{d\mathbf{k}}{dt} &= -\frac{\partial \omega}{\partial \mathbf{r}} = \frac{ck}{2n^3}\frac{\partial n^2}{\partial \mathbf{r}} = \frac{\omega}{n}\frac{\partial n}{\partial \mathbf{r}}\,. && (3.2.15) \end{aligned}$$

After a short manipulation, these canonical equations combine into

$$\frac{n^2}{c}\frac{d}{dt}\left(\frac{n^2}{c}\frac{d\mathbf{r}}{dt}\right) = \frac{1}{2}\frac{\partial n^2}{\partial \mathbf{r}}\,. \qquad (3.2.16)$$

In terms of a different variable time increment $cdt = n^2 d\tau$, equation (3.2.16) may also be expressed in the form of

$$\frac{d^2\mathbf{r}}{d\tau^2} = \frac{1}{2}\frac{\partial n^2}{\partial \mathbf{r}} \qquad \textit{(Newton's 2nd Law)} \qquad (3.2.17)$$

If instead of τ we define the variable time increment $cdt = nd\sigma$, then equation (3.2.16) takes the form of the ***eikonal equation*** (1.1.7) for the paths of light rays in geometric optics, $\mathbf{r}(\sigma) \in \mathbb{R}^3$ as

$$\frac{d}{d\sigma}\left(n(\mathbf{r})\frac{d\mathbf{r}}{d\sigma}\right) = \frac{\partial n}{\partial \mathbf{r}} \qquad \textit{(Eikonal equation)} \qquad (3.2.18)$$

As discussed in Chapter 1, this equation follows from Fermat's principle of stationarity of the optical length under variations of the ray paths,

$$\delta \int_A^B n(\mathbf{r}(\sigma))\, d\sigma = 0 \qquad \textit{(Fermat's principle)} \qquad (3.2.19)$$

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 (3.2.14) yields

$$n(\mathbf{r})\frac{d\mathbf{r}}{d\sigma} = \nabla S(\mathbf{r}) \qquad \textit{(Huygens equation)} \tag{3.2.20}$$

from which the eikonal equation (3.2.18) 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 (3.2.11) includes and unifies the ideas that originated with Fermat, Huygens and Newton.

Remark 3.2.24 (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 (3.2.12), over all possible paths between the initial and final points. Earlier Dirac [Dirac1933] 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, Dirac1981].

3.3 Differential forms and Lie derivatives

3.3.1 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 3.3.1 (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 3.3.2 (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 3.3.3 (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 3.3.4 *[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 3.3.5 (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 3.3.6 (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 3.3.7 *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 3.3.8 (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 3.3.9 *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 3.3.10 (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 3.3.11 (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 3.3.12 (Dual basis)
As in Definition 3.3.11, 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 3.3.13 (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 3.3.14 (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^*_xM dual to T_xM. The dual space T^*_xM is called the ***cotangent space*** *of M at x.*

Definition 3.3.15 (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.*

3.3.2 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:

$$(\phi_* X)\big(\phi(z)\big) = T_z\phi \cdot X(z)\,. \tag{3.3.1}$$

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

$$(\phi_* X)^l = \frac{\partial \phi^l(z)}{\partial z^J} X^J(z) =: (T_z\phi \cdot X(z))^l\,. \tag{3.3.2}$$

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:

$$\phi^* df = d(f \circ \phi) = d(\phi^* f)\,. \tag{3.3.3}$$

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

$$(T_z\phi \cdot dz)^l = \frac{\partial \phi^l(z)}{\partial z^J}\, dz^J\,. \tag{3.3.4}$$

3.3.3 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

$$\alpha_m = \alpha_{i_1 \ldots i_k}(m) dx^{i_1} \wedge \cdots \wedge dx^{i_k} \in \Lambda^k\,,\ i_1 < i_2 < \cdots < i_k\,, \tag{3.3.5}$$

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 $k \leq n$ on an n-dimensional manifold are summarised in the following proposition.

Proposition 3.3.16 (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* $i_1 < i_2 < \cdots < i_{k+l}$.

3.3.4 Pull-back & push-forward of differential forms

Smooth invertible maps act on differential forms by the operations of pull-back and push-forward.

Definition 3.3.17 (Diffeomorphism)
A smooth invertible map whose inverse is also smooth is said to be a ***diffeomorphism****.*

Definition 3.3.18 (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*

$$\phi^*\alpha_m = \alpha_{i_1 \dots i_k}(\phi(m))(T_m\phi \cdot dx)^{i_1} \wedge \cdots \wedge (T_m\phi \cdot dx)^{i_k} \,, \quad (3.3.6)$$

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 3.3.19 *In the definition (3.3.6) 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 (3.3.2) and (3.3.4),

$$\begin{aligned}
\phi^*\big(\mathbf{v}(\mathbf{x})\cdot d\mathbf{x}\big) &= \mathbf{v}\big(\phi(\mathbf{x})\big)\cdot d\phi(\mathbf{x}) \\
&= v_{i_1}\big(\phi(\mathbf{x})\big)\left(\frac{\partial \phi^{i_1}(\mathbf{x})}{\partial x^{i_A}}\, dx^{i_A}\right) \\
&= \mathbf{v}\big(\phi(\mathbf{x})\big)\cdot(T_{\mathbf{x}}\phi\cdot d\mathbf{x})\,.
\end{aligned}$$

Pull-backs of other differential forms may be built up from their basis elements, by the following.

Proposition 3.3.20 (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:

$$\phi^*(\alpha\wedge\beta) = \phi^*\alpha\wedge\phi^*\beta\,. \tag{3.3.7}$$

Remark 3.3.21 *The Definition 3.2.12 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 3.1.6 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

$$\phi_t^*(dq\wedge dp) = dq\wedge dp\,.$$

3.3.5 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 3.3.22

Note that Lie derivative commutes with exterior derivative. That is,

$$d(\pounds_X\alpha) = \pounds_X d\alpha\,, \quad \textit{for} \quad \alpha \in \Lambda^k(M) \quad \textit{and} \quad X \in \mathfrak{X}(M)\,.$$

3.3.6 Contraction, or interior product

Definition 3.3.23 (Contraction, or interior product)
Let $\alpha \in \Lambda^k$ be a k-form on a manifold M

$$\alpha = \alpha_{i_1 \dots i_k} dx^{i_1} \wedge \dots \wedge dx^{i_k} \in \Lambda^k \,, \quad \textit{with} \quad i_1 < i_2 < \dots < i_k \,,$$

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

$$X \,\lrcorner\, \alpha = X^j \alpha_{j i_2 \dots i_k} dx^{i_2} \wedge \dots \wedge dx^{i_k} \,. \tag{3.3.8}$$

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 3.3.24 (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

$$\begin{aligned} X \,\lrcorner\, \alpha &= X^1 \alpha_{123} dx^2 \wedge dx^3 + \textit{cyclic permutations,} \\ &= X^j \alpha_{j i_2 i_3} dx^{i_2} \wedge dx^{i_3} \quad \textit{with} \quad i_2 < i_3 \,. \end{aligned}$$

(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 3.3.25 (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 (3.3.8), 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 3.3.26 (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

$$X \,\lrcorner\, (\alpha\wedge\beta) = (X \,\lrcorner\, \alpha)\wedge\beta + (-1)^k\alpha\wedge(X \,\lrcorner\, \beta)\,. \tag{3.3.9}$$

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 3.3.27 (Contraction commutes with pull-back)
That is,

$$\phi^*(X(m) \,\lrcorner\, \alpha) = X(\phi(m)) \,\lrcorner\, \phi^*\alpha\,. \tag{3.3.10}$$

Proof. Direct verification using the relation between pull-back of forms and push-forward of vector fields. ■

Definition 3.3.28 (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.

$$X \,\lrcorner\, \alpha = i_X\alpha = \alpha(X, \underbrace{\cdot\,,\cdot\,,\ldots\,,\cdot}_{k-1\ \text{slots}}) \in \Lambda^{k-1}\,. \tag{3.3.11}$$

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 3.3.29 (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 3.3.30 *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.*

3.3.7 Exterior derivative

Definition 3.3.31 (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 3.3.32 (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*** *,*

$$d(\alpha \wedge \beta) = (d\alpha) \wedge \beta + (-1)^k \alpha \wedge d\beta\,, \tag{3.3.12}$$

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.*

3.3.8 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:

$$
\begin{aligned}
X \lrcorner\, d\mathbf{x} &= \mathbf{X}\,, \\
X \lrcorner\, d\mathbf{S} &= \mathbf{X} \times d\mathbf{x}\,, \\
(\text{or, } \; X \lrcorner\, dS_i &= \epsilon_{ijk} X^j dx^k) \\
Y \lrcorner\, X \lrcorner\, d\mathbf{S} &= \mathbf{X} \times \mathbf{Y}\,, \\
X \lrcorner\, d^{\,3}x &= \mathbf{X} \cdot d\mathbf{S} = X^k dS_k\,, \\
Y \lrcorner\, X \lrcorner\, d^{\,3}x &= \mathbf{X} \times \mathbf{Y} \cdot d\mathbf{x} = \epsilon_{ijk} X^i Y^j dx^k\,, \\
Z \lrcorner\, Y \lrcorner\, X \lrcorner\, d^{\,3}x &= \mathbf{X} \times \mathbf{Y} \cdot \mathbf{Z}\,.
\end{aligned}
$$

★

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 (A.7.1) 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

$$\begin{aligned} d\big((\mathbf{A}\cdot d\mathbf{x})\wedge(\mathbf{B}\cdot d\mathbf{x})\big) &= (\operatorname{curl}\mathbf{A})\cdot d\mathbf{S}\wedge\mathbf{B}\cdot d\mathbf{x} \\ &\quad -\mathbf{A}\cdot d\mathbf{x}\wedge(\operatorname{curl}\mathbf{B})\cdot d\mathbf{S} \\ &= \big(\mathbf{B}\cdot\operatorname{curl}\mathbf{A} - \mathbf{A}\cdot\operatorname{curl}\mathbf{B}\big)d^{\,3}x \\ &= d\big((\mathbf{A}\times\mathbf{B})\cdot d\mathbf{S}\big) \\ &= \operatorname{div}(\mathbf{A}\times\mathbf{B})\, d^{\,3}x\,. \end{aligned}$$

These calculations return the familiar formulas from vector calculus for quantities curl(grad), div(curl), $\operatorname{curl}(f\mathbf{A})$ and $\operatorname{div}(\mathbf{A}\times\mathbf{B})$. ★

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,

$$\int_a^b df \quad = \quad \int_a^b \nabla f \cdot d\mathbf{x} = f(b) - f(a)\,.$$

(2) ***Classical Stokes theorem***, for a compact surface S with boundary ∂S:

$$\int_S (\operatorname{curl}\mathbf{v}) \cdot d\mathbf{S} \quad = \quad \oint_{\partial S} \mathbf{v} \cdot d\mathbf{x}\,.$$

(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:

$$\int_D (\operatorname{div}\mathbf{A})\, d^{\,3}x \quad = \quad \oint_{\partial D} \mathbf{A} \cdot d\mathbf{S}\,.$$

★

These exercises illustrate the following.

Theorem 3.3.33 (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

$$\int_M d\alpha = \oint_{\partial M} \alpha\,.$$

3.4 Lie derivative

Definition 3.4.1 (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*

$$\pounds_X \alpha = \frac{d}{dt}\bigg|_{t=0} (\phi_t^* \alpha)\,. \tag{3.4.1}$$

Remark 3.4.2 *This is the definition we have been using all along in Chapter 1 in defining vector fields by their characteristic equations.*

Definition 3.4.3 (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*

$$\pounds_X \alpha = X \lrcorner\, d\alpha + d(X \lrcorner\, \alpha)\,. \tag{3.4.2}$$

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.

3.4.1 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 3.4.4 *Poincaré's Theorem 3.2.11 for preservation of the symplectic form ω may be rewritten using Lie derivative notation as*

$$
\begin{aligned}
0 = \frac{d}{dt}\phi_t^*\omega\Big|_{t=0} = \pounds_{X_H}\omega \quad &:= \quad d(X_H \,\lrcorner\, \omega) + X_H \,\lrcorner\, d\omega \\
&=: \quad (\mathrm{div} X_H)\,\omega\,. \qquad (3.4.3)
\end{aligned}
$$

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 3.4.5

- *Relation (3.4.3) expresses Hamiltonian dynamics as the symplectic flow in phase space of the divergenceless Hamiltonian vector field X_H.*

- *The Lie derivative operation defined in (3.4.3) 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 3.4.6 (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

$$\frac{d\omega_n}{dt}\bigg|_{t=0} = \pounds_{X_H} \omega_n := \underbrace{d(X_H \,\lrcorner\, \omega_n)}_{d(dH)\,=\,0} + \underbrace{X_H \,\lrcorner\, d\omega_n}_{=\,0} = 0 \,, \qquad (3.4.4)$$

where $\omega_n \equiv dq_n \wedge dp_n$ is closed $(d\omega_n = 0)$ for all n. ■

Remark 3.4.7 *Many of the following exercises may be solved (or checked) by equating the dynamical definition of Lie derivative in equation (3.4.1) with its geometrical definition by Cartan's formula (3.4.2)*

$$\begin{aligned} \pounds_X \alpha &= \frac{d}{dt}\bigg|_{t=0} (\phi_t^* \alpha) \\ &= X \,\lrcorner\, d\alpha + d(X \,\lrcorner\, \alpha) \,, \end{aligned}$$

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 (3.3.3), wedge product as in (3.3.7) and contraction as in (3.3.10). That is, for $m \in M$,*

$$
\begin{aligned}
d(\phi_t^* \alpha) &= \phi_t^* d\alpha \,, \\
\phi_t^*(\alpha \wedge \beta) &= \phi_t^* \alpha \wedge \phi_t^* \beta \,, \\
\phi_t^*(X(m) \,\lrcorner\, \alpha) &= X(\phi_t(m)) \,\lrcorner\, \phi_t^* \alpha \,.
\end{aligned}
$$

3.4.2 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) $\pounds_X \left(\mathbf{v} \cdot d\mathbf{x}\right) = \big(-\ \mathbf{X} \times \operatorname{curl} \mathbf{v} + \nabla(\mathbf{X} \cdot \mathbf{v})\big) \cdot d\mathbf{x}$,

(c) $\pounds_X(\boldsymbol{\omega} \cdot d\mathbf{S}) = \big(-\ \operatorname{curl}(\mathbf{X} \times \boldsymbol{\omega}) + \mathbf{X} \operatorname{div} \boldsymbol{\omega}\big) \cdot d\mathbf{S}$,

(d) $\pounds_X(f\, d^{\,3}x) = (\operatorname{div} f\mathbf{X})\, d^{\,3}x$.

★

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,

$$\pounds_X(Y \,\lrcorner\, d\mathbf{x}) = \pounds_X Y \,\lrcorner\, d\mathbf{x} + Y \,\lrcorner\, \pounds_X d\mathbf{x} \,, \quad (3.4.5)$$

which implies the relation,

$$\pounds_X Y = [X,\, Y]\,, \qquad (3.4.6)$$

where $[X,\, Y]$ is the Jacobi-Lie bracket (1.11.3) 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,

$$\pounds_X(\alpha \wedge \beta) = (\pounds_X \alpha) \wedge \beta + \alpha \wedge \pounds_X \beta\,. \qquad (3.4.7)$$

The product rule for Lie derivative (3.4.7) also follows immediately from its dynamical definition (3.4.1).

(g) Use

$$[X\,,\, Y] \lrcorner \alpha = \pounds_X (Y \lrcorner \alpha) - Y \lrcorner (\pounds_X \alpha)\,, \qquad (3.4.8)$$

as verified in part **(d)** in concert with the definition(s) of Lie derivative to show,

$$\pounds_{[X\,,Y]}\alpha = \pounds_X \pounds_Y \alpha - \pounds_Y \pounds_X \alpha\,. \qquad (3.4.9)$$

(h) Use the result of **(g)** to verify the ***Jacobi identity for the Lie derivative***,

$$\pounds_{[\,Z\,,[X\,,Y]]}\,\alpha + \pounds_{[\,X\,,[Y\,,Z]]}\,\alpha + \pounds_{[\,Y\,,[Z\,,X]]}\,\alpha = 0\,.$$

★

3.5 Formulations of ideal fluid dynamics

3.5.1 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

$$\underbrace{\partial_t\,\mathbf{u} + \mathbf{u}\cdot\nabla\mathbf{u}}_{\text{Acceleration}} = \underbrace{\mathbf{u}\times 2\boldsymbol{\Omega}}_{\text{Coriolis}} - \underbrace{\nabla p}_{\text{Pressure}} \; . \tag{3.5.1}$$

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,

$$\begin{aligned} -\Delta p &= \mathrm{div}\big(\mathbf{u}\cdot\nabla\mathbf{u} - \mathbf{u}\times 2\boldsymbol{\Omega}\big)\,, \\ &= u_{i,j}u_{j,i} - \mathrm{div}\,\big(\mathbf{u}\times 2\boldsymbol{\Omega}\big)\,, \\ &= \mathrm{tr}\,\mathsf{S}^2 - \frac{1}{2}|\mathrm{curl}\mathbf{u}|^2 - \mathrm{div}\,\big(\mathbf{u}\times 2\boldsymbol{\Omega}\big)\,, \end{aligned} \tag{3.5.2}$$

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 (3.5.1) may be rearranged into an alternative form,

$$\partial_t\,\mathbf{v} - \mathbf{u}\times\boldsymbol{\omega} + \nabla\Big(p + \frac{1}{2}|\mathbf{u}|^2\Big) = 0\,, \tag{3.5.3}$$

where we denote

$$\mathbf{v} \equiv \mathbf{u} + \mathbf{R}\,, \quad \boldsymbol{\omega} = \mathrm{curl}\mathbf{v} = \mathrm{curl}\mathbf{u} + 2\boldsymbol{\Omega}\,, \tag{3.5.4}$$

and introduce the ***Lamb vector***,

$$\boldsymbol{\ell} := -\,\mathbf{u}\times\boldsymbol{\omega}\,, \tag{3.5.5}$$

which represents the nonlinearity in Euler's fluid equation (3.5.3). The Poisson equation (3.5.2) for pressure p may now be expressed in terms of the divergence of the Lamb vector,

$$-\,\Delta\big(p + \tfrac{1}{2}|\mathbf{u}|^2\big) = \mathrm{div}(-\,\mathbf{u}\times\mathrm{curl}\,\mathbf{v}) = \mathrm{div}\,\boldsymbol{\ell}\,. \tag{3.5.6}$$

Remark 3.5.1 (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*

$$\partial_n\big(p + \tfrac{1}{2}|\mathbf{u}|^2\big) + \hat{\mathbf{n}} \cdot \boldsymbol{\ell} = 0\,, \tag{3.5.7}$$

at a fixed boundary with unit outward normal vector $\hat{\mathbf{n}}$.

Remark 3.5.2 (Helmholtz vorticity dynamics)

Taking the curl of the Euler fluid equation (3.5.3) yields the ***Helmholtz vorticity equation***

$$\partial_t\, \boldsymbol{\omega} - \operatorname{curl}\left(\mathbf{u} \times \boldsymbol{\omega}\right) = 0\,, \tag{3.5.8}$$

whose geometrical meaning will emerge in discussing Stokes Theorem 3.5.5 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 3.5.3 (Kelvin's circulation theorem)
The Euler equations (3.5.1) preserve the circulation integral $I(t)$ *defined by*

$$I(t) = \oint_{c(\mathbf{u})} \mathbf{v} \cdot d\mathbf{x}\,, \tag{3.5.9}$$

where $c(\mathbf{u})$ *is a closed circuit moving with the fluid at velocity* $\mathbf{u}$.

Proof. The dynamical definition of Lie derivative (3.4.1) yields the following for the time rate of change of this circulation integral,

$$\begin{aligned}
\frac{d}{dt}\oint_{c(\mathbf{u})}\mathbf{v}\cdot d\mathbf{x} &= \oint_{c(\mathbf{u})}\Big(\frac{\partial}{\partial t}+\pounds_{\mathbf{u}}\Big)(\mathbf{v}\cdot d\mathbf{x})\\
&= \oint_{c(\mathbf{u})}\Big(\frac{\partial \mathbf{v}}{\partial t}+\frac{\partial \mathbf{v}}{\partial x^j}u^j+v_j\frac{\partial u^j}{\partial \mathbf{x}}\Big)\cdot d\mathbf{x}\\
&= -\oint_{c(\mathbf{u})}\nabla\Big(p+\frac{1}{2}|\mathbf{u}|^2-\mathbf{u}\cdot\mathbf{v}\Big)\cdot d\mathbf{x}\\
&= -\oint_{c(\mathbf{u})}d\Big(p+\frac{1}{2}|\mathbf{u}|^2-\mathbf{u}\cdot\mathbf{v}\Big)=0\,.
\end{aligned}\tag{3.5.10}$$

The Cartan formula (3.4.2) 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,

$$\begin{aligned}
\pounds_{\mathbf{u}}(\mathbf{v}\cdot d\mathbf{x}) &= \big(\mathbf{u}\cdot\nabla\mathbf{v}+v_j\nabla u^j\big)\cdot d\mathbf{x}\\
&= u \,\lrcorner\, d(\mathbf{v}\cdot d\mathbf{x})+d(u\,\lrcorner\,\mathbf{v}\cdot d\mathbf{x})\\
&= u\,\lrcorner\, d(\mathrm{curl}\,\mathbf{v}\cdot d\mathbf{S})+d(\mathbf{u}\cdot\mathbf{v})\\
&= \big(-\ \mathbf{u}\times\mathrm{curl}\,\mathbf{v}+\nabla(\mathbf{u}\cdot\mathbf{v})\big)\cdot d\mathbf{x}\,.
\end{aligned}\tag{3.5.11}$$

This identity recasts Euler's equation into the following geometric form,

$$\begin{aligned}
\Big(\frac{\partial}{\partial t}+\pounds_{\mathbf{u}}\Big)(\mathbf{v}\cdot d\mathbf{x}) &= \big(\partial_t\mathbf{v}-\ \mathbf{u}\times\mathrm{curl}\,\mathbf{v}+\nabla(\mathbf{u}\cdot\mathbf{v})\big)\cdot d\mathbf{x}\\
&= -\nabla\Big(p+\frac{1}{2}|\mathbf{u}|^2-\mathbf{u}\cdot\mathbf{v}\Big)\cdot d\mathbf{x}\\
&= -\,d\Big(p+\frac{1}{2}|\mathbf{u}|^2-\mathbf{u}\cdot\mathbf{v}\Big).
\end{aligned}\tag{3.5.12}$$

This finishes the last step in the proof (3.5.10), because the integral of an exact differential around a closed loop vanishes. ∎

The exterior derivative of the Euler fluid equation in the form (3.5.12) yields Stokes theorem, after using the commutativity of the exterior and Lie derivatives $[d,\ \pounds_{\mathbf{u}}]=0$,

$$
\begin{aligned}
d\pounds_{\mathbf{u}}(\mathbf{v}\cdot d\mathbf{x}) &= \pounds_{\mathbf{u}}\, d(\mathbf{v}\cdot d\mathbf{x}) \\
&= \pounds_{\mathbf{u}}(\mathrm{curl}\mathbf{v}\cdot d\mathbf{S}) \\
&= -\,\mathrm{curl}\,\big(\mathbf{u}\times\mathrm{curl}\mathbf{v}\big)\cdot d\mathbf{S} \\
&= \big[\mathbf{u}\cdot\nabla\mathrm{curl}\mathbf{v}+\mathrm{curl}\mathbf{v}(\mathrm{div}\mathbf{u})-(\mathrm{curl}\mathbf{v})\cdot\nabla\mathbf{u}\big]\cdot d\mathbf{S} \\
(\text{by } \mathrm{div}\mathbf{u}=0) &= \big[\mathbf{u}\cdot\nabla\mathrm{curl}\mathbf{v}-(\mathrm{curl}\mathbf{v})\cdot\nabla\mathbf{u}\big]\cdot d\mathbf{S} \\
&=: \ [\,u,\ \mathrm{curl}v\,]\cdot d\mathbf{S}\,,
\end{aligned}
\tag{3.5.13}
$$

where $[\,u,\ \mathrm{curl}v\,]$ denotes the ***Jacobi-Lie bracket*** (1.11.3) of the vector fields u and $\mathrm{curl}v$. This calculation proves the following.

Theorem 3.5.4 *Euler's fluid equations (3.5.3) imply that*

$$
\frac{\partial\omega}{\partial t}=-\,[\,u,\ \omega\,]
\tag{3.5.14}
$$

where $[\,u,\ \omega\,]$ denotes the Jacobi-Lie bracket (1.11.3) of the divergenceless vector fields u and $\omega:=\mathrm{curl}v$.

The exterior derivative of Euler's equation in its geometric form (3.5.12) is equivalent to the curl of its vector form (3.5.3). That is,

$$
d\Big(\frac{\partial}{\partial t}+\pounds_{\mathbf{u}}\Big)(\mathbf{v}\cdot d\mathbf{x})=\Big(\frac{\partial}{\partial t}+\pounds_{\mathbf{u}}\Big)(\mathrm{curl}\mathbf{v}\cdot d\mathbf{S})=0\,.
\tag{3.5.15}
$$

Hence from the calculation in (3.5.13) and the Helmholtz vorticity equation (3.5.15) we have

$$
\Big(\frac{\partial}{\partial t}+\pounds_{\mathbf{u}}\Big)(\mathrm{curl}\mathbf{v}\cdot d\mathbf{S})=\Big(\partial_t\,\boldsymbol{\omega}-\mathrm{curl}\,(\mathbf{u}\times\boldsymbol{\omega})\Big)\cdot d\mathbf{S}=0\,,
\tag{3.5.16}
$$

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 3.5.5 (Stokes theorem for vorticity of a rotating fluid)

$$
\begin{aligned}
\frac{d}{dt}\iint_{S(\mathbf{u})}\mathrm{curl}\mathbf{v}\cdot d\mathbf{S} &= \iint_{S(\mathbf{u})}\Big(\frac{\partial}{\partial t}+\pounds_{\mathbf{u}}\Big)(\mathrm{curl}\mathbf{v}\cdot d\mathbf{S}) \qquad (3.5.17)\\
&= \iint_{S(\mathbf{u})}\Big(\partial_t\,\boldsymbol{\omega}-\mathrm{curl}\,(\mathbf{u}\times\boldsymbol{\omega})\Big)\cdot d\mathbf{S}=0\,,
\end{aligned}
$$

where the surface $S(\mathbf{u})$ is bounded by an arbitrary circuit $\partial S = c(\mathbf{u})$ moving with the fluid.

3.5.2 Steady solutions: Lamb surfaces

According to Theorem 3.5.4, Euler's fluid equations (3.5.3) imply that

$$\frac{\partial \omega}{\partial t} = -\,[\,u,\,\omega\,]\,. \tag{3.5.18}$$

Consequently, the vector fields $u,\ \omega$ 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 (3.5.5) satisfies

$$\boldsymbol{\ell} := -\,\mathbf{u} \times \mathrm{curl}\mathbf{v} = \nabla H(\mathbf{x})\,, \tag{3.5.19}$$

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*

$$\pounds_u H = \mathbf{u} \cdot \nabla H = 0 \quad \text{and} \quad \pounds_{\mathrm{curl}v} H = \mathrm{curl}\mathbf{v} \cdot \nabla H = 0\,. \tag{3.5.20}$$

The Lamb surface condition (3.5.19) 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 3.5.6 (Lamb surface condition [HaMe1998])
The Lamb surface condition (3.5.19) is equivalent to the following double substitution of vector fields into the volume form,

$$dH = u \,\lrcorner\, \mathrm{curl} v \,\lrcorner\, d^3x \,. \tag{3.5.21}$$

Proof. Recall that the contraction of vector fields with forms yields the following useful formula for the surface element:

$$\nabla \,\lrcorner\, d^3x = d\mathbf{S} \,. \tag{3.5.22}$$

Then using results from previous exercises in vector calculus operations one finds by direct computation that

$$\begin{aligned} u \,\lrcorner\, \mathrm{curl} v \,\lrcorner\, d^3x &= u \,\lrcorner\, (\mathrm{curl}\mathbf{v} \cdot d\mathbf{S}) \\ &= -\big(\mathbf{u} \times \mathrm{curl}\mathbf{v}\big) \cdot d\mathbf{x} \\ &= \nabla H \cdot d\mathbf{x} \\ &= dH \,. \end{aligned} \tag{3.5.23}$$

■

Remark 3.5.7 *Formula (3.5.23)*

$$u \,\lrcorner\, (\mathrm{curl}\mathbf{v} \cdot d\mathbf{S}) = dH$$

is to be compared with

$$X_h \,\lrcorner\, \omega = dH \,,$$

in the definition of a Hamiltonian vector field in equation (3.2.4) of Theorem 3.2.6. Likewise, the stationary case of the Helmholtz vorticity equation (3.5.15), namely,

$$\pounds_{\mathbf{u}}(\mathrm{curl}\mathbf{v} \cdot d\mathbf{S}) = 0 \,. \tag{3.5.24}$$

is to be compared with the proof of Poincaré's theorem in Proposition 3.4.4

$$\pounds_{X_h}\omega = d(X_h \,\lrcorner\, \omega) = d^2 H = 0 \,.$$

Thus, the 2-form $\mathrm{curl}\mathbf{v} \cdot d\mathbf{S}$ *plays the same role for stationary Euler fluid flows as the symplectic form* $dq \wedge dp$ *plays for canonical Hamiltonian flows.*

Definition 3.5.8 *The* ***Clebsch representation*** *of the 1-form* $\mathbf{v} \cdot d\mathbf{x}$ *is defined by*

$$\mathbf{v} \cdot d\mathbf{x} = -\,\Pi\, d\Xi + d\Psi\,. \tag{3.5.25}$$

The functions Ξ*,* Π *and* Ψ *are called* ***Clebsch potentials*** *for the vector* $\mathbf{v}$*.*[2]

In terms of the Clebsch representation (3.5.25) 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,

$$\mathrm{curl}\mathbf{v} \cdot d\mathbf{S} = d\Xi \wedge d\Pi\,. \tag{3.5.26}$$

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,

$$\mathrm{curl}\mathbf{v} = \nabla\Xi \times \nabla\Pi\,, \tag{3.5.27}$$

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 3.5.9 (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 (3.5.26) and they may be envisioned along with the flow lines of the vector field $\mathrm{curl}v$ in $\mathbb{R}^3$ by using formula (3.5.27). The main result is the following.

[2] The Clebsch representation is another example of a momentum map. For more discussion of this aspect of fluid flows, see [MaWe83, HoMa2004].

Theorem 3.5.10 (Lamb surfaces are symplectic manifolds)
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 (3.5.23) to coordinates on a total vorticity flux surface (3.5.26) yields the exterior derivative of the Hamiltonian,

$$\begin{aligned}
dH(\Xi,\,\Pi) &= u\,\lrcorner\,(\mathrm{curl}\mathbf{v}\cdot d\mathbf{S})\\
&= u\,\lrcorner\,(d\Xi\wedge d\Pi)\\
&= (\mathbf{u}\cdot\nabla\Xi)\,d\Pi-(\mathbf{u}\cdot\nabla\Pi)\,d\Xi\\
&=: \frac{d\Xi}{dT}\,d\Pi-\frac{d\Pi}{dT}\,d\Xi\\
&= \frac{\partial H}{\partial\Pi}\,d\Pi+\frac{\partial H}{\partial\Xi}\,d\Xi\,,
\end{aligned}\tag{3.5.28}$$

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,

$$(\mathbf{u}\cdot\nabla\Xi)=\pounds_u\Xi \;\;=:\;\; \frac{d\Xi}{dT}=\frac{\partial H}{\partial\Pi}=\big\{\Xi,\,H\big\}\,,\tag{3.5.29}$$

$$(\mathbf{u}\cdot\nabla\Pi)=\pounds_u\Pi \;\;=:\;\; \frac{d\Pi}{dT}=-\,\frac{\partial H}{\partial\Xi}=\big\{\Pi,\,H\big\}\,,\tag{3.5.30}$$

where $\{\,\cdot\,,\,\cdot\,\}$ is the canonical Poisson bracket for the symplectic form $d\Xi\wedge d\Pi$. ∎

Corollary 3.5.11 *The vorticity flux* $d\Xi\wedge d\Pi$ *is invariant under the flow of the velocity vector field* u.

Proof. By (3.5.28), 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 3.4.4 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, $\mathrm{curl}\mathbf{v} \cdot d\mathbf{S} = d\Xi \wedge d\Pi$. ■

3.5.3 Helicity in incompressible fluids

Definition 3.5.12 (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*

$$\Lambda[\mathrm{curl}\mathbf{v}] = \int_D \mathbf{v} \cdot \mathrm{curl}\mathbf{v}\, d^{\,3}x\,, \qquad (3.5.31)$$

where $\mathbf{v}$ *is a divergence-free vector-potential for the field* $\mathrm{curl}\mathbf{v}$.

Remark 3.5.13 *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 3.5.14 (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*

$$\Lambda[\mathrm{curl}\mathbf{v}] = \int_D \mathbf{v} \cdot \mathrm{curl}\mathbf{v}\, d^{\,3}x\,, \tag{3.5.32}$$

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 (3.5.12) for rotating incompressible flow with unit mass density in terms of the circulation 1-form $v := \mathbf{v} \cdot d\mathbf{x}$ as

$$(\partial_t + \pounds_u)v = -\,d\left(p + \frac{1}{2}|\mathbf{u}|^2 - \mathbf{u}\cdot\mathbf{v}\right) =: -\,d\,\varpi\,, \tag{3.5.33}$$

and $\pounds_u\, d^{\,3}x = 0$, where ϖ is an augmented pressure variable,

$$\varpi := p + \frac{1}{2}|\mathbf{u}|^2 - \mathbf{u}\cdot\mathbf{v}\,. \tag{3.5.34}$$

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

$$v \wedge dv = \mathbf{v}\cdot\mathrm{curl}\mathbf{v}\, d^{\,3}x = \lambda\, d^{\,3}x\,, \quad \text{with} \quad \lambda = \mathbf{v}\cdot\mathrm{curl}\mathbf{v}\,, \tag{3.5.35}$$

obeys the dynamics it inherits from the Euler equations,

$$\big(\partial_t + \pounds_u\big)(v\wedge dv) = -d\varpi\wedge dv - v\wedge d^2\varpi = -d(\varpi\, dv)\,, \tag{3.5.36}$$

after using $d^2\varpi = 0$ and $d^2v = 0$. In vector form, this result may be expressed as a conservation law,

$$\big(\partial_t\lambda + \mathrm{div}\lambda\mathbf{u}\big)d^{\,3}x = -\,\mathrm{div}(\varpi\,\mathrm{curl}\mathbf{v})\, d^{\,3}x\,. \tag{3.5.37}$$

Consequently, the time derivative of the integrated helicity in a domain D obeys

$$\begin{aligned}\frac{d}{dt}\Lambda[\mathrm{curl}\mathbf{v}] &= \int_D \partial_t\lambda\, d^{\,3}x = -\int_D \mathrm{div}(\lambda\mathbf{u} + \varpi\,\mathrm{curl}\mathbf{v})\, d^{\,3}x \\ &= -\oint_{\partial D}(\lambda\mathbf{u} + \varpi\,\mathrm{curl}\mathbf{v})\,\cdot\, d\mathbf{S}\,,\end{aligned} \tag{3.5.38}$$

which vanishes when homogeneous or periodic boundary conditions are imposed on ∂D. ∎

Remark 3.5.15 *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 3.5.16 *A flux of total vorticity* curl$\mathbf{v}$ *into the domain is a source of helicity.*

> **Exercise.** Use Cartan's formula (3.4.2) to compute $\pounds_u(v\wedge dv)$ in equation (3.5.36). ★

> **Exercise.** Compute the helicity for the 1-form $v = \mathbf{v}\cdot d\mathbf{x}$ in the Clebsch representation (3.5.25). What does this mean for the linkage of the vortex lines that admit the Clebsch representation? ★

Remark 3.5.17 (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 3.5.18 (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 3.5.19 (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 3.5.20 (Beltrami flows)
Equilibrium Euler fluid flows whose velocity and total vorticity are collinear are called ***Beltrami flows****.*

Theorem 3.5.21 (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 3.5.22 (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.

3.5.4 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 $\mathrm{curl}\mathbf{R} = 2\boldsymbol{\Omega}$ are given by

$$\underbrace{\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla\mathbf{u}}_{\text{Acceleration}} = \underbrace{-\,gb\nabla z}_{\text{Buoyancy}} + \underbrace{\mathbf{u} \times 2\boldsymbol{\Omega}}_{\text{Coriolis}} - \underbrace{\nabla p}_{\text{Pressure}} \tag{3.5.39}$$

where $-g\nabla z$ is the constant downward acceleration of gravity and b is the bouyancy, which satisfies the ***advection relation***,

$$\partial_t b + \mathbf{u} \cdot \nabla b = 0\,. \tag{3.5.40}$$

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,

$$-\,\Delta\Big(p + \frac{1}{2}|\mathbf{u}|^2\Big) = \mathrm{div}(-\,\mathbf{u} \times \mathrm{curl}\,\mathbf{v}) + g\partial_z b\,, \tag{3.5.41}$$

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 (3.5.39) may be rearranged as

$$\partial_t \mathbf{v} -\ \mathbf{u} \times \mathrm{curl}\,\mathbf{v} + gb\nabla z + \nabla\Big(p + \frac{1}{2}|\mathbf{u}|^2\Big) = 0\,, \tag{3.5.42}$$

where $\mathbf{v} \equiv \mathbf{u} + \mathbf{R}$ and $\nabla \cdot \mathbf{u} = 0$. Geometrically, this is

$$\big(\partial_t + \pounds_u\big)v + gbdz + d\varpi = 0\,, \tag{3.5.43}$$

where ϖ is defined in (3.5.33). In addition, the buoyancy satisfies

$$\big(\partial_t + \pounds_u\big)b = 0\,, \quad \text{with} \quad \pounds_u\, d^3x = 0\,. \tag{3.5.44}$$

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 (3.5.43) are written as

$$\big(\partial_t + \pounds_u\big)dv = -gdb \wedge dz \quad \text{and} \quad \big(\partial_t + \pounds_u\big)db = 0\,. \tag{3.5.45}$$

Consequently, one finds from the product rule for Lie derivatives (3.4.7) that

$$\big(\partial_t + \pounds_u\big)(dv \wedge db) = 0 \quad \text{or} \quad \partial_t\, q + \mathbf{u}\cdot\nabla q = 0\,, \tag{3.5.46}$$

in which the quantity

$$q = \nabla b \cdot \mathrm{curl}\mathbf{v} \tag{3.5.47}$$

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 3.5.23 (Silberstein-Ertel theorem)
The constancy of the scalar quantities b and q on fluid parcels implies conservation of the spatially integrated quantity,

$$C_\Phi = \int_D \Phi(b,q)\, d^{\,3}x\,, \tag{3.5.48}$$

for any smooth function Φ for which the integral exists.

Remark 3.5.24 (Energy conservation)
*In addition to C_Φ, the Euler-Boussinesq fluid equations (3.5.42) also conserve the **total energy***

$$E = \int_D \frac{1}{2}|\mathbf{u}|^2 + gbz\ d^{\,3}x\,, \tag{3.5.49}$$

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 3.5.25 (Energy-Casimir criteria for equilibria)
Critical points of the conserved sum $E_\Phi = E + C_\Phi$, namely,

$$E_\Phi = \int_D \frac{1}{2}|\mathbf{u}|^2 + gbz\ d^{\,3}x + \int_D \Phi(b,q) + \varkappa q\, d^{\,3}x\,, \tag{3.5.50}$$

are equilibrium solutions of the Euler-Boussinesq fluid equations in (3.5.42). The function Φ in the Casimir and the Bernoulli function K in (3.5.54) for the corresponding fluid equilibrium are related by $q\Phi_q - \Phi = K$.

Proof. The last term in (3.5.50) 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

$$\begin{aligned} \delta E_\Phi &= \int_D \Big(\mathbf{u}_e - \Phi_{qq}\nabla b_e \times \nabla q_e\Big)\cdot \delta\mathbf{u}\, d^3x \\ &\quad + \int_D \Big(gz + \Phi_b - \mathrm{curl}\mathbf{v}_e \cdot \nabla\Phi_q\Big)\,\delta b\, d^3x \\ &\quad + \Big(\Phi_q\Big|_{\partial D} + \varkappa\Big)\oint_{\partial D}\Big(\delta b\,\mathrm{curl}\mathbf{v}_e + b_e\,\mathrm{curl}\,\delta\mathbf{u}\Big)\cdot\hat{\mathbf{n}}\, dS\,, \end{aligned} \tag{3.5.51}$$

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,

$$\begin{aligned} \delta\mathbf{u}: &\qquad \mathbf{u}_e = \Phi_{qq}\nabla b_e \times \nabla q_e\,, \\ \delta b: &\qquad gz + \Phi_b = \mathrm{curl}\mathbf{v}_e\cdot\nabla\Phi_q\,, \\ \partial D: &\qquad \Phi_q\Big|_{\partial D} + \varkappa = 0\,. \end{aligned} \tag{3.5.52}$$

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 (3.5.42)

$$\mathbf{u}_e \times \mathrm{curl}\,\mathbf{v}_e = -\,gz\nabla b_e + \nabla K\,, \tag{3.5.53}$$

which summons the ***Bernoulli function***,

$$K(b_e, q_e) = p_e + \frac{1}{2}|\mathbf{u}_e|^2 + g b_e z\,, \tag{3.5.54}$$

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 (3.5.53) yields

$$\mathbf{u}_e = \frac{1}{q_e}\nabla b_e \times \nabla K(b_e, q_e) = \nabla b_e \times \nabla \Phi_q(b_e, q_e)\,, \tag{3.5.55}$$

where the last relation uses the critical point condition arising from the variations of velocity, $\delta\mathbf{u}$. By equation (3.5.55), critical points of E_Φ are steady solutions of the Euler-Boussinesq fluid equations (3.5.42) and the function Φ in the Casimir is related to the Bernoulli function K in (3.5.54) for the corresponding steady solution by

$$q_e\Phi_{qq}(b_e, q_e) = K_q(b_e, q_e)\,. \tag{3.5.56}$$

This equation integrates to

$$q_e\Phi_q - \Phi = K + F(b_e)\,. \tag{3.5.57}$$

The vector product of ∇q_e with the steady flow relation (3.5.53) yields

$$\Phi_{qq}(b_e, q_e)(\nabla q_e \cdot \mathrm{curl}\mathbf{v}_e) = gz - K_b(b_e, q_e)\,. \tag{3.5.58}$$

Combining this result with the critical point condition for δb in (3.5.52) yields

$$q_e\Phi_q - \Phi = K + G(q_e)\,. \tag{3.5.59}$$

Subtracting the two relations (3.5.59) and (3.5.56) eliminates the integration functions F and G, and establishes

$$q_e\Phi_q(b_e, q_e) - \Phi(b_e, q_e) = K(b_e, q_e)\,, \tag{3.5.60}$$

as the relation between critical points of E_Φ and equilibrium solutions of the Euler-Boussinesq fluid equations. ∎

Remark 3.5.26
The energy-Casimir stability method was implemented for Euler-Boussinesq fluid equilibria in [AbHoMaRa1986]. See also [HoMaRaWe1985] for additional examples.

3.6 Hodge star operator on $\mathbb{R}^3$

Definition 3.6.1 *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 3.6.2 *Note that $**\,\alpha = \alpha$ for these k-forms.*

Definition 3.6.3 (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*

$$(\alpha,\, \beta) := \int_M \alpha \wedge *\beta = \int_M \langle \alpha, \beta \rangle \; d^3x \,, \tag{3.6.1}$$

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,

$$\begin{aligned}
(f,\, g) = \int_M f \wedge *g \quad &:= \quad \int_M fg\, d^3x \,, \\
(\mathbf{u} \cdot d\mathbf{x},\, \mathbf{v} \cdot d\mathbf{x}) = \int_M \mathbf{u} \cdot d\mathbf{x} \wedge *(\mathbf{v} \cdot d\mathbf{x}) \quad &:= \quad \int_M \mathbf{u} \cdot \mathbf{v}\, d^3x \,.
\end{aligned}$$

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 3.6.4 *The **codifferential**, denoted as δ, is defined for a k-form $\alpha\in\Lambda^k$ as*

$$\delta\alpha = (-1)^{k+1+k(3-k)} * \, d*\alpha\,. \tag{3.6.2}$$

Note that the sign is positive for $k=1$ and negative for $k=2$.

Exercise. Verify that $\delta^2=0$. ★

Remark 3.6.5 *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 3.6.6 *The codifferential is the adjoint of the exterior derivative, in that*

$$(\delta\alpha,\,\beta) = (\alpha,\,d\beta)\,. \tag{3.6.3}$$

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 3.6.7 *The **Laplace-Beltrami operator** on smooth functions is defined to be $\nabla^2=\operatorname{div}\operatorname{grad}=\delta\,d$. Thus, one finds,*

$$\nabla^2 f = \delta\,df = *\,d*df\,, \tag{3.6.4}$$

for a smooth function f.

Definition 3.6.8 *The **Laplace-deRham operator** is defined by*

$$\Delta := d\delta + \delta d\,. \tag{3.6.5}$$

Exercise. Show that the ***Laplace-deRham operator*** on a 1-form $\mathbf{v}\cdot d\mathbf{x}$ expresses the ***Laplacian of a vector,***

$$(d\delta+\delta d)(\mathbf{v}\cdot d\mathbf{x}) = (\nabla \mathrm{div}\,\mathbf{v} - \mathrm{curl}\,\mathrm{curl}\,\mathbf{v})\cdot d\mathbf{x} =: (\Delta \mathbf{v})\cdot d\mathbf{x}\,.$$

Use this expression to define the inverse of the curl operator applied to a divergenceless vector function as

$$\mathrm{curl}^{-1}\mathbf{v} = \mathrm{curl}(-\Delta^{-1}\mathbf{v}) \quad \text{when} \quad \mathrm{div}\,\mathbf{v} = 0. \tag{3.6.6}$$

This is the ***Biot-Savart Law*** often used in electo-magnetism and incompressible fluid dynamics. ★

Remark 3.6.9 *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

$$\begin{aligned}\int -\Delta\alpha\cdot\beta d^3x &= \int (-\nabla \mathrm{div}\alpha + \mathrm{curlcurl}\alpha)\cdot\beta d^3x \\ &= \int (\mathrm{div}\alpha\cdot\mathrm{div}\beta + \mathrm{curl}\alpha\cdot\mathrm{curl}\beta d^3x\,.\end{aligned}$$

Conclude that the Laplace-deRham operator $-\Delta$ is non-negative, by setting $\alpha = \beta$.

★

Exercise. Use formula (3.6.2) 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 (3.5.41)? ★

3.7 Poincaré's Lemma: Closed vs exact differential forms

Definition 3.7.1 (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 3.7.2 (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 3.7.3 *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 (3.6.1) 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 3.6.6. ■

Remark 3.7.4 *Not all closed forms are globally exact on a given manifold M.*

Example 3.7.5 (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 3.7.6 (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 3.7.7 (Poincaré's lemma)
Any closed form on a manifold M *is locally exact.*

Remark 3.7.8 *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 3.7.9 *In the example of* ***helicity*** *above, the one-form* $v = f dg + \psi d\phi$ *may always be written locally as* $v = f dg + c d\phi$ *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 3.7.10 *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 3.7.11 (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,*

$$\begin{aligned}
\beta &= X \mathbin{\lrcorner} d^{\,3}x = \mathbf{x} \cdot \hat{\mathbf{n}}\, dS \\
&= xdy \wedge dz + ydz \wedge dx + zdx \wedge dy \\
&= \frac{1}{2}\epsilon_{abc} x^a\, dx^b \wedge dx^c \\
&= r\partial_r \mathbin{\lrcorner} r^2 dr \wedge d\phi \wedge d\cos\theta = r^3 d\phi \wedge d\cos\theta\,.
\end{aligned}$$

One computes the exterior derivative

$$\begin{aligned}
d\beta &= d(X \mathbin{\lrcorner} d^{\,3}x) = d(\mathbf{x} \cdot \hat{\mathbf{n}}\, dS) = \operatorname{div} \mathbf{x}\, d^{\,3}x \\
&= 3\, d^{\,3}x = 3\, r^2 dr \wedge d\phi \wedge d\cos\theta \neq 0\,.
\end{aligned}$$

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 3.7.12

- *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 3.7.13 *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,*

$$\begin{aligned} \beta = X \mathbin{\lrcorner} d^3n := \mathbf{n} \cdot \frac{\partial}{\partial \mathbf{n}} \mathbin{\lrcorner} d^3n &= \frac{1}{2}\epsilon_{abc}\, n^a\, dn^b \wedge dn^c \\ &= \frac{1}{2}\epsilon_{abc}\, n^a \nabla n^b \times \nabla n^c \cdot d\mathbf{S}(\mathbf{x})\,. \end{aligned}$$

One computes the exterior derivative once again,

$$d\beta = d(X \mathbin{\lrcorner} d^3x) = \operatorname{div} \mathbf{n}\, d^3n = \det\,[\nabla \mathbf{n}]\, d^3x\,.$$

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,

$$\beta = d\phi \wedge d\cos\theta = d(\phi\, d\cos\theta)\,. \tag{3.7.1}$$

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 3.7.14 (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]. The Hopf fibration will re-emerge naturally and play an important role as the* ***orbit manifold*** *in our studies of the dynamics of* ***coupled resonant oscillators*** *in Chapter 4.*

3.8 Euler's equations in Maxwell form

Exercise. (Maxwell form of Euler's fluid equations)
Show that by making the following identifications

$$\begin{aligned}
\mathbf{B} &:= \boldsymbol{\omega} + \operatorname{curl}\mathbf{A}_0 \\
\mathbf{E} &:= \boldsymbol{\ell} + \nabla\big(p + \tfrac{1}{2}|\mathbf{u}|^2\big) + \big(\nabla\phi - \partial_t\mathbf{A}_0\big) \\
\mathbf{D} &:= \boldsymbol{\ell} \\
\mathbf{H} &:= \nabla\psi\,,
\end{aligned} \tag{3.8.1}$$

the Euler fluid equations (3.5.3) and (3.5.6) imply the

Maxwell form

$$
\begin{aligned}
\partial_t \mathbf{B} &= -\operatorname{curl}\mathbf{E} \\
\partial_t \mathbf{D} &= \operatorname{curl}\mathbf{H} + \mathbf{J} \\
\operatorname{div}\mathbf{B} &= 0 \\
\operatorname{div}\mathbf{E} &= 0 \\
\operatorname{div}\mathbf{D} &= \rho \;=\; -\Delta\big(p + \tfrac{1}{2}|\mathbf{u}|^2\big) \\
\mathbf{J} &= \mathbf{E}\times\mathbf{B} + (\operatorname{curl}^{-1}\mathbf{E})\times\operatorname{curl}\mathbf{B}\,,
\end{aligned}
\tag{3.8.2}
$$

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 3.8.1

- The first term in the current density $\mathbf{J}$ in the Maxwell form of Euler's fluid equations in (3.8.2) 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 (3.6.6).

- The divergence of the $\mathbf{D}$-equation in the Maxwell form (3.8.2) of the Euler fluid equations implies a conservation equation, given by
$$\partial_t\rho = \operatorname{div}\mathbf{J}\,. \tag{3.8.3}$$
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 (3.8.3) for $\rho = \mathrm{div}\boldsymbol{\ell}$ 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.

3.9 Euler's equations in Hodge-star form in $\mathbb{R}^4$

Definition 3.9.1 *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 3.9.2 *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

$$
F = *u_\mu \omega_\nu dx^\mu \wedge dx^\nu = \boldsymbol{\ell} \cdot d\mathbf{x} \wedge dt + \boldsymbol{\omega} \cdot d\mathbf{S} \,. \qquad (3.9.1)
$$

★

Exercise. Show that Euler's fluid equations (3.5.3) imply

$$F = d\Big(\mathbf{v}\cdot d\mathbf{x} - (p + \tfrac{1}{2}u^2)dt\Big)\,, \tag{3.9.2}$$

in the Euler fluid notation of equation (3.5.4). ★

After the preparation of having solved these exercises, it is an easy computation to show that the Helmholtz vorticity equation (3.5.15) 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 (3.8.2). 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 $\mathbb{R}^4$

$$G = \boldsymbol{\ell}\cdot d\mathbf{S} + d\chi\wedge dt, \tag{3.9.3}$$

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 (3.8.2). First, the exterior derivative of G, given by

$$\begin{aligned} dG &= \partial_t\boldsymbol{\ell}\cdot d\mathbf{S}\wedge dt + \mathrm{div}\,\boldsymbol{\ell}\,d^3x && (3.9.4)\\ &= \mathbf{J}\cdot d\mathbf{S}\wedge dt + \rho\,d^3x =: \boldsymbol{J}\,, && (3.9.5)\end{aligned}$$

recovers the two relations $\partial_t\boldsymbol{\ell} = \mathbf{J}$ and $\mathrm{div}\,\boldsymbol{\ell} = \rho$ in (3.8.2). 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

$$d^2G = \Big(\mathrm{div}\,\mathbf{J} - \partial_t\rho\Big)d^3x\wedge dt = 0\,.$$

This recovers the conservation law in (3.8.3) 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,

$$dF = 0 \quad \text{and} \quad dG = \boldsymbol{J}\,, \tag{3.9.6}$$

where the 2-forms F, G and the 3-form $\boldsymbol{J}$ are given in (3.9.1), (3.9.3) and (3.9.5), respectively.

Exercise. Show that equations (3.9.6) for the differential representation of Euler's fluid equations in $\mathbb{R}^4$ may be written as a pair of partial differential equations,

$$\partial_\mu F^{\mu\nu} = 0 \quad \text{and} \quad \partial_\mu G^{\mu\nu} = J^\nu, \tag{3.9.7}$$

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 (3.9.6) 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. ★

Chapter 4

Resonances and S^1 reduction

Resonant harmonic oscillators play a central role in physics. This is largely because the linearised dynamics of small excitations always leads to an eigenvalue problem. Excitations oscillate. Nonlinear oscillations resonate. Under changes of parameters, resonant oscillations bifurcate. The application of these ideas in physics is immense in scope, ranging from springs, to swings, to molecules, to lasers, to coherent states in nuclear physics, to Bose-Einstein condensed (BEC) systems, to nonlinear optics, to telecommunication, to qubits in quantum computing.

This chapter studies resonance of two coupled nonlinear oscillators, using polarisation states of travelling waves in nonlinear optical fibres as our final physical application.

4.1 Dynamics of two coupled oscillators on $\mathbb{C}^2$

4.1.1 Oscillator variables in $\mathbb{C}^2$

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

$\mathbb{C}^2$ is defined by

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} q_1 + ip_1 \\ q_2 + ip_2 \end{bmatrix} = \mathbf{q} + i\mathbf{p} \in \mathbb{C}^2 \,. \tag{4.1.1}$$

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

$$\begin{aligned} dq_j \wedge dp_j &= \frac{1}{(-2i)}(dq_j + idp_j) \wedge (dq_j - idp_j) \\ &= \frac{1}{(-2i)} da_j \wedge da_j^* \,. \end{aligned} \tag{4.1.2}$$

Likewise, the Poisson bracket transforms by the chain rule as

$$\begin{aligned} \{a_j, a_k^*\} &= \{q_j + ip_j \,,\, q_k - ip_k\} \\ &= -2i\,\{q_j \,,\, p_k\} = -2i\,\delta_{jk} \,. \end{aligned} \tag{4.1.3}$$

Thus, in oscillator variables Hamilton's canonical equations become

$$\begin{aligned} \dot{a}_j &= \{a_j,\, H\} = -2i\,\frac{\partial H}{\partial a_j^*} \,, \\ \text{and} \quad \dot{a}_j^* &= \{a_j^*,\, H\} = 2i\,\frac{\partial H}{\partial a_j} \,. \end{aligned} \tag{4.1.4}$$

The corresponding Hamiltonian vector field is

$$X_H = \{\cdot\,, H\} = -2i\left(\frac{\partial H}{\partial a_j^*}\frac{\partial}{\partial a_j} - \frac{\partial H}{\partial a_j}\frac{\partial}{\partial a_j^*}\right), \tag{4.1.5}$$

satisfying the defining property of a Hamiltonian system,

$$X_H \,\lrcorner\, \frac{1}{(-2i)} da_j \wedge da_j^* = dH \,. \tag{4.1.6}$$

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

$$\Big\langle\!\!\Big\langle \cdot\,, \cdot \Big\rangle\!\!\Big\rangle :\, \mathbb{C}^2 \otimes \mathbb{C}^2 \mapsto \mathbb{R} \,. \tag{4.1.7}$$

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

$$\Big\langle\!\!\Big\langle \mathbf{a}\,,\,\mathbf{b} \Big\rangle\!\!\Big\rangle = \mathbf{a}^* \cdot \mathbf{b} = \big[a_1^*,\, a_2^*\big] \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = a_1^* b_1 + a_2^* b_2 \,. \tag{4.1.8}$$

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

$$|\mathbf{a}|^2 = \langle \mathbf{a}\,,\,\mathbf{a}\rangle = \mathbf{a}^* \cdot \mathbf{a} = a_1^* a_1 + a_2^* a_2 \,. \tag{4.1.9}$$

For additional introductory discussion of simple harmonic motion, see Appendix A.

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

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

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

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

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

The ***determinant*** $\det U$ of a unitary matrix U is a complex number of unit modulus; that is, $\det U$ belongs to $U(1)$.

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

$$\begin{bmatrix} e^{i(r+s)} & 0 \\ 0 & e^{i(r-s)} \end{bmatrix} = \begin{bmatrix} e^{ir} & 0 \\ 0 & e^{ir} \end{bmatrix} \begin{bmatrix} e^{is} & 0 \\ 0 & e^{-is} \end{bmatrix} , \tag{4.1.10}$$

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

Definition 4.1.1 (Diagonal S^1 action)
The action of a $U(1)_{diag}$ phase shift $S^1 : \mathbb{C}^2 \mapsto \mathbb{C}^2$ on a complex two-vector with real angle parameter s is given by

$$\mathbf{a}(s) = e^{-2is}\mathbf{a}(0) \quad \textit{and} \quad \mathbf{a}^*(s) = e^{2is}\mathbf{a}^*(0)\,, \tag{4.1.11}$$

which solves,

$$\frac{d}{ds}\mathbf{a}(s) = -2i\mathbf{a} \quad \textit{and} \quad \frac{d}{ds}\mathbf{a}^*(s) = 2i\mathbf{a}^*\,. \tag{4.1.12}$$

This operation is generated canonically by the Hamiltonian vector field

$$X_R = \{\cdot\,, R\} = -2i\,\mathbf{a}\cdot\frac{\partial}{\partial\mathbf{a}} + 2i\,\mathbf{a}^*\cdot\frac{\partial}{\partial\mathbf{a}^*} =: \frac{d}{ds}\,, \tag{4.1.13}$$

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

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

*in (4.1.3). The variable R canonically conjugate to the **angle** s is called its corresponding **action** and is given by*

$$R = |a_1|^2 + |a_2|^2 = |\mathbf{a}|^2 \quad \textit{for which} \quad \{s, R\} = 1\,. \tag{4.1.14}$$

Definition 4.1.2 ($1:1$ resonance dynamics)

- *The flow ϕ_t^R of the Hamiltonian vector field $X_R = \{\,\cdot\,,\,R\}$ is the $1:1$ resonant phase shift. In this flow, the phases of both simple harmonic oscillations (a_1, a_2) precess counterclockwise at the same constant rate.*
- *In $1:1$ resonant motion on $\mathbb{C}^2 = \mathbb{C}\times\mathbb{C}$, the complex two amplitudes a_1 and a_2 stay in phase, because they both oscillate at the same rate.*

Remark 4.1.3 (Restricting to S^1-invariant variables)
Choosing a particular value of R restricts $\mathbb{C}^2 \to S^3$. As we shall see, further restriction to S^1-invariant variables will locally *map*

$$\mathbb{C}^2/S^1 \to S^3/S^1 = S^2.$$

4.1.3 The S^1-invariant Hermitian coherence matrix

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

$$Q = \mathbf{a} \otimes \mathbf{a}^* = \begin{bmatrix} a_1 a_1^* & a_1 a_2^* \\ a_2 a_1^* & a_2 a_2^* \end{bmatrix} . \tag{4.1.15}$$

Its matrix properties are

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

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

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

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

4.1.4 The Poincaré sphere $S^2 \in S^3$

We expand the Hermitian matrix Q in (4.1.15) as a linear combination of the four 2×2 Pauli spin matrices $(\sigma_0\,, \boldsymbol{\sigma})$, with $\boldsymbol{\sigma} = (\sigma_1,\, \sigma_2,\, \sigma_3)$ given by

$$\begin{aligned} \sigma_0 &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \\ \sigma_2 &= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \end{aligned} \tag{4.1.16}$$

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

The result of this linear expansion is, in vector notation,

$$Q = \frac{1}{2}\Big(R\sigma_0 + \mathbf{Y}\cdot\boldsymbol{\sigma}\Big)\,. \tag{4.1.17}$$

Consequently, the vector $\mathbf{Y}$ is determined by the trace formula,

$$\mathbf{Y} = \mathrm{tr}\,(Q\,\boldsymbol{\sigma}) = a_k^*\boldsymbol{\sigma}_{kl}a_l\,. \tag{4.1.18}$$

Definition 4.1.4 (Poincaré sphere)
The coefficients $(R,\,\mathbf{Y})$ in the expansion of the matrix Q in (4.1.17) are the four quadratic S^1 invariants,

$$\begin{aligned} R &= |a_1|^2 + |a_2|^2\,, \\ Y_3 &= |a_1|^2 - |a_2|^2\,, \quad \textit{and} \\ Y_1 + i\,Y_2 &= 2a_1^*a_2\,. \end{aligned} \tag{4.1.19}$$

These satisfy the relation,

$$\det Q = R^2 - |\mathbf{Y}|^2 = 0\,, \quad \textit{with} \quad |\mathbf{Y}|^2 \equiv Y_1^2 + Y_2^2 + Y_3^2\,, \tag{4.1.20}$$

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

Definition 4.1.5 (Stokes vector)
The matrix decomposition (4.1.17) of the Hermitian matrix Q in equation (4.1.15) into the Pauli spin matrix basis (4.1.16) may be written explicitly as

$$\begin{aligned} Q = \mathbf{a}\otimes\mathbf{a}^* &= \begin{bmatrix} a_1a_1^* & a_1a_2^* \\ a_2a_1^* & a_2a_2^* \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} R+Y_3 & Y_1 - iY_2 \\ Y_1 + iY_2 & R - Y_3 \end{bmatrix}, \end{aligned} \tag{4.1.21}$$

or, on factoring $R > 0$,

$$\begin{aligned} Q &= \frac{R}{2}\left(\sigma_0 + \frac{\mathbf{Y}\cdot\boldsymbol{\sigma}}{R}\right) \\ &= \frac{R}{2}\begin{bmatrix} 1+s_3 & s_1 - is_2 \\ s_1 + is_2 & 1 - s_3 \end{bmatrix} \\ &= \frac{R}{2}\begin{bmatrix} 1+\cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & 1-\cos\theta \end{bmatrix}, \end{aligned} \tag{4.1.22}$$

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

$$\begin{aligned} \mathbf{Y}/R = (Y_1, Y_2, Y_3)/R &= (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \\ &=: (s_1, s_2, s_3) = \mathbf{s}\,. \end{aligned} \tag{4.1.23}$$

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

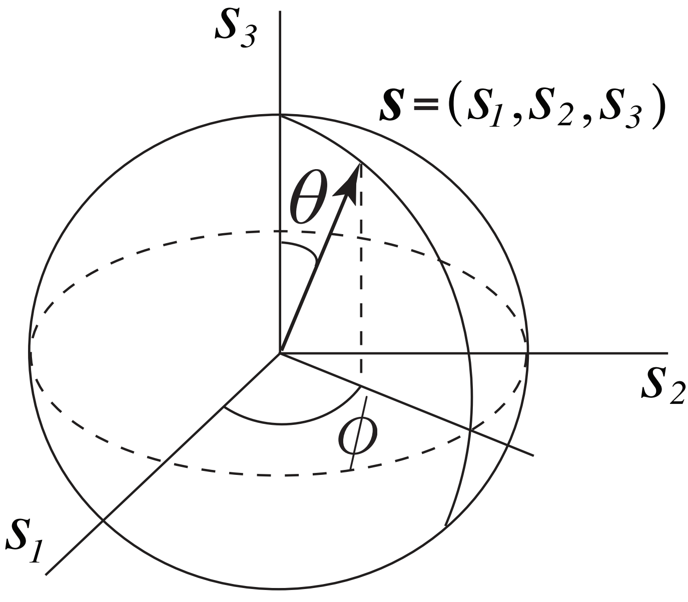


Figure 4.1: The Stokes vector on the unit Poincaré sphere.

4.1.5 $1:1$ resonance: Quotient map and orbit manifold

Definition 4.1.6 (Quotient map)
The $1:1$ ***quotient map*** $\pi:\ \mathbb{C}^2\backslash\{0\}\to\mathbb{R}^3\backslash\{0\}$ *is defined (for* $R>0$*) by means of the formula (4.1.19). This map may be expressed as*

$$\mathbf{Y} := (Y_1 + iY_2,\, Y_3) = \pi(\mathbf{a})\,. \tag{4.1.24}$$

Explicitly, this is

$$\mathbf{Y} = \operatorname{tr} Q\,\boldsymbol{\sigma} = a_k^* \boldsymbol{\sigma}_{kl} a_l\,, \tag{4.1.25}$$

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

Remark 4.1.7 *Since* $\{\mathbf{Y},\, R\} = 0$*, the quotient map* π *in (4.1.24) collapses each* $1:1$ *orbit to a point. The converse also holds, namely that the inverse of the quotient map* $\pi^{-1}\mathbf{Y}$ *for* $\mathbf{Y}\in\operatorname{Image}\pi$ *consists of a* $1:1$ *orbit* (S^1)*.*

Definition 4.1.8 (Orbit manifold)
The image in $\mathbb{R}^3$ *of the quotient map* $\pi:\ \mathbb{C}^2\backslash\{0\}\to\mathbb{R}^3\backslash\{0\}$ *in (4.1.24) is the* ***orbit manifold*** *for the* $1:1$ *resonance.*

Remark 4.1.9 ($1:1$ orbit manifold is the Poincaré sphere)
The image of the quotient map π *in (4.1.24) may be conveniently displayed as the zero-level set of the relation* $\det Q = 0$ *using the* S^1*-invariant variables in equation (4.1.21)*

$$\begin{aligned}\det Q &= C(Y_1,\, Y_2,\, Y_3,\, R)\\ &:= (R+Y_3)(R-Y_3) - \left(Y_1^2 + Y_2^2\right) = 0\,.\end{aligned} \tag{4.1.26}$$

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

For the $1:1$ *resonance, the image in* $\mathbb{R}^3$ *of the quotient map* π *in (4.1.24) is the* ***Poincaré sphere*** S^2*.*

Remark 4.1.10 (The Poincaré sphere $S^3 \simeq S^2 \times S^1$)
Being invariant under the flow of the Hamiltonian vector field $X_R = \{\,\cdot\,,\,R\}$, each point on the Poincaré sphere S^2 consists of a resonant S^1 orbit under the $1:1$ *circle action*

$$\begin{aligned} \phi_{1:1} : \mathbb{C}^2 \mapsto \mathbb{C}^2 \quad \textit{as} \quad & (a_1, a_2) \to (e^{i\phi} a_1, e^{i\phi} a_2) \\ \textit{and} \quad & (a_1^*, a_2^*) \to (e^{-i\phi} a_1^*, e^{-i\phi} a_2^*)\,. \end{aligned} \tag{4.1.27}$$

Exercise. What corresponds to the quotient map, orbit manifold (image of the quotient map) and Poincaré sphere for the transmission of optical rays by Fermat's principle in an axisymmetric, translation-invariant medium in Chapter 1? ★

4.1.6 The basic qubit: Quantum computing in the Bloch ball

In quantum mechanics, the Poincaré sphere is known as the Bloch sphere, and it corresponds to a two-level atomic system [FeVeHe1957, AlEb1975, DaBaBi1992, BeBuHo1994]. In particular, the Hermitian matrix Q in (4.1.22) is the analog of the ***density matrix*** (denoted as $\boldsymbol{\rho}$) in quantum mechanics. The density matrix $\boldsymbol{\rho}$ is a key element in the applications of quantum mechanics, including quantum computing. In traditional computing, a ***bit*** is a scalar that can assume either of the values 0, or 1. In quantum computing, a ***qubit*** is a vector in a two-dimensional complex Hilbert space that can assume the values *up* and *down*, as well as all other values intermediate between them in a certain sense. The basic qubit in quantum computing is a two-level spin system whose density matrix $\boldsymbol{\rho}$ was first introduced in [FeVeHe1957] for describing MASER dynamics.[2] This is the 2×2 positive Hermitian matrix with unit trace,

$$\boldsymbol{\rho} = \frac{1}{2}\Big(\sigma_0 + \mathbf{r}\cdot\boldsymbol{\sigma}\Big) = \frac{1}{2}\left[\begin{array}{cc} 1+r_3 & r_1 - ir_2 \\ r_1 + ir_2 & 1 - r_3 \end{array}\right], \tag{4.1.28}$$

so that $\mathbf{r} = \operatorname{tr} \boldsymbol{\rho}\,\boldsymbol{\sigma}$.

[2]The acronym MASER stands for *Microwave Amplification by Stimulated Emission of Radiation*.

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

$$\det \boldsymbol{\rho} = 1 - |\mathbf{r}|^2 \,, \tag{4.1.29}$$

will also be positive, provided one requires

$$|\mathbf{r}|^2 = r_1^2 + r_2^2 + r_3^2 \leq 1 \,, \tag{4.1.30}$$

which defines the ***Bloch ball***.

Definition 4.1.11 (Bloch ball)
The Bloch ball is the locus of states for a given density matrix $\boldsymbol{\rho}$ in (4.1.28) satisfying $|\mathbf{r}| \leq 1$.

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

$$\psi(t) = a_1(t) \begin{bmatrix} 1 \\ 0 \end{bmatrix} + a_2(t) \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a_1(t) \\ a_2(t) \end{bmatrix} .$$

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

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

Restriction of the density matrix $\boldsymbol{\rho}|_{r=1}$ to the surface of the Bloch sphere $|\mathbf{r}| = 1$ recovers the Hermitian matrix Q in (4.1.15) whose vanishing determinant defines the unit Poincaré sphere, with $R = 1$.

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

Definition 4.1.12 (Pure vs impure states)
*A given state corresponding to density matrix $\boldsymbol{\rho}$, is **impure**, or **mixed**, if*

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

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

Remark 4.1.13 (Optics vs quantum mechanics)
The quantum mechanical case deals with mixed and pure states inside and on the surface of the Bloch ball, respectively. The dynamics of optical polarisation directions on the Poincaré sphere deals exclusively with pure states.

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

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

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

Here, the quantity

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

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

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

Consequently, ξ alone (without multiple i) is a 2×2 traceless Hermitian matrix, $\xi^\dagger = \xi$, which may be written as a sum over the Pauli matrices in (4.1.16) as

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

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

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

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

Definition 4.2.1 (Momentum map $J : T^*\mathbb{C}^2 \mapsto su(2)^*$)
*The **momentum map**, $J(\mathbf{a}^*, \mathbf{a}) : T^*\mathbb{C}^2 \simeq \mathbb{C}^2 \times \mathbb{C}^2 \mapsto su(2)^*$ for the action of $SU(2)$ on $\mathbb{C}^2$ is defined by*

$$
\begin{aligned}
J^\xi(\mathbf{a}^*, \mathbf{a}) &:= \Big\langle J(\mathbf{a}^*, \mathbf{a}),\, \xi \Big\rangle_{su(2)^* \times su(2)} = \Big\langle\!\!\Big\langle \mathbf{a}^*,\, \pounds_{\xi_U}\mathbf{a} \Big\rangle\!\!\Big\rangle_{\mathbb{C}^2 \times \mathbb{C}^2} \\
&:= a_k^* \xi_{kl} a_l = a_l a_k^* \xi_{kl} \\
&:= \operatorname{tr}\big((\mathbf{a} \otimes \mathbf{a}^*)\xi\big) = \operatorname{tr}\big(Q\xi\big)\,. \qquad (4.2.1)
\end{aligned}
$$

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

Remark 4.2.2 *This map may also be expressed using the canonical symplectic form, $\Omega(\mathbf{a}, \mathbf{b}) = \operatorname{Im}(\mathbf{a}^* \cdot \mathbf{b})$ on $\mathbb{C}^2$ as in [MaRa1994]*

$$
\begin{aligned}
J^\xi(\mathbf{a}^*, \mathbf{a}) &:= \Omega(\mathbf{a}, \xi_U(\mathbf{a})) = \Omega(\mathbf{a}, i\xi\mathbf{a}) \\
&= \operatorname{Im}(a_k^*(i\xi)_{kl}a_l) = a_k^*\xi_{kl}a_l = \operatorname{tr}\big(Q\xi\big)\,. \qquad (4.2.2)
\end{aligned}
$$

Remark 4.2.3 (Removing the trace)
Being traceless, ξ has zero pairing with any multiple of the identity; so one may remove the trace of Q by subtracting $\operatorname{tr} Q$ times a multiple of the identity. Thus, the momentum map

$$
J(\mathbf{a}^*, \mathbf{a}) = \tilde{Q} = Q - \frac{1}{2} Id \operatorname{tr} Q \in su(2)^* \qquad (4.2.3)
$$

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

$$
\Big\langle \tilde{Q},\, \xi \Big\rangle = \operatorname{tr}\big(\tilde{Q}\xi\big) \quad \textit{for} \quad \tilde{Q} \in su(2)^* \textit{ and } \xi \in su(2)\,. \qquad (4.2.4)
$$

Remark 4.2.4 (Momentum map and Poincaré sphere)
A glance at equation (4.1.17) reveals that the ***momentum map*** $\mathbb{C}^2\times\mathbb{C}^2 \mapsto su(2)^*$ *for the action of* $SU(2)$ *acting on* $\mathbb{C}^2$ *in equation (4.2.3) is none other than the map* $\mathbb{C}^2 \mapsto S^2$ *to the Poincaré sphere. To see this, one simply replaces* $\xi \in su(2)$ *by the vector of Pauli matrices* $\boldsymbol{\sigma}$ *in equation (4.1.17) to recover the* ***Hopf map****,*

$$\tilde{Q} = Q - \frac{1}{2}R\sigma_0 = \frac{1}{2}\,\mathbf{Y}\cdot\boldsymbol{\sigma} \tag{4.2.5}$$

in which

$$\mathbf{Y} = \operatorname{tr}\tilde{Q}\,\boldsymbol{\sigma} = a_k\boldsymbol{\sigma}_{kl}a_l^*\,. \tag{4.2.6}$$

Thus, the traceless *momentum map,* $J\,:\, T^*\mathbb{C}^2 \simeq \mathbb{C}^2\times\mathbb{C}^2 \mapsto su(2)^*$ *in equation (4.2.3) explicitly recovers the components of the vector* $\mathbf{Y}$ *on the* ***Poincaré sphere*** $|\mathbf{Y}|^2 = R^2$*. Namely,*

$$\tilde{Q} = \frac{1}{2}\left[\begin{array}{cc} Y_3 & Y_1 - iY_2 \\ Y_1 + iY_2 & -Y_3 \end{array}\right]. \tag{4.2.7}$$

Remark 4.2.5 (Poincaré sphere and Hopf fibration of S^3)
The sphere $|\mathbf{Y}|^2 = R^2$ *defined by the map* $\mathbb{C}^2/S^1 \mapsto S^2 \simeq S^3/S^1$ *for 1:1 resonance was first introduced by Poincaré to describe the two transverse polarisation states of light [Po1892, BoWo1965]. It was later studied by Hopf [Ho1931], who showed that it has interesting topological properties. Namely, it is a* ***fibration*** *of* $S^3 \simeq SU(2)$*. That is, the Poincaré sphere* $S^2 = S^3/S^1$ *has an* S^1*-fibre sitting over every point on the sphere [Ho1931]. This reflects the* $SU(2)$ *group decomposition which locally factorises into* $S^2\times S^1$ *at each point on the sphere* S^2*. For* $R=1$*, this may be expressed as a matrix factorisation for any* $A\in SU(2)$ *with* $s,\phi\in[0,2\pi)$ *and* $\theta\in[0,\pi/2)$*, as*

$$\begin{aligned} A &= \begin{pmatrix} a_1^* & -a_2 \\ a_2^* & a_1 \end{pmatrix} \\ &= \begin{pmatrix} -i\cos\theta & -\sin\theta e^{i\phi} \\ \sin\theta e^{-i\phi} & i\cos\theta \end{pmatrix}\begin{pmatrix} \exp(-is) & 0 \\ 0 & \exp(is) \end{pmatrix}. \end{aligned} \tag{4.2.8}$$

For further discussion of the ***Hopf fibration*** *in the context of geometric dynamics for resonant coupled oscillators, see [Ku1976, Ku1986, El2006].*

4.2.1 Coherence matrix dynamics for the $1:1$ resonance

For a given Hamiltonian $H:\ \mathbb{C}^2\mapsto\mathbb{R}$, the matrix $Q(t)=\mathbf{a}\otimes\mathbf{a}^*(t)$ in (4.1.15) evolves canonically according to

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

Thus, by the decomposition (4.1.17) of $Q(t)$ in an $su(2)$ basis of Pauli matrices for an S^1-invariant Hamiltonian $H(\mathbf{Y},R)$ we have

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

where we have used $\dot{R}=\{R,\,H\}=0$ which follows because of the S^1 symmetry of the Hamiltonian. The Poisson brackets of the components $(Y_1,Y_2,Y_3)\in\mathbb{R}^3$ are computed by the ***product rule*** to close among themselves. This is expressed in tabular form as

$$\{Y_i,\,Y_j\}=\begin{array}{|c|ccc|}\hline \{\,\cdot\,,\,\cdot\,\} & Y_1 & Y_2 & Y_3\\ \hline Y_1 & 0 & 4Y_3 & -4Y_2\\ Y_2 & -4Y_3 & 0 & 4Y_1\\ Y_3 & 4Y_2 & -4Y_1 & 0\\ \hline\end{array} \qquad (4.2.9)$$

or, in index notation,

$$\{Y_k\,,\,Y_l\}=4\epsilon_{klm}Y_m\,. \qquad (4.2.10)$$

Proof. The proof of the result (4.2.9) is a direct verification using the chain rule for Poisson brackets,

$$\{Y_i,\,Y_j\}=\frac{\partial Y_i}{\partial z_A}\{z_A,\,z_B\}\frac{\partial Y_j}{\partial z_A}\,, \qquad (4.2.11)$$

for the invariant bilinear functions $Y_i(z_A)$ in (1.4.5). In (4.2.11), one denotes $z_A=(a_A,a_A^*)$, with $A=1,2,3$. ■

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

$$\{F\,,\,H\}(\mathbf{Y})=4\mathbf{Y}\cdot\frac{\partial F}{\partial\mathbf{Y}}\times\frac{\partial H}{\partial\mathbf{Y}}\,, \qquad (4.2.12)$$

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

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

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

$$\dot{\mathbf{Y}} = -\,4\mathbf{Y}\times\frac{\partial H}{\partial \mathbf{Y}}\,, \tag{4.2.13}$$

which of course preserves the radius of the Poincaré sphere $|\mathbf{Y}| = R$. Equation (4.2.13) proves the following.

Theorem 4.2.6 (The momentum map (4.2.5) is Poisson)
The traceless momentum map, $J\,:\,T^*\mathbb{C}^2\simeq\mathbb{C}^2\times\mathbb{C}^2\mapsto su(2)^*$ in equation (4.2.5) is a Poisson map. That is, it satisfies the Poisson property (2.3.26) for smooth functions F and H,

$$\big\{F\circ J\,,\,H\circ J\big\} = \big\{F\,,\,H\big\}\circ J\,. \tag{4.2.14}$$

*This relation defines a **Lie-Poisson bracket** on $su(2)^*$ that inherits the defining properties of a Poisson bracket from the canonical relations*

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

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

Remark 4.2.7 *The Hamiltonian vector field in $\mathbb{R}^3$ for the evolution of the components of the coherence matrix $\tilde{Q}$ in (4.2.7) has the same form, modulo the factor of $(-\,4)$, as Euler's equations (2.4.5) for a rigid body.*

4.2.2 Poisson brackets on the surface of a sphere

The $\mathbb{R}^3$-bracket (4.2.12) for functions of the vector $\mathbf{Y}$ on the sphere $|\mathbf{Y}|^2 = const$ is expressible as

$$\begin{aligned}\{F,\, H\}(\mathbf{Y}) \;=\; & 4Y_3\Big(\frac{\partial F}{\partial Y_1}\frac{\partial H}{\partial Y_2} - \frac{\partial F}{\partial Y_2}\frac{\partial H}{\partial Y_1}\Big) + 4Y_1\Big(\frac{\partial F}{\partial Y_2}\frac{\partial H}{\partial Y_3} - \frac{\partial F}{\partial Y_3}\frac{\partial H}{\partial Y_2}\Big)\\ & + 4Y_2\Big(\frac{\partial F}{\partial Y_3}\frac{\partial H}{\partial Y_1} - \frac{\partial F}{\partial Y_1}\frac{\partial H}{\partial Y_3}\Big)\,.\end{aligned} \tag{4.2.15}$$

This expression may be rewritten equivalently as

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

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

> **Exercise.** In spherical coordinates,
>
> $$Y_1 = r\cos\phi\sin\theta, \quad Y_2 = r\sin\phi\sin\theta, \quad Y_3 = r\cos\theta,$$
>
> where $Y_1^2 + Y_2^2 + Y_3^2 = r^2$. The volume element is,
>
> $$d^{\,3}Y = dY_1 \wedge dY_2 \wedge dY_3 = \frac{1}{3}\, dr^3 \wedge d\phi \wedge d\cos\theta \,.$$
>
> Show that the area element on the sphere satisfies
>
> $$r^2 d\phi \wedge d\cos\theta = \frac{r}{3}\left(\frac{dY_1 \wedge dY_2}{Y_3} + \frac{dY_2 \wedge dY_3}{Y_1} + \frac{dY_3 \wedge dY_1}{Y_2}\right).$$
>
> Explain why the canonical Poisson bracket
>
> $$\{F,\, H\} = \Big(\frac{\partial F}{\partial\cos\theta}\frac{\partial H}{\partial\phi} - \frac{\partial F}{\partial\phi}\frac{\partial H}{\partial\cos\theta}\Big)$$
>
> might be expected to be related to the $\mathbb{R}^3$ bracket in (4.2.15), up to a constant multiple. ★

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

$$\begin{aligned} \{F,\, H\}\, d^{\,3}Y &= \frac{1}{3}\, dr^3 \wedge dF \wedge dH(\cos\theta, \phi) \\ &= \frac{1}{3}\, dr^3 \wedge \Big(\frac{\partial F}{\partial\cos\theta}\frac{\partial H}{\partial\phi} - \frac{\partial F}{\partial\phi}\frac{\partial H}{\partial\cos\theta}\Big) d\cos\theta \wedge d\phi \,. \end{aligned} \tag{4.2.16}$$

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

$$\begin{aligned} dF \wedge dH(\cos\theta, \phi) &= \{F,\, H\}\, d\cos\theta \wedge d\phi \\ &= \Big(\frac{\partial F}{\partial\cos\theta}\frac{\partial H}{\partial\phi} - \frac{\partial F}{\partial\phi}\frac{\partial H}{\partial\cos\theta}\Big) d\cos\theta \wedge d\phi \,. \end{aligned} \tag{4.2.17}$$

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

4.2.3 Riemann projection of Poincaré sphere

The Riemann sphere may be visualised as the unit Poincaré sphere $Y_1^2 + Y_2^2 + Y_3^2 = 1$ in the three-dimensional real space $\mathbb{R}^3$. To this end, consider the stereographic projection from the unit sphere minus the point $P(\infty) = (0,0,1)$ onto the plane $Y_3 = 0$, which we identify with the complex plane by setting $\zeta = Y_1 + iY_2$. In Cartesian coordinates (Y_1, Y_2, Y_3) and coordinates (θ, ϕ) on the sphere defined by the polar angle θ and the azimuthal angle ϕ, the stereographic projection is found to be

$$\zeta = \frac{Y_1 + iY_2}{1 - Y_3} = \cot(\theta/2)\, e^{i\phi}\,. \tag{4.2.18}$$

This formula may be verified by an argument using similar triangles and the relation

$$\frac{\sin\theta}{1-\cos\theta} = \frac{\sqrt{1+\cos\theta}}{\sqrt{1-\cos\theta}} = \frac{\cos(\theta/2)}{\sin(\theta/2)} = \cot(\theta/2)\,, \tag{4.2.19}$$

as shown in Figure 4.2.

> **Exercise.** Use equation (4.2.19) and the unit sphere condition to show that
>
> $$|\zeta|^2 = \frac{1+Y_3}{1-Y_3}\,. \tag{4.2.20}$$
>
> Hence, deduce from this equation and (4.2.18) that
>
> $$Y_3 = \frac{|\zeta|^2 - 1}{|\zeta|^2 + 1} \quad \text{and} \quad Y_1 + iY_2 = \frac{2\zeta}{|\zeta|^2+1}\,. \tag{4.2.21}$$
>
> Check that $Y_3^2 + |Y_1 + iY_2|^2 = 1$. ★

Theorem 4.2.8 *Stereographic projection of the symplectic dynamics on the sphere yields the following Poisson bracket on the complex plane:*

$$\{\zeta,\, \zeta^*\} = \frac{1}{4i}\left(|\zeta|^2 + 1\right)^2. \tag{4.2.22}$$

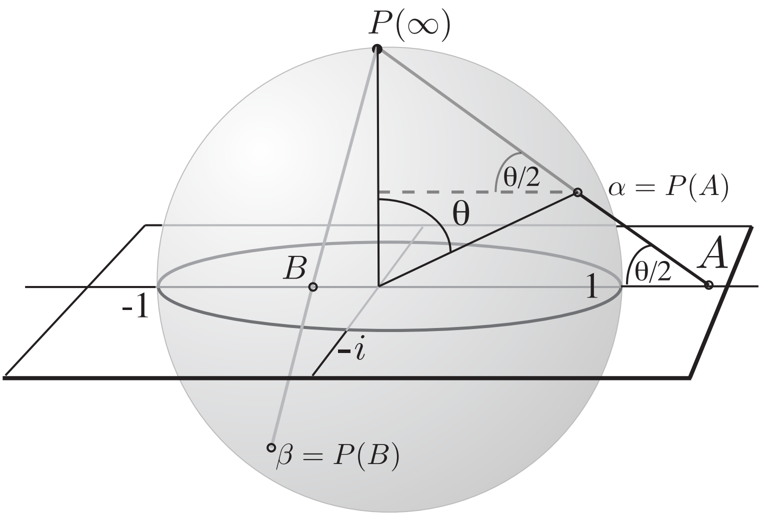


Figure 4.2: In the stereographic projection of complex numbers A and B onto points α and β of the Riemann sphere, complex numbers lying outside (resp., inside) the unit circle are projected onto the upper (resp., lower) hemisphere. The circle at complex $|\zeta| \to \infty$ is projected stereographically onto the upper pole ($\theta \to 0$) of the Riemann sphere.

Proof. By formulas (4.2.18) and (4.2.19), one finds

$$\begin{aligned} d\zeta &= \frac{e^{i\phi}}{2}(1+\cos\theta)^{-1/2}(1-\cos\theta)^{-3/2}d\cos\theta \\ &\quad + (1+\cos\theta)^{1/2}(1-\cos\theta)^{-1/2}e^{i\phi}(i\,d\phi)\,. \end{aligned}$$

Consequently, one finds the following wedge-product relations by a direct calculation,

$$\begin{aligned} d\zeta \wedge d\zeta^*(\cos\theta,\phi) &= \frac{1}{4i}\left(\frac{2}{1-\cos\theta}\right)^2 d\cos\theta\wedge d\phi \\ &= \frac{1}{4i}\big(|\zeta|^2+1\big)^2 d\cos\theta\wedge d\phi \\ &= \big\{\zeta,\,\zeta^*\big\}\, d\cos\theta\wedge d\phi\,, \end{aligned}$$

which proves equation (4.2.22) of the theorem. ■

Corollary 4.2.9 *The Riemann projection of symplectic Hamiltonian dynamics on the Poincaré sphere is given by the Poisson bracket relation,*

$$\frac{d\zeta}{dt} = \{\zeta,\, H\} = \frac{1}{4i}\left(|\zeta|^2+1\right)^2\frac{\partial H}{\partial \zeta^*}\,, \tag{4.2.23}$$

where $H(\zeta,\zeta^*)$ is the result of stereographic projection of the corresponding Hamiltonian defined on the sphere.

Remark 4.2.10 *The Hermitian coherence matrix in (4.1.22) for the* ***Poincaré sphere*** *with $|\mathbf{Y}|^2 = 1$ takes the following form in the complex Riemann variables,*

$$Q = \frac{1}{2}\begin{bmatrix} 1+Y_3 & Y_1 - iY_2 \\ Y_1 + iY_2 & 1-Y_3 \end{bmatrix} = \frac{1}{|\zeta|^2+1}\begin{bmatrix} |\zeta|^2 & \zeta^* \\ \zeta & 1 \end{bmatrix}. \tag{4.2.24}$$

Exercise. Show that the Poisson bracket (4.2.22) resulting from stereographic projection of the complex plane onto the Riemann sphere satisfies the relations,

$$\left\{\frac{\zeta}{|\zeta|^2+1},\,\frac{\zeta^*}{|\zeta|^2+1}\right\} = 2i\,\frac{|\zeta|^2-1}{|\zeta|^2+1}\,,$$

and

$$\left\{\frac{\zeta}{|\zeta|^2+1},\,\frac{|\zeta|^2-1}{|\zeta|^2+1}\right\} = \frac{4i\,\zeta}{|\zeta|^2+1}\,.$$

★

4.3 The geometric and dynamic S^1 phases

S^1-invariant dynamical solutions for $1:1$ resonance on $\mathbb{C}^2$ may be represented as motion along the intersections of a level surface of the Hamiltonian $H(\mathbf{Y})$ and the Poincaré sphere for various values of the intensity R, the radius of the sphere.

However, the Poincaré sphere does not represent the ***phase*** of the $\mathbb{C}^2$ oscillation: all states related by an S^1 phase factor correspond to a single point on the Poincaré sphere. The geometric determination of the phase evolution of a particular motion on the sphere is a story in itself.

Imagine reconstructing the solutions on $\mathbb{C}^2$ from the S^1-reduced system on the Poincaré sphere. Consider a periodic solution. As a result of travelling over one period in the S^1-reduced space (on a level surface of the Poincaré sphere) the solution may not return to its original state, because phase shifts may be generated. These phases are associated with the S^1-group action for the $1:1$ resonance symmetry responsible for the reduction. Thus, the Poincaré sphere represents a manifold of S^1 orbits. That is, the Poincaré sphere is an ***orbit manifold***, because each point on it is a $1:1$ resonant S^1 motion.

4.3.1 Geometric phase

The $1:1$ resonant S^1 reduction produces the symplectic form on the Poincaré sphere

$$\omega \;=\; \frac{1}{2}\,\mathrm{Im}\,(da_j \wedge da_j^*) = d\cos\theta \wedge\, d\phi + \frac{1}{2}dR \wedge ds\,.$$

This symplectic form arises from the following canonical one-form

$$\frac{1}{2}\,\mathrm{Im}\,(\mathbf{a}\cdot d\mathbf{a}^*) = \cos\theta\, d\phi + \frac{1}{2}\,Rds\,. \tag{4.3.1}$$

Thus, we may compute the total phase change Δs around a closed periodic orbit on the Poincaré sphere of radius R as the sum of two integrations, namely

$$\oint \frac{1}{2}\,R\,ds = \frac{R}{2}\,\Delta s = \underbrace{-\oint \cos\theta\, d\phi}_{\text{Geometric phase}} \;+\; \underbrace{\oint \frac{1}{2}\,\mathrm{Im}\,(\mathbf{a}\cdot d\mathbf{a}^*)\,.}_{\text{Dynamic phase}} \tag{4.3.2}$$

Equivalently, one may denote the total change in phase as the sum

$$\Delta s = \Delta s_{geom} + \Delta s_{dyn}\,,$$

by identifying the corresponding terms in the previous formula.

Remark 4.3.1 *By the Stokes theorem, one sees that the geometric phase associated with a periodic motion on the Poincaré sphere is proportional to the solid angle that it encloses. Thus, the name:* ***geometric phase****. Quite a large literature has built up during the past few years on this subject. For a guide to learning about this literature, the interested reader might want to consult the collection of articles in [ShWi1989].*

4.3.2 Dynamic phase

The other term in the phase formula (4.3.2) is called the ***dynamic phase***. One computes the dynamic phase as

$$\begin{aligned} \frac{R}{2}\,\Delta s_{dyn} &= \oint \frac{1}{2}\,\mathrm{Im}\,(\mathbf{a}\cdot\dot{\mathbf{a}}^*)dt = \oint \frac{1}{2}\,\mathrm{Im}\,(\mathbf{a}\cdot\{\mathbf{a}^*,\,H\})dt \qquad (4.3.3) \\ &= \oint \frac{1}{2}\,\mathrm{Im}\left(2i\,\mathbf{a}\cdot\frac{\partial H}{\partial \mathbf{a}}\right)dt \quad \text{(depends on the Hamiltonian).} \end{aligned}$$

For Hamiltonians that are S^1-invariant polynomials in C^2, the dynamic phase may be computed in terms of the orbit-averaged Hamiltonian, as in the following example. Suppose the Hamiltonian is the real-valued S^1-invariant polynomial in $\mathbb{C}^2$ in [DaHoTr1990]

$$\begin{aligned} H &= H_{linear} + H_{int} \\ &= a^*\cdot\chi^{(1)}\cdot a + \frac{3}{2}\,a^*a^*\,:\chi^{(3)}\,:aa \qquad (4.3.4) \\ &= a_j^*\chi_{jk}^{(1)}a_k + \frac{3}{2}\,a_j^*a_k^*\chi_{jklm}^{(3)}a_la_m\,. \end{aligned}$$

This is the Hamiltonian for a polarised optical travelling wave moving in a continuously varying medium with linear and nonlinear optical susceptibilities $\chi^{(1)}$ and $\chi^{(3)}$, respectively, to be treated in Section 4.5. These susceptibilities are required to be Hermitian in the appropriate indices so that H_{linear} and H_{int} are real quantities. One computes the dynamic phase for this Hamiltonian as

$$\begin{aligned} \frac{R}{2}\,\Delta s_{dyn} &= \oint \frac{1}{2}\,\mathrm{Im}\left(2i\,\mathbf{a}\cdot\frac{\partial H}{\partial \mathbf{a}}\right)dt\,, \qquad (4.3.5) \\ \text{(for this Hamiltonian)} \quad &= \oint (H + H_{int})\,dt = T(H + \langle\, H_{int}\,\rangle\,)\,, \end{aligned}$$

where T is the period of the closed orbit and $\langle \cdot \rangle := T^{-1} \oint (\,\cdot\,)\, dt$ denotes time average along the orbit.

4.3.3 Total phase

The total phase shift of a closed orbit on the Poincaré sphere is proportional to the sum of the geometric phase given by the solid angle captured by the orbit,

$$\Omega_{orbit} = - \oint_{orbit = \partial S} \cos\theta \, d\phi = \iint_S d\phi \wedge d\cos\theta \,,$$

plus the dynamic phase given by the product of the period with the sum of the constant total energy and the time-averaged interaction energy, namely,

$$\frac{R}{2} \Delta s \quad = \quad \Omega_{orbit} \; + T(H + \langle \, H_{int} \, \rangle \,) \,. \tag{4.3.6}$$

Remark 4.3.2 (S^1 gauge invariance)
The geometric phase is given by the solid angle subtended by the orbit on the Poincaré sphere. Thus, it is independent of the S^1 parameterisation of the orbit.

*On the other hand, the dynamic phase depends on the period of the orbit and also on properties of the Hamiltonian averaged in time over the orbit. The dynamic phase also depends on the parameterisation of the S^1 phase along the orbit. In particular, under an S^1**gauge transformation** given by*

$$\mathbf{a}(s) \to e^{i\psi(s)} \mathbf{a}(s) \,, \tag{4.3.7}$$

*along the orbit $\mathbf{a}(s)$ the Poincaré sphere is invariant by construction. This means the geometric phase (the area subtended by the orbit on the sphere) is also left invariant. However, the dynamic phase **changes** under this transformation, as follows:*

$$\mathrm{Im}\,(\mathbf{a} \cdot d\mathbf{a}^*) \to \mathrm{Im}\,(\mathbf{a} \cdot d\mathbf{a}^*) + d\psi \,. \tag{4.3.8}$$

This change in the dynamical phase for a problem with S^1 symmetry is reminiscent of the transformation of the vector potential in electrodynamics under the change of gauge that leaves the electric and magnetic fields invariant. As for the electromagnetic fields in that case, the geometric phase is invariant under S^1 gauge transformations.

4.4 Kummer shapes for $n : m$ resonances

This section reviews work by M. Kummer [Ku1981, Ku1986] that extends the representation of the orbit manifold for the reduced phase flow of the $(1 : 1)$ resonance in a Hamiltonian system composed of two harmonic oscillators of equal frequency to the case of $(n : m)$ resonance of any type. The orbit manifolds for $(n : m)$ resonant oscillators are two-dimensional surfaces of revolution in $\mathbb{R}^3$ called ***Kummer shapes***. These are:

(a) the Poincaré sphere S^2 for the $(1 : 1)$ resonance;

(b) spheres pinched at one pole for the $(1 : m)$ resonances when $m > 1$; and

(c) spheres pinched at both poles for the $(n : m)$ resonances when $m \geq n > 1$.

Resonant orbit manifolds represented by the corresponding Kummer shapes also exist for pseudo-oscillators – when the resonance Hamiltonian is an indefinite quadratic form. In this case, the orbit manifold is an unbounded surface of revolution, as we saw in Section 1.5 also occurs for the *hyperbolic onion* ray optics.

Resonances of this kind include the Fermi $(1 : 2)$ resonance in the CO_2 molecule. Low-order resonances are important in any physical studies of perturbed equilibria of Hamiltonian sytems and such systems have been systematically studied, see, e.g., [Ku1981, Ku1986].

Recall the canonical Poisson bracket relations (4.1.3) for oscillator variables on $\mathbb{C}^2$,

$$\{a_j, a_k^*\} = -\,2i\delta_{jk} \quad \text{and} \quad \{a_j^*, a_k\} = 2i\delta_{jk}\,. \tag{4.4.1}$$

Consider the evolution equations generated by the ***resonance phase space function*** $R: \mathbb{C}^2 \to \mathbb{R}$ given explicitly by

$$R = \frac{n}{2}\left|a_1\right|^2 + \frac{m}{2}\left|a_2\right|^2, \qquad a_j = q_j + ip_j \quad j = 1, 2\,. \tag{4.4.2}$$

The following are the ***canonical equations of*** $n : m$ ***resonant motion***:

$$\dot{a}_1 = \{a_1, R\} = -2i\,\frac{\partial R}{\partial a_1^*} = -\,nia_1\,, \tag{4.4.3}$$

$$\dot{a}_2 = \{a_2, R\} = -2i\frac{\partial R}{\partial a_2^*} = -\,mia_2\,, \tag{4.4.4}$$

whose solutions are,

$$a_1\left(t\right) = e^{-int}a_1\left(0\right) \quad \text{and} \quad a_2\left(t\right) = e^{-imt}a_2\left(0\right)\,. \tag{4.4.5}$$

These solutions represent $S^1 \times S^1 = T^2$ toroidal motion, wrapping m times around in one S^1 direction and n times around in the other.

Definition 4.4.1 ($n : m$ **resonance dynamics)**

- *The flow* ϕ_t^R *of the Hamiltonian vector field* $X_R = \{\,\cdot\,,\,R\}$ *is the* $n : m$ *resonant phase shift. In this flow, the phases of both simple harmonic oscillations* (a_1, a_2) *precess counterclockwise at constant rates, with* $a_1 \in \mathbb{C}$ *at rate* n *and* $a_2 \in \mathbb{C}$ *at rate* m*, in arbitrary frequency units.*

- *In* $n : m$ *resonant motion on* $\mathbb{C}^2 = \mathbb{C} \times \mathbb{C}$*, the variable* a_1 *oscillates* n *times while* a_2 *oscillates* m *times.*

- *Specifically, the canonical Poisson brackets with* R *generate the following* ***Hamiltonian vector field****,*

$$\begin{aligned} X_R &= \{\,\cdot\,,\,R\} \\ &= \dot{a}_1\frac{\partial}{\partial a_1} + \dot{a}_2\frac{\partial}{\partial a_2} + c.c. \\ &= -\,ina_1\frac{\partial}{\partial a_1} - \,ima_2\frac{\partial}{\partial a_2} + c.c. \end{aligned}$$

Definition 4.4.2 *The set $\{I_{n:m}\}$ of $n : m$* ***resonance invariants*** *comprises the set of independent phase space functions whose Poisson bracket vanishes when taken with R in equation (4.4.2) for a given choice of n and m. That is,*

$$I_{n:m} = \{I : \mathbb{C}^2 \to \mathbb{R} \,|\, X_R I = \{\, I\,,\, R\} = 0\}\,.$$

Definition 4.4.3 *The $n : m$ resonance invariants may be expressed conveniently by introducing a new variable $\mathbf{b} \in \mathbb{C}^2$ with two complex components,*

$$b_1 = \frac{a_1^m}{m|a_1|^{(m-1)}}\,, \quad b_2 = \frac{a_2^n}{n|a_2|^{(n-1)}}\,. \tag{4.4.6}$$

These components of $\mathbf{b} \in \mathbb{C}^2$ satisfy the canonical Poisson bracket relations; in particular,

$$\{b_1,\, b_1^*\} = -\,\frac{2i}{m}\,, \quad \{b_2,\, b_2^*\} = -\,\frac{2i}{n}\,, \tag{4.4.7}$$

as may be verified directly. However, this transformation is ***not globally defined****, because of its multivaluedness at the North and South poles. By writing these expressions in polar coordinates, as*

$$\begin{aligned} a_1 &= r_1 e^{i\phi_1}\,, \quad a_2 = r_2 e^{i\phi_2}\,, \\ b_1 &= \frac{r_1}{m}\, e^{im\phi_1}\,, \quad b_2 = \frac{r_2}{n}\, e^{in\phi_2}\,, \end{aligned} \tag{4.4.8}$$

one sees that this multivaluedness derives from the ambiguities of the polar coordinates at the North and South poles.

> **Exercise.** Compute the relations between the symplectic forms $da_j \wedge da_j^*$ and $db_j \wedge db_j^*$ for the canonical transformation (4.4.6) and thereby prove the Poisson bracket relations (4.4.7). Hint: Try this calculation in polar coordinates (4.4.8). ★

Proposition 4.4.4 *In terms of the two sets of variables $(a_1,\, a_2)$ and $(b_1,\, b_2)$, the four independent $n : m$ resonance invariants may be expressed equivalently as*

$$\begin{aligned} I_{n:m} &\in \left\{\, |b_1|^2\,, |b_2|^2\,, b_1 b_2^*, b_1^* b_2 \right\} \\ &:= \left\{ \frac{|a_1|^2}{m^2}, \frac{|a_2|^2}{n^2}, \frac{a_1^m a_2^{*n}}{D_{mn}}\,, \frac{a_1^{*m} a_2^n}{D_{mn}} \right\}, \end{aligned} \tag{4.4.9}$$

where the denominator D_{nm} is defined as

$$D_{nm} = mn|a_1|^{m-1}|a_2|^{n-1} \,. \tag{4.4.10}$$

Proof. Each of these variables in (4.4.9) is invariant under the $n : m$ resonance motion in (4.4.3) and (4.4.4), which is canonically generated by the Hamiltonian vector field X_R given by

$$\begin{aligned} X_R &= -\, ina_1 \frac{\partial}{\partial a_1} - ima_2 \frac{\partial}{\partial a_2} + c.c. \\ &= -\, imnb_1 \frac{\partial}{\partial b_1} - imnb_2 \frac{\partial}{\partial b_2} + c.c. \end{aligned}$$

That is, each of these invariants Poisson-commutes with the phase space function R. ■

Remark 4.4.5 *Although the functionally independent members of the set of resonance invariants $I \in I_{n:m}$ each Poisson-commute with the phase space function R, they do not Poisson-commute amongst themselves. The Poisson-bracket relations among these resonance invariants are the subject of Section 4.4.3.*

4.4.1 Poincaré map analog for $n : m$ resonances

The Poincaré map (4.2.3) for the $1 : 1$ resonance is the momentum map $J : T^*\mathbb{C}^2 \simeq \mathbb{C}^2 \times \mathbb{C}^2 \mapsto su(2)^*$. Accordingly, the map $Q = \mathbf{b} \otimes \mathbf{b}^* \in su(2)^*$ sends $\mathbb{C}^2 \to S^2$. We shall construct the analog of Poincaré's map to the sphere $\mathbb{C}^2 \to S^2$ for the $n : m$ resonance by transforming variables in the $1 : 1$ momentum map, as

$$\begin{aligned} Q = \mathbf{b} \otimes \mathbf{b}^* &= \begin{pmatrix} b_1 b_1^* & b_1 b_2^* \\ b_2 b_1^* & b_2 b_2^* \end{pmatrix} \\ &= \begin{pmatrix} |a_1|^2/m^2 & a_1^m a_2^{*n}/D_{mn} \\ a_1^{*m} a_2^n/D_{mn} & |a_2|^2/n^2 \end{pmatrix}, \end{aligned} \tag{4.4.11}$$

where again $D_{mn} = mn|a_1|^{m-1}|a_2|^{n-1}$ in equation (4.4.10). In the variables $\mathbf{a} \in \mathbb{C}^2$, the Poincaré map $\mathbb{C}^2 \to S^2$ in the variables $\mathbf{b} \in \mathbb{C}^2$ will transform into $\mathbb{C}^2 \to S^2|_{\text{pinch}}$, where $S^2|_{\text{pinch}}$ is a ***pinched sphere,*** which will have singularities at its North and South poles created by being pinched into corners for $n, m = 2$ and cusps for $n, m \geq 2$. These singularities are inherited from the multivalued nature of the transformation in equation (4.4.6) at the North and South poles.

4.4.2 The $n : m$ resonant orbit manifold

We shall construct the ***orbit manifold*** for $n : m$ resonance by following the same procedure of transforming to S^1 invariant variables as we did in in Section 4.1.3 when transforming $\mathbb{C}^2 \mapsto S^3 \mapsto S^2 \times S^1$ for the Poincaré sphere, or Hopf map, for $n : m = 1 : 1$. For this purpose, we first define the following $n : m$ resonant S^1-invariant quantities:

$$
\begin{aligned}
R &= \frac{n}{2}\left|a_1\right|^2 + \frac{m}{2}\left|a_2\right|^2 \,, && (4.4.12)\\
Z &= \frac{n}{2}\left|a_1\right|^2 - \frac{m}{2}\left|a_2\right|^2 \,, && (4.4.13)\\
X - iY &= 2a_1^m a_2^{*n} \,. && (4.4.14)
\end{aligned}
$$

Definition 4.4.6 (Quotient map)
The $n : m$ ***quotient map*** *$\pi : \mathbb{C}^2\backslash\{0\} \to \mathbb{R}^3\backslash\{0\}$ is defined (for $R > 0$) by means of the formulas (4.4.13) and (4.4.14). This map may be rewritten succinctly as*

$$
\mathbf{X} := (X - iY,\, Z) = \pi(\mathbf{a})\,. \tag{4.4.15}
$$

Remark 4.4.7 *Since $\{K,\, R\} = 0$ for $K \in \{X, Y, Z\}$, the quotient map π in (4.4.15) collapses each $n : m$ orbit to a point. The converse also holds, namely that the inverse of the quotient map $\pi^{-1}\mathbf{X}$ for $\mathbf{X} \in \mathrm{Image}\,\pi$ consists of an $n : m$ orbit.*

Definition 4.4.8 (Orbit manifold)
The image in $\mathbb{R}^3$ of the quotient map $\pi : \mathbb{C}^2\backslash\{0\} \to \mathbb{R}^3\backslash\{0\}$ in (4.4.15) is the ***orbit manifold*** *for the $n : m$ resonance.*

Remark 4.4.9 *The* ***orbit manifold*** *for the $n : m$ resonance may be conveniently described as the zero level set of a function in $\mathbb{R}^3$, defined via the* *analogy with the Poincaré sphere. Namely, apply the relation* $\det Q = 0$ *for the* $n : m$ *momentum map (4.4.11), which is equivalent to*

$$
\begin{aligned}
|X - iY|^2 &= X^2 + Y^2\,, \\
&= 4\,|a_1|^{2m}\,|a_2|^{2n}\,, \\
&= 4\left(\frac{R+Z}{n}\right)^m \left(\frac{R-Z}{m}\right)^n .
\end{aligned}
$$

This identity defines the functional relation for the orbit manifold for $n : m$ *resonance,*

$$
C(X,\,Y,\,Z,\,R) := X^2 + Y^2 - 4\left(\frac{R+Z}{n}\right)^m \left(\frac{R-Z}{m}\right)^n = 0\,. \quad (4.4.16)
$$

The result recovers the Poincaré sphere relation (4.1.20) for the $1 : 1$ *resonance. Consequently, a level set of* R *in the* $n : m$ *resonance map* $\mathbb{C}^2 \to S^2|_{\rm pinch}$ *obtained by transforming to* S^1 *invariants yields an orbit manifold defined by* $C(X,\,Y,\,Z,\,R) = 0$ *in 3D. A selection of these orbit manifolds is plotted in Figure 4.3. Since they were introduced by M. Kummer in [Ku1981, Ku1986] these orbit manifolds are called* ***Kummer shapes****.*

Proposition 4.4.10 *The orbit manifold* $C(X,\,Y,\,Z,\,R) = 0$ *in equation (4.4.16) is a surface of revolution about the* Z*-axis. For* $n : m = 1 : 1$*, this surface is a sphere. When either or both of the integers* $(n,\,m)$ *exceeds unity, this surface is closed and is equivalent to a sphere which has been* ***pinched*** *along the* Z*-axis at either or both of its poles at the points* $Z = \pm R$*. These points correspond to placing all of the conserved intensity* R *into only one of the oscillators. When either of* $(n,\,m)$ *is negative, the shape* $C(X,\,Y,\,Z,\,R) = 0$ *is an open surface of revolution.*

Proof. To see these azimuthally symmetric surfaces are *pinched*, solve $C(X,\,0,\,Z,\,R) = 0$ for $X(Z)$ in the $Y = 0$ plane for a constant value of R, to find

$$
X(Z) = 2\left(\frac{R+Z}{n}\right)^{m/2} \left(\frac{R-Z}{m}\right)^{n/2} .
$$

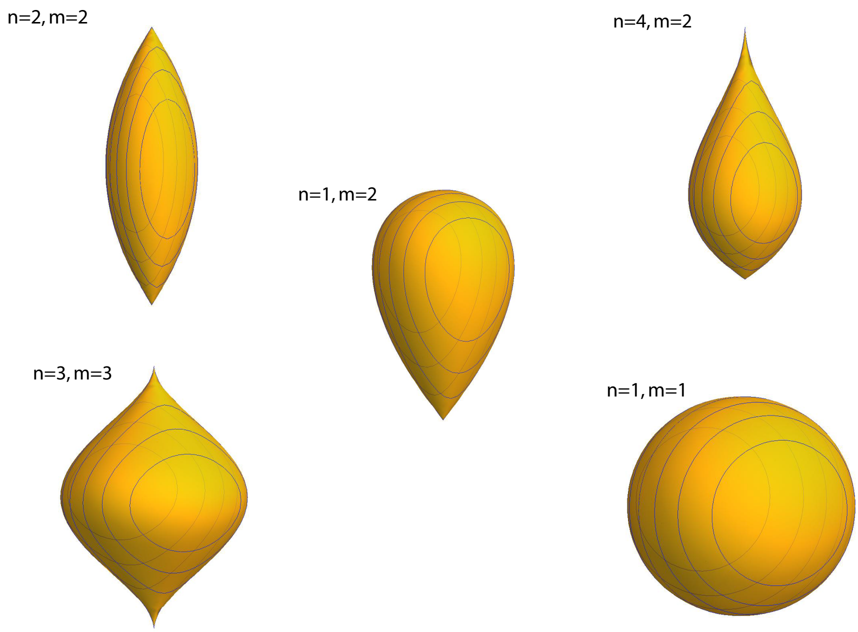


Figure 4.3: Kummer shapes are the orbit manifolds defined in equation (4.4.16) for different $n : m$ resonance values. The points at the top and bottom ($Z = \pm, R$) represent periodic solutions of period $2\pi/n$ with $a_2 = 0$ and $2\pi/m$ with $a_1 = 0$, respectively. These points on the reduced manifolds are singular *pinches* of the surfaces for $n, m \neq 1$. The curves drawn on these surfaces of revolution are intersections with vertical planes representing different level surfaces of the simple Hamiltonian $H = X$. The intersection of a vertical plane through one of the singular points is a homoclinic, or heteroclinic orbit for this Hamiltonian. The other closed intersections are periodic orbits of the quotient flow defined by the Nambu bracket in equation (4.4.17).

From this expression, one may compute the derivatives in the $Y = 0$ plane as

$$\left.\frac{dX}{dZ}\right|_{Y=0} = \frac{(R+Z)^{(m-2)/2}\,(R-Z)^{(n-2)/2}}{n^{m/2}m^{n/2}}\left[m\,(R-Z) - n\,(R+Z)\right] .$$

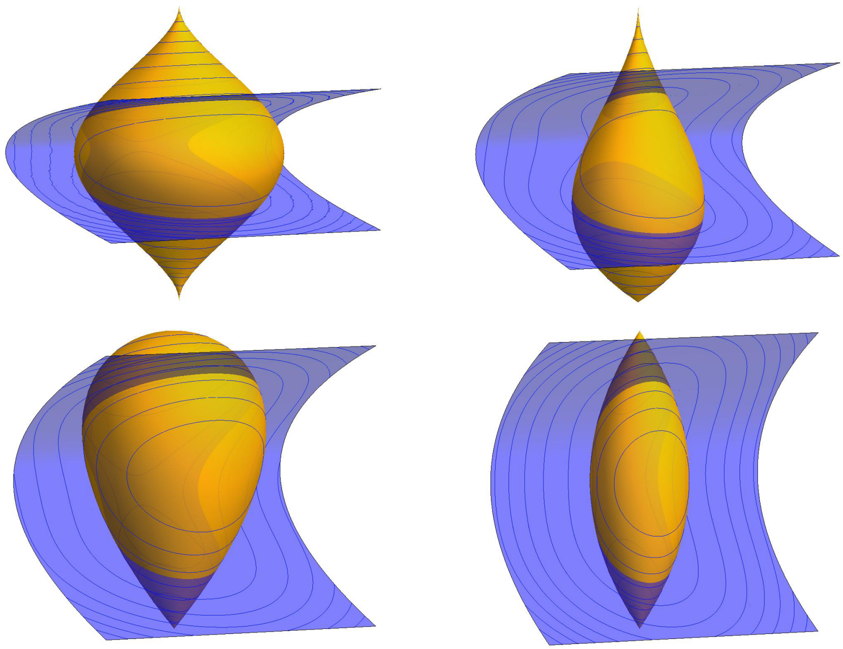


Figure 4.4: Intersections are shown of level sets of a Hamiltonian represented by a family of horizontal parabolic cylinders (in blue) with orbit manifolds for $n:m$ resonance values of $3:3$, $3:2$, $1:2$ and $2:2$, respectively, (in gold) clockwise from the upper left. If the radius of curvature of the parabola is less than that of the orbit manifold at the tangent point in the back, then the critical point is hyperbolic (unstable). Otherwise, it is elliptic (stable).

At the poles, $Z = \pm R$, these derivatives determine cusps for $n > 2$, corners for $n = 2$ and spheres for $n = 1$,

$$\left.\begin{aligned} \left.\frac{dX}{dZ}\right|_{Y=0,\,Z=R} &= 0 \quad \text{for } n > 2 \\ \left.\frac{dX}{dZ}\right|_{Y=0,\,Z=-R} &= 0 \quad \text{for } m > 2 \end{aligned}\right\} \text{Cusps at both poles (onions).}$$

$$\left.\frac{dX}{dZ}\right|_{Y=0,\,Z=\pm R} < \infty \quad \text{for } m, n = 2\,,$$

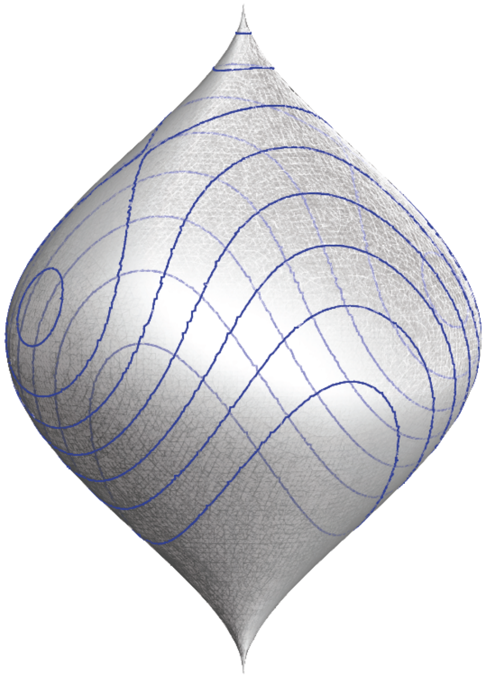


Figure 4.5: Because of its two circles of inflection points, the quotient flow on the Kummer shape for the $3:3$ resonance has an interesting ***double pitchfork bifurcation*** near its upper singular point, for a Hamiltonian represented by an ellipsoid of revolution about the vertical axis whose center is shifted upward. In this situation, the singular points at the top and bottom are always centres. The double pitchfork bifurcation creates a centre and *two* saddle-centre pairs from a single centre at the top singular point.

with a "beet" shape for $m = 1,\ n = 2.$

$$\left.\frac{dX}{dZ}\right|_{Y=0,\ Z=\pm R} \to \infty \quad \text{for } m, n \to 1 \text{ yields the Poincaré sphere.}$$

The azimuthally symmetric, closed Kummer shapes in three dimensions are reminiscent of the natural shapes of fruits and vegetables. One may indicate what happens at the singular points for $m, n \neq 1$, by identifying the shapes with fruits and vegetables, as follows:

$m = 2, n > 2 \to$ beet;

$m = 2, n = 2 \rightarrow$ lemon;

$m, n > 2 \rightarrow$ onion;

$m = 2, n = 1 \rightarrow$ turnip;

$m = 1, n = 1 \rightarrow$ orange.

■

Exercise. What happens to the Kummer shapes when $mn < 0$? The $1 : -1$ resonance produces the *hyperbolic onion* of Chapter 1. For hints about how to answer this question more generally for the $n : -m$ resonance, refer to papers by Cushman, Kummer, Elipe, Lanchares, Deprit, Miller and their collaborators, as cited in the references. ★

4.4.3 $n : m$ Poisson bracket relations

Exercise.

1. Show that the canonical Poisson bracket relations,

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

imply the useful identities

$$\left\{|a_1|^2\,,\, a_1^m\right\} = 2i\,m\,a_1^m\,,$$

and

$$\left\{|a_1|^2\,,\, a_1^{*\,m}\right\} = -\,2i\,m\,a_1^{*\,m}\,,$$

as well as

$$\begin{aligned}
\{a_1^{*m}, a_1^n\} &= m\,a_1^{m-1}\,\{a_1^*, a_1^n\} \\
&= mn\,a_1^{*m-1}\,\{a_1^*, a_1\}\,a_1^{n-1} \\
&= 2i\,mn\,a_1^{*\,m-1}a_1^{n-1}\,.
\end{aligned}$$

2. Show that the $n : m$ resonance invariants,

$$
\begin{aligned}
R &= \frac{n}{2}\left|a_1\right|^2 + \frac{m}{2}\left|a_2\right|^2 \,, \\
Z &= \frac{n}{2}\left|a_1\right|^2 - \frac{m}{2}\left|a_2\right|^2 \,, \\
X - iY &= 2a_1^m a_2^{*n}
\end{aligned}
$$

satisfy the Poisson bracket relations,

$$
\big\{R, X\big\} = 0 = \big\{R, Y\big\} = \big\{R, Z\big\}\,,
$$

and

$$
\{Z, X - iY\} = i\, mn\,(X - iY)\,.
$$

3. Finally, show that

$$
\begin{aligned}
\big\{X, Y\big\} &= i\big\{a_1^m a_2^{*n} + a_1^{*m} a_2^n\,,\, a_1^m a_2^{*n} - a_1^{*m} a_2^n\big\} \\
&= -2i\big\{a_1^m a_2^{*n}, a_1^{*\,m} a_2^n\big\} \\
&= 4\left|a_1\right|^{2(m-1)}\left|a_2\right|^{2(n-1)}\left(n^2\left|a_1\right|^2 - m^2\left|a_2\right|^2\right) \\
&= -\,4nm(X^2+Y^2)\Big(\frac{m}{R+Z} - \frac{n}{R-Z}\Big)\,.
\end{aligned}
$$

The last step is evaluated on $C(X,\,Y,\,Z,\,R) = 0$. Conclude that the $\{X, Y\}$ brackets do not close linearly, but they do close in the space of $n : m$-invariant coordinates $X, Y, Z \in \mathbb{R}^3$. What role is played by the Jacobi identity in ensuring that the last bracket would close in terms of the $n : m$ invariants?

★

Question 4.4.11 *Why do these brackets close? What motion do they generate? For the answers to these questions, we return to the $\mathbb{R}^3$-bracket theory of Nambu [Na1973] introduced in Section 1.11.3.*

4.4.4 Nambu, or $\mathbb{R}^3$-bracket for $n : m$ resonance

The flow induced by Hamiltonian $H : \mathbb{R}^3 \to \mathbb{R}$ on the orbit manifold, the so-called ***quotient flow***, is governed by the Euler-like equations

$$\dot{X} = \{X,\, H\} = \nabla C \times \nabla H\,. \tag{4.4.17}$$

A quick qualitative picture of the quotient flow lines (i.e., the flow lines of H on $C = 0$) is obtained by cutting the corresponding Kummer shape with the appropriate level sets of the Hamilton $H = const.$

Here the Casimir surface for fixed R is given by the zero level set corresponding to the functional relation found above in equation (4.4.16),

$$C(X, Y, Z, R) = X^2 + Y^2 - 4\left(\frac{R+Z}{n}\right)^m \left(\frac{R-Z}{m}\right)^n = 0\,. \tag{4.4.18}$$

Both C and H are preserved by the quotient flow,

$$\begin{aligned} \dot{C} &= \nabla C \cdot \nabla C \times \nabla H = 0\,, \\ \dot{H} &= \nabla C \cdot \nabla H \times \nabla H = 0\,. \end{aligned}$$

Remark 4.4.12 *The $n : m$ resonant motion takes place in $\mathbb{R}^3$ along intersections of a surface $C(X, Y, Z, R) = 0$ (the orbit manifold for a given value R =constant) and a level surface of the Hamiltonian, H =constant.*

Remark 4.4.13 *The equations*

$$\dot{X} = \nabla C \times \nabla H$$

are unchanged by taking linear combinations

$$\dot{X} = \nabla\left(\alpha C + \beta H\right) \times \left(\gamma C + \varepsilon H\right)$$

provided

$$det \begin{pmatrix} \alpha & \beta \\ \gamma & \varepsilon \end{pmatrix} = 1\,.$$

This remark may produce some interesting simplifications when judiciously applied. See the application to the rotor and pendulum in Section 2.4.4.

Remark 4.4.14 *The singularities at the pinches in the Kummer shapes may apparently be removed by transforming to the* $\mathbf{b}$ *variables. However, this apparent regularisation of the singularities comes at the cost of introducing a multiple-sheeted relationship between the* $\mathbf{b}$ *variables for the resulting Poincaré sphere and the original* $\mathbf{a}$ *variables for the Kummer vegetable surfaces. For more details, see the enhanced coursework problem A.5.4.*

4.5 Optical travelling-wave pulses

4.5.1 Background

The problem of a single, polarised, optical laser pulse propagating as a travelling wave in an anisotropic, cubically nonlinear, lossless medium may be investigated as a Hamiltonian system,

$$\frac{d\mathbf{a}}{d\tau} = \{\mathbf{a},\, H\}\,,$$

where $\mathbf{a} \in \mathbb{C}^2$ is the complex two-component electric field amplitude of the optical pulse and s is the travelling wave variable. This Hamiltonian system describes the nonlinear travelling-wave dynamics of two complex oscillator modes (the two polarisations). Since the two polarisations of a single optical pulse must have the same natural frequency, they are in $1 : 1$ resonance. An S^1 phase invariance of the Hamiltonian for the interaction of the optical pulse with the optical medium in which it propagates will reduce the phase space to the Poincaré sphere, S^2, on which the problem is completely integrable. In this section, the fixed points and bifurcations of the phase portrait on S^2 of this system are studied as the beam intensity and medium parameters are varied. This study reveals bifurcations involving the creation or destruction of homoclinic and heteroclinic connections. For the complete discussion, see [DaHoTr1990].

Our approach to this problem uses the Stokes description [St1852, BoWo1965] of polarisation dynamics. In this approach Hamiltonian methods are used to reduce the four-dimensional phase space $\mathbb{C}^2$ (the two-component, complex-vector electric field amplitude) for the travelling-wave dynamics to a spherical surface S^2 (the Poincaré sphere). Bifurcations of the phase portrait on S^2 may then be determined as functions of material properties and beam intensity. The formation of homoclinic and heteroclinic orbits connecting hyperbolic fixed points may also be identified. These homoclinic and heteroclinic orbits are ***separatrices*** (i.e., stable and unstable manifolds of hyperbolic fixed points) which separate regions on S^2 with different types of periodic behaviour in the travelling-wave frame. The bifurcations of these fixed points signal the onset of a change in behaviour of the optical pulse in different parameter regimes, due to the nonlinearity of its interaction with the optical medium.

Remark 4.5.1 (Utility of the Poincaré sphere)

- *The Poincaré sphere is still useful in both linear and nonlinear optics today. For example, it is used in visualising polarisation dynamics for telecommunications using travelling-wave pulses in optical fibres [DaHoTr1990]. In this application, a given state of polarisation is represented by a unit vector (Stokes vector) on the Poincaré sphere, normalised by its intensity R. It is conventional in optics applications to permute the labels of the Stokes axes as $(1,2,3) \to (3,1,2)$. This cyclic permutation of the order of the Pauli matrices in (4.1.16) will cause no confusion, provided the conventions are checked carefully when making explicit applications of formulas based on the Poincaré sphere. See, for example, Section 4.5.*

- *After the permutation of the order of the Pauli matrices in (4.1.16),* ***linear polarisation*** *is represented on the equator of the Poincaré sphere, while right (resp. left)* ***circular polarisation*** *is located at its North (resp. South) pole. Super-positions of these polarisations generate* ***elliptical polarisations****, which are uniquely represented by other directions on the Poincaré sphere (4.1.22). Orthogonal polarisations are diametrically opposite across the sphere. Polarisation dynamics is then strikingly represented as motion along a curve on the Poincaré sphere.*

- *The surface of the Poincaré sphere corresponds to completely polarised light. Its interior corresponds to partially polarised, or partially incoherent, light.*

4.5.2 Hamiltonian formulation

Propagation of an optical travelling-wave pulse in a cubically nonlinear medium is described by the following Hamiltonian system of equations [Sh1984, Bl1965].

$$\frac{i}{2}\,\frac{da_j}{d\tau} = \chi^{(1)}_{jk}\,a_k + 3\chi^{(3)}_{jklm}a_k a_l a^*_m\,. \tag{4.5.1}$$

Its explicit Hamiltonian form is[3]

$$\begin{aligned} \frac{da_j}{d\tau} &= \{a_j,\, H\} = -\,2i\,\frac{\partial H}{\partial a^*_j}\,, && (4.5.2)\\ H &= a^*_j\chi^{(1)}_{jk}\,a_k + \frac{3}{2}a^*_j a_k \chi^{(3)}_{jklm} a_l a^*_m\,, && (4.5.3) \end{aligned}$$

where τ is the independent variable for travelling waves, the indices j, k, l, m take values $1, 2$ for the two polarisations and the complex two-vector $\mathbf{a} = (a_1, a_2)^T \in \mathbb{C}^2$ represents the electric field amplitude.

Remark 4.5.2 *Although we will follow only the formulation of optical pulse propagation, we note that the Hamiltonian (4.5.3) also includes the well known* ***Darling-Dennison Hamiltonian*** *for molecular dynamics [AlEb1975]. In the molecular context, the two modes of oscillation are called symmetric and antisymmetric stretch.*

Remark 4.5.3 (Complex susceptibility tensors)
The complex susceptibility tensors $\chi^{(1)}_{jk}$ and $\chi^{(3)}_{jklm}$ in equation (4.5.3) parameterise the linear and nonlinear polarisability of the medium, respectively. The susceptibility tensors are taken to be constant and Hermitian

[3] The optics convention $(1, 2, 3) \to (3, 1, 2)$ is applied in labelling the Stokes axes and Pauli matrices in this section.

in each $\mathbf{a} - \mathbf{a}^*$ *pair, as required for the Hamiltonian* H *to take real values. In addition,* $\chi^{(3)}_{jklm}$ *possesses a permutation symmetry arising from relabelling of indices:*

$$\chi^{(1)}_{jk} = \chi^{(1)*}_{\;\;kj}\,, \quad \chi^{(3)}_{jklm} = \chi^{(3)*}_{\;\;kjml}\,, \quad \chi^{(3)}_{jklm} = \chi^{(3)}_{lkjm} = \chi^{(3)}_{jmlk} = \chi^{(3)}_{lmjk}\,.$$

Besides the Hamiltonian, the intensity,

$$R = |\mathbf{a}|^2 = |a_1|^2 + |a_2|^2\,,$$

is also conserved because of the $1:1$ *resonance* S^1 *symmetry of the Hamiltonian* H.

4.5.3 Stokes vectors in polarisation optics

Following [DaHoTr1990] for this problem, we introduce the three-component Stokes vector, $\mathbf{u}$, given by

$$\mathbf{u} = a_j^* \boldsymbol{\sigma}_{jk} a_k = \operatorname{tr} Q\boldsymbol{\sigma}\,, \tag{4.5.4}$$

with $\boldsymbol{\sigma} = (\sigma_3, \sigma_1, \sigma_2)$ the standard Pauli spin matrices in equation (4.1.16) but in ***permuted order***, $(1,2,3) \to (3,1,2)$. As described in Section 4.1.4, the order $(3,1,2)$ is important in the identification of the types of optical polarisations, linear, circular and elliptical.

The transformation to Stokes vectors is carried out using the optical decomposition (4.5.4) of the coherency matrix, as follows.

$$\begin{aligned}
H &= \operatorname{tr}(Q\chi^{(1)}) + \frac{3}{2}\operatorname{tr}(Q\cdot\chi^{(3)}\cdot Q) \\
&= \frac{1}{2}\operatorname{tr}\Big(\big(R\sigma_0 + \mathbf{u}\cdot\boldsymbol{\sigma}\big)\chi^{(1)}\Big) \\
&\quad + \frac{3}{8}\operatorname{tr}\Big(\big(R\sigma_0 + \mathbf{u}\cdot\boldsymbol{\sigma}\big)\cdot\chi^{(3)}\cdot\big(R\sigma_0 + \mathbf{u}\cdot\boldsymbol{\sigma}\big)\Big) \\
&= \frac{R}{2}\operatorname{tr}\big(\chi^{(1)}\big) + \frac{1}{2}\mathbf{u}\cdot\operatorname{tr}\big(\boldsymbol{\sigma}\chi^{(1)}\big) + \frac{3R^2}{8}\operatorname{tr}\big(\sigma_0\cdot\chi^{(3)}\cdot\sigma_0\big) \\
&\quad + \frac{3R}{4}\mathbf{u}\cdot\operatorname{tr}\big(\boldsymbol{\sigma}\cdot\chi^{(3)}\cdot\sigma_0\big) + \frac{3}{8}\mathbf{u}\cdot\operatorname{tr}\big(\boldsymbol{\sigma}\cdot\chi^{(3)}\cdot\boldsymbol{\sigma}\big)\cdot\mathbf{u}\,.
\end{aligned}$$

Thus, in terms of the Stokes parameters $\mathbf{u}$ the Hamiltonian function H in equation (4.5.4) may be rewritten as

$$H = \mathbf{b} \cdot \mathbf{u} + \frac{1}{2}\mathbf{u} \cdot W \cdot \mathbf{u} \quad \text{with} \quad \mathbf{b} = \mathbf{f} + |\mathbf{u}|\mathbf{c} = \mathbf{f} + R\mathbf{c}\,. \qquad (4.5.5)$$

Geometrically, a level set of H in $\mathbb{R}^3$ is an offset ellipsoid (or hyperboloid for negative eigenvalues of W), whose centre has been shifted by $\mathbf{u} \to \mathbf{u} + W^{-1}\mathbf{b}$. Terms in $|\mathbf{u}| = R$ and $|\mathbf{u}|^2 = R^2$ may be treated as constants in H because they will Poisson commute with the Stokes vector $\mathbf{u}$ and, thus, will not contribute to its dynamics. In this transformed Hamiltonian, the constant vectors $\mathbf{f}$ and $\mathbf{c}$, and the constant symmetric tensor W are given by

$$\begin{aligned}
\mathbf{f} &= \frac{1}{2}\,\mathrm{tr}\left(\boldsymbol{\sigma}\chi^{(1)}\right) = \frac{1}{2}\,\boldsymbol{\sigma}_{kj}\chi^{(1)}_{jk}\,,\\
\mathbf{c} &= \frac{3}{4}\,\mathrm{tr}\left(\boldsymbol{\sigma}\cdot\chi^{(3)}\cdot\sigma_0\right) = \frac{3}{4}\,\boldsymbol{\sigma}_{kj}\chi^{(3)}_{jkll}\,,\\
W &= \frac{3}{4}\,\mathrm{tr}\left(\boldsymbol{\sigma}\cdot\chi^{(3)}\cdot\boldsymbol{\sigma}\right) = \frac{3}{4}\,\boldsymbol{\sigma}_{kj}\chi^{(3)}_{jklm}\boldsymbol{\sigma}_{ml}\,.
\end{aligned} \qquad (4.5.6)$$

Remark 4.5.4 (Stokes material parameters)
The material parameters $\mathbf{f}$, $\mathbf{c}$, and W are all real. According to equation (4.5.6), the parameters $\mathbf{f}$ and $\mathbf{c}$ represent the effects of anisotropy in the linear and nonlinear polarisability, respectively. They lead to precession of the Stokes vector $\mathbf{u}$ with (vector) frequency

$$\mathbf{b} = \mathbf{f} + R\mathbf{c}\,,$$

which is a function of intensity R.

- *In the low intensity case, one ignores $\chi^{(3)}$. Then the precession rate of the Stokes vector due to anisotropy of the linear polarisability (linear birefringence) is $\mathbf{f}$, which is constant and independent of intensity. This effect is called **Faraday rotation** of the polarisation vector. In some materials, Faraday rotation may be enhanced by applying an external magnetic field, in which case it is called a **magneto-optical** effect.*

- *The nonlinear effects of* $\chi^{(3)}$ *are parameterised by* $\mathbf{c}$ *and* W *in the Hamiltonian (4.5.5). The tensor* W *is symmetric, so a polarisation basis may always be assumed in which* W *is diagonal,* $W = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$*, in analogy to the principal moments of inertia of a rigid body. However, unlike the rigid body, the entries of* W *need not all be positive.*

The equations of motion for the Stokes vector $\mathbf{u}$ may be expressed in Hamiltonian form as

$$\frac{d\mathbf{u}}{d\tau} = \{\mathbf{u}, H\} \quad \text{with} \quad \{F,\, H\} = 4\mathbf{u} \cdot \nabla F(\mathbf{u}) \times \nabla H(\mathbf{u})\,. \tag{4.5.7}$$

This is written in triple scalar product form, just as in the case of the rigid body, but with an inessential factor of (-4). As expected, the intensity $R = |\mathbf{u}|$ is the Casimir function for this $\mathbb{R}^3$ Poisson bracket. That is, the intensity R Poisson-commutes with all functions of $\mathbf{u}$ in the applications of the Poisson bracket (4.5.7). Consequently, the intensity R in the Stokes description of lossless polarised optical beam dynamics may be regarded simply as a constant parameter. However, the intensity R is also an experimental control parameter which may be varied in the study of bifurcations and stability of travelling-wave polarisation states.

The travelling-wave equation (4.5.1) for polarisation dynamics thus becomes

$$\frac{d\mathbf{u}}{d\tau} = \{\mathbf{u},\, H\} = (\mathbf{b} + W \cdot \mathbf{u}) \times \mathbf{u} \quad \text{with} \quad \mathbf{b} = \mathbf{f} + R\mathbf{c}\,, \tag{4.5.8}$$

in the Stokes vector representation.

4.5.4 Further reduction to the Poincaré sphere

The system of polarisation vector equations (4.5.8) reduces further to the Poincaré sphere S^2 of radius R upon transforming to spherical coordinates, defined by

$$(u_1, u_2, u_3) = (R\sin\theta\sin\phi, R\cos\theta, R\sin\theta\cos\phi)\,, \tag{4.5.9}$$

written in the nonlinear optics convention with polar angle measured from the 2-axis. In these coordinates, the symplectic Poisson bracket on S^2 is expressible as in equation (4.2.16) modulo an overall sign,

$$\big\{F,\,H\big\} = \frac{\partial F}{\partial\cos\theta}\frac{\partial H}{\partial\phi} - \frac{\partial F}{\partial\phi}\frac{\partial H}{\partial\cos\theta}\,, \tag{4.5.10}$$

and the reduced Hamiltonian function (4.5.5) is

$$\begin{aligned} H \;&=\; \frac{1}{2}\,R^2\Big[\big(\lambda_1\sin^2\phi + \lambda_3\cos^2\phi\big)\sin^2\theta + \lambda_2\cos^2\theta\Big] \\ &\qquad + R\sin\theta\big(b_1\sin\phi + b_3\cos\phi\big) + b_2 R\cos\theta\,. \end{aligned} \tag{4.5.11}$$

When $\mathbf{b} = (b_1, b_2, b_3) = 0$, this expression reduces to the Hamiltonian for the rigid body restricted to a level surface of angular momentum, modulo the permuted coordinates axes $(1,2,3) \to (3,1,2)$ in (4.5.9) for the optics convention.

Example 4.5.5 (Pitchfork bifurcation for optical gyrostat)
The system (4.5.8) is solved easily when (i) two eigenvalues of W coincide, and (ii) one or more of the components of $\mathbf{b}$ vanish. This case recovers the symmetric Lagrange gyrostat, or rigid body with a flywheel discussed in Section A.1.4.

For definiteness, we set $W = \omega\,\mathrm{diag}\,(1,1,2)$ and $\mathbf{b} = (b_1, b_2, 0)$. Then equations (4.5.8) read

$$\begin{aligned} \frac{du_1}{d\tau} &= (b_2 - \omega u_2)u_3\,, \\ \frac{du_2}{d\tau} &= (\omega u_1 - b_1)u_3\,, \\ \frac{du_3}{d\tau} &= b_1 u_2 - b_2 u_1\,. \end{aligned}$$

*Hence, a **Duffing equation** emerges for u_3,*

$$\frac{d^2 u_3}{d\tau^2} = A u_3(B - u_3^2)\,, \tag{4.5.12}$$

with constant parameters,

$$A = \frac{1}{2}\omega^2\,,\quad B = \frac{2H}{\omega} - R^2 - \frac{2(b_1^2 + b_2^2)}{\omega^2}\,. \tag{4.5.13}$$

The other two components of $\mathbf{u}$ *may be determined algebraically from the two constants of motion* R *and* H*. When* B *increases through zero, the* ***Duffing equation*** *(4.5.12) develops a pair of orbits, homoclinic to the fixed point* u_3 *(see, e.g., [GuHo1983, Wi1988]). This is the pitchfork bifurcation. This particular case suffices to demonstrate that the system (4.5.8) possesses bifurcations in which homoclinic orbits are created. For more details and references, see Appendix A of [DaHoTr1990]. See also [HoKoSu1991, AcHoKoTi1997] for other applications of Hamiltonian bifurcations and related ideas in nonlinear optics.*

4.5.5 Bifurcation analysis

We now specialise to the case of a non-parity-invariant material with C_4 rotation symmetry about the axis of propagation (the z-axis), for which material constants take the form $W = (\lambda_1, \lambda_2, \lambda_3)$ and $\mathbf{b} = (0, b_2, 0)$. (See [DaHoTr1989, DaHoTr1990] for additional developments of what follows.)

Let us introduce the following parameters:

$$\mu = \lambda_3 - \lambda_1\,, \quad \lambda = \frac{\lambda_2 - \lambda_1}{\lambda_3 - \lambda_1}\,, \quad \beta = \frac{b_2}{R(\lambda_3 - \lambda_1)}\,. \tag{4.5.14}$$

In terms of these parameters, the Hamiltonian, Poisson bracket and equations of motion transform, as follows. First, the Hamiltonian (4.5.11) for $\mathbf{b} = (0, b_2, 0)$ reduces to

$$H = \frac{1}{2}\,\mu\big[(R^2 - u^2)\cos^2\phi + \lambda u^2 + 2\beta R u\big] + \frac{1}{2}\,\lambda_1 R^2\,, \tag{4.5.15}$$

with $u = R\cos\theta$. The canonical Poisson bracket corresponds to a symplectic form that is equal to the area element on the sphere,

$$\{F,\, H\} = \frac{\partial F}{\partial \phi}\frac{\partial H}{\partial u} - \frac{\partial H}{\partial \phi}\frac{\partial F}{\partial u}\,. \tag{4.5.16}$$

The canonical equations of motion on the sphere are then

$$\begin{aligned} \frac{d\phi}{d\tau} &= \{\phi,\, H\} = \frac{\partial H}{\partial u} = \mu[\beta R - (\cos^2\phi - \lambda)u]\,, && (4.5.17)\\ \frac{du}{d\tau} &= \{u,\, H\} = -\,\frac{\partial H}{\partial \phi} = \mu(R^2 - u^2)\cos\phi\sin\phi\,. && (4.5.18) \end{aligned}$$

We construct the phase portrait of the system and explain how this portrait changes as the parameters in the equations vary. Fixed points of (4.5.17) and (4.5.18) occur when the right-hand sides vanish. These points are easily located and classified. The locations and types of these fixed points are listed in Table 4.1, for $\mu \neq 0$.

Fixed point	Coordinates	Saddle	Centre
F, B	$\phi = 0,\ \pi$ $\cos\theta = \beta/(1-\lambda)$	$\lambda > 1$	$\lambda < 1$
L, R	$\phi = \pi/2,\ -\pi/2$ $\cos\theta = -\beta/\lambda$	$\lambda < 0$	$\lambda > 0$
N	$\cos^2\phi = \lambda + \beta$ $\theta = 0$	$\beta \in (-\lambda, 1-\lambda)$	$\beta \notin (-\lambda, 1-\lambda)$
S	$\cos^2\phi = \lambda - \beta$ $\theta = \pi$	$\beta \in (-\lambda, 1-\lambda)$	$\beta \notin (-\lambda, 1-\lambda)$

Table 4.1: The fixed points of the Hamiltonian system (4.5.17) and (4.5.18) change their types (that is, they bifurcate) as the parameters β and λ defined in (4.5.14) are varied for $\mu \neq 0$. The fixed points F,B (resp. L,R) are constrained by $\beta^2 < (1-\lambda)^2$ (resp. $\beta^2 < \lambda^2$).

A circle of fixed points when $\mu = 0$

The special case where $\mu = 0$, i.e., $\lambda_3 - \lambda_1 = 0$, requires a separate analysis. In that case, the right-hand side of equation (4.5.8) vanishes identically so that the set of fixed points of the system is the circle

$$\cos\theta = \frac{b_2}{R(\lambda_2 - \lambda_1)} = \frac{\beta}{\lambda}\,. \tag{4.5.19}$$

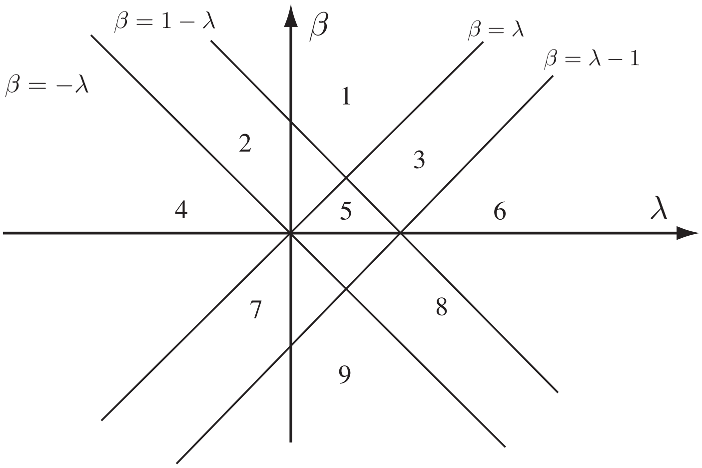


Figure 4.6: Pairs of of fixed points at the North and South poles appear or vanish as the lines $\beta = \pm(1 - \lambda)$ and $\beta = \pm\lambda$ are crossed in the (λ, β) parameter plane. The (λ, β) parameter plane is partitioned into nine distinct regions $R1 - R9$ separated by four critical lines that intersect in pairs at four points. Because of discrete symmetries it is sufficient to consider the quarter plane given by $\lambda < 1/2$ and $\beta > 0$, i.e., to restrict attention to regions $R1$, $R2$, $R4$, and $R5$. Compare with the bifurcations shown in Figure 4.7.

Two essential parameters (λ, β) when $\mu \neq 0$

The phase portrait depends on two essential parameters, λ and β, or equivalently, $\lambda_2 - \lambda_1$, and b_2/R. Bifurcations of the phase portrait occur when the inequality constraints in the third column of Table 4.1 become equalities; hence we observe that the pairs of fixed points (F, B) and (L, R) appear or vanish as the lines $\beta = \pm(1 - \lambda)$ and $\beta = \pm\lambda$ are crossed in the (λ, β) parameter plane (see Figure 4.6).

4.5.6 Nine regions in the (λ, β) parameter plane

The (λ, β) parameter plane is partitioned into nine distinct regions $R1 - R9$ separated by four critical lines that intersect in pairs at four

S_1 :	$\lambda < 0$	$R1 \leftrightarrow R2 \leftrightarrow R4 \leftrightarrow R7 \leftrightarrow R9$
S_2 :	$\lambda = 0$	$R1 \leftrightarrow R2 \leftrightarrow R7 \leftrightarrow R9$
S_3 :	$0 < \lambda < \frac{1}{2}$	$R1 \leftrightarrow R2 \leftrightarrow R4 \leftrightarrow R7 \leftrightarrow R9$
S_4 :	$\lambda = \frac{1}{2}$	$R1 \leftrightarrow R5 \leftrightarrow R9$
S_5 :	$\frac{1}{2} < \lambda < 1$	$R1 \leftrightarrow R3 \leftrightarrow R5 \leftrightarrow R8 \leftrightarrow R9$
S_6 :	$\lambda = 1$	$R1 \leftrightarrow R3 \leftrightarrow R8 \leftrightarrow R9$
S_7 :	$\lambda > 1$	$R1 \leftrightarrow R3 \leftrightarrow R6 \leftrightarrow R8 \leftrightarrow R9$

Table 4.2: A choreography of seven possible bifurcation sequences $S_1 - S_7$ are shown that may be executed along vertical lines in the (λ, β) plane of Figure 4.6 by varying the beam intensity R. The sensitivity of the equilibrium solutions for polarisation dynamics to the beam intensity could be of special relevance in telecom applications of polarised optical pulses.

points. Typical phase portraits corresponding to each of these regions are shown in Figure 4.7. The phase portraits of the system (4.5.18) and (4.5.17) are invariant under the following discrete symmetry transformations:

$$
\begin{aligned}
&\phi \to \phi \pm \pi\,; \\
&\phi \to \phi \pm \pi\,, \quad \theta \to \pi - \theta\,, \quad \beta \to -\beta\,; \\
&\phi \to \phi \pm \pi/2\,, \quad \lambda \to 1 - \lambda\,, \quad \beta \to -\beta\,; \\
&\phi \to \phi \pm \pi/2\,, \quad \lambda \to 1 - \lambda\,, \quad \theta \to \pi - \theta\,.
\end{aligned}
$$

Thus, as far as the configurations of critical orbits on the phase sphere are concerned, it is sufficient to consider the quarter plane given by $\lambda < 1/2$ and $\beta > 0$. That is, we may restrict attention to regions $R1$, $R2$, $R4$, and $R5$.

The rigid-body limit lies along the λ-axis

Along the λ-axis ($\beta = 0$ in the parameter plane) the set of fixed points does not change except at $\lambda = 0$ and $\lambda = 1$. That is, no bifurcations occur in general when the λ-axis is crossed. Nonetheless, this line is special. Indeed, in the interval $\lambda \in (0, 1)$, that is in region $R5$, both poles are hyperbolic; to each of them is attached to a

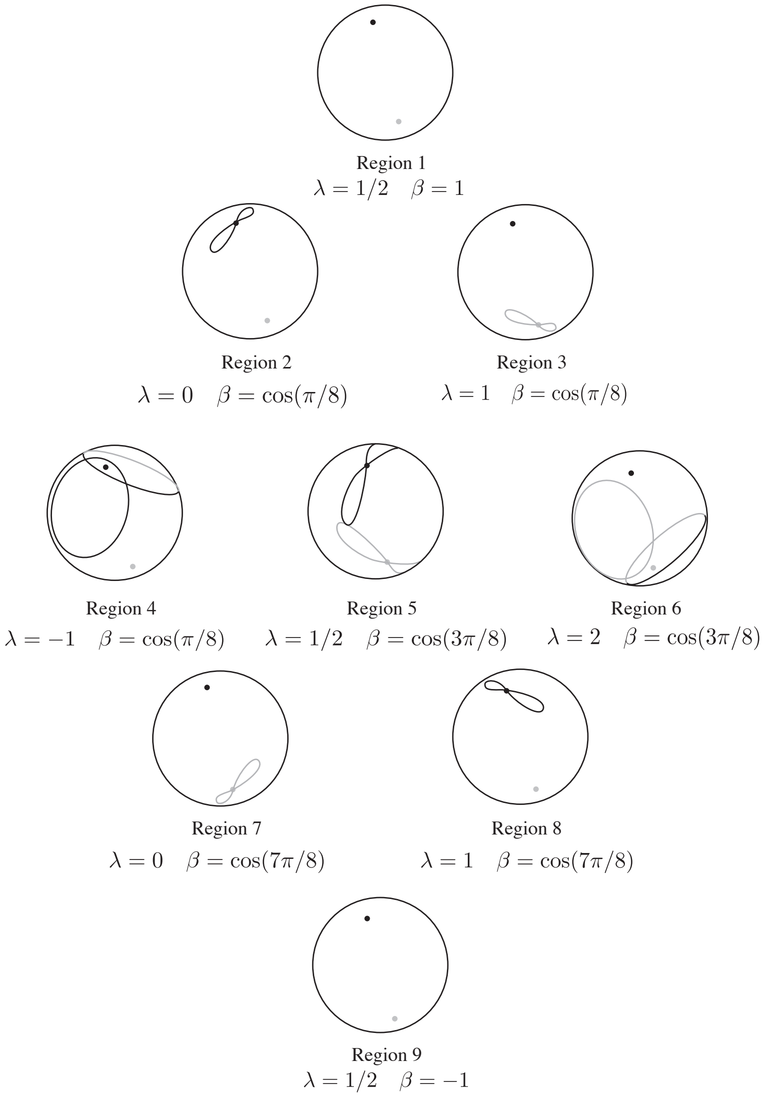


Figure 4.7: The choreography of bifurcations on the Poincaré sphere is shown for a polarised optical travelling wave in a cubically nonlinear lossless medium.

pair of homoclinic loops. When β vanishes, these homoclinic loops merge together. For $\beta = 0$ and $0 < \lambda < 1$, these merged homoclinic loops form four heteroclinic lines that connect the North pole with the South pole. This is the familiar phase portrait on the angular momentum sphere for the rigid body with three unequal moments of inertia.

On the λ-axis when $\beta = 0$, polarisation dynamics on S^2 reduces *exactly* to rigid-body dynamics. In that case, the phase portrait consists of the poles N and S, and the four other fixed points are located on the equator of the Poincaré sphere. (This configuration of fixed four points distributed symmetrically on the equator is obtained only on the λ-axis.) Two of these, (N, S) or (F, B) or (R, L), are unstable while the other four are stable. Which pair of critical points is unstable is decided by the value of $\lambda = (\lambda_2 - \lambda_1)(\lambda_3 - \lambda_1)$. The pair (F, B) is hyperbolic when $0 < \lambda < 1$, and the pair (R, L) is hyperbolic whenever $\lambda > 1$. In each of these cases, the unstable direction is specified by the λ_k, whose value is intermediate among the three.

Remark 4.5.6 *By representing the orbits of constant angular momentum as a flow on the Poincaré sphere, one may see at a glance the qualitative motion for any set of initial conditions. This visualisation also allows one to characterise the stability properties of the flow geometrically as either elliptic (stable) or hyperbolic (unstable). One may also see how these solutions bifurcate as the parameters are varied. Thus, one visualises the equilibrium points, the homoclinic orbits and the bifurcations for the problem considered, all on a single figure.*

Bifurcations that depend on beam intensity R

As the beam intensity R is varied, a vertical line is traced in the parameter plane along which various bifurcation sequences may occur. A ***choreography*** of the seven possible bifurcation sequences is shown in Figure 4.2. The figures are plotted as functions of the bifurcation parameters $\lambda = (\lambda_2 - \lambda_1)/(\lambda_3 - \lambda_1)$ (horizontal) and $\beta = (b_2/R)/(\lambda_3 - \lambda_1)$ (vertical). Compare with Figure 4.6. See [DaHoTr1990] for more information about the bifurcations that may take place on the Poincaré sphere when beam intensity R is varied, on tracking along these vertical lines in the (λ, β) parameter plane.

Chapter 5

Elastic spherical pendulum

5.1 Introduction and problem formulation

This chapter discusses the swinging spring, or elastic spherical pendulum, which has a 1:1:2 resonance arising at cubic order in its phase-averaged Lagrangian. The corresponding modulation equations turn out to be the famous three-wave equations that also apply, for example, in laser-matter interaction in a cavity. Thus, analysis of the elastic spherical pendulum combines many of the ideas we have developed so far, and it suggests a strategy for solving even more examples.

The chapter formulates and analyses the equations for the motion of an elastic spherical pendulum, as follows.

- Identify degrees of freedom.
- Approximate the Lagrangian for small excitations at cubic order in nonlinearity.
- Average the Lagrangian over its oscillating phases. (The averaging process introduces another S^1 phase symmetry.)
- Introduce an additional S^1 1:1:2 resonance symmetry by phase averaging that allows integration of the corresponding Euler-Lagrange equations.

- Derive the resulting three-wave equations and study their solution behaviour.

5.1.1 Problem statement, approach and results

The elastic pendulum or swinging spring is a simple mechanical system that exhibits rich dynamics. It consists of a heavy mass suspended from a fixed point by a light spring which can stretch but cannot bend, moving under gravity. In this chapter, we investigate the 1:1:2 resonance dynamics of this system in three dimensions. Its characteristic feature – the regular step-wise precession of its azimuthal angle – is also discussed, following [HoLy2002].

When the Lagrangian for the elastic pendulum is approximated to cubic order and then averaged over its fast dynamics, the resulting modulation equations for the three degrees of freedom have three independent constants of motion and thus they are completely integrable. Integrability arises because the averaging process introduces an additional S^1 symmetry. These modulation equations turn out to be identical to the three-wave equations for the $1:1:2$ resonant triad interactions that also occur in fluids and plasmas, and in laser-matter interaction, as discussed in Chapter 6.

The averaged system is reduced to a form amenable to analytical solution and the full solution – including the phase – is reconstructed. The geometry of the solutions in phase space is examined and used to classify the available motions. As might be expected from the analysis of the $1:2$ resonance in Chapter 4, the geometry of the orbit manifolds for the $1:1:2$ resonant approximation of the elastic pendulum has a singular point. At this singular point, the azimuthal angle of the pendulum is undefined, which leads to the regular step-wise precession seen in its swing-spring motions nearby.

5.1.2 History of the problem

The first comprehensive analysis of the elastic pendulum appeared in [ViGo1933]. These authors were inspired by the analogy between this system and the $1:2$ Fermi resonance in a carbon-dioxide molecule between its bending and vibrating modes. This chapter is also connected with other physical systems of current interest. For example, the modulation equations for the averaged motion of the swinging spring may be transformed into the equations for three-wave interactions.

These three-wave equations also appear in analysing fluid and plasma systems, and in laser-matter interaction. For example, the three complex equations in this system are identical to the Maxwell-Schrödinger envelope equations for the interaction between radiation and a two-level resonant medium in a microwave cavity [HoKo1992]. The three-wave equations also govern the envelope dynamics of light waves in an inhomogeneous material [DaHoTr1990, AlLuMaRo1998, AlLuMaRo1999]. For the special case where the Hamiltonian takes the value zero, the equations reduce to Euler's equations for a freely rotating rigid body. Finally, the equations are also equivalent to a complex (unforced and undamped) version of the Lorenz [Lo1963] three-component model, which has been the subject of many studies [Sp1982].

Thus, the simple spring pendulum, whose planar dynamics were first studied as a classical analogue of the quantum phenomenon of Fermi resonance in the CO_2 molecule, now provides a concrete mechanical system which simulates a wide range of physical phenomena that become tractable by using methods of geometric dynamics. In turn, the geometric analysis of the classical singularities of the spring pendulum has recently produced a new insight into the corresponding quantum behaviour of the CO_2 molecule [Cu-etal-2004]. Namely, the most important features in the quantisation of a classical system may occur as a result of its singularities.

5.2 Equations of motion

5.2.1 Approaches of Newton, Lagrange & Hamilton

The physical system under investigation is an elastic pendulum, or swinging spring, consisting of a heavy mass m suspended from a fixed point by a light spring which can stretch but not bend. The mass swings as a pendulum under the force of gravity with constant acceleration g and it oscillates under the restoring force of the spring. The unstretched spring has length ℓ_0 and its spring constant is denoted as k. In coordinates centred at the point of support, the mass at the end of the spring has position $\mathbf{X} = (X, Y, Z) \in \mathbb{R}^3$. Thus, the configuration space for this problem is $\mathbb{R}^3$.

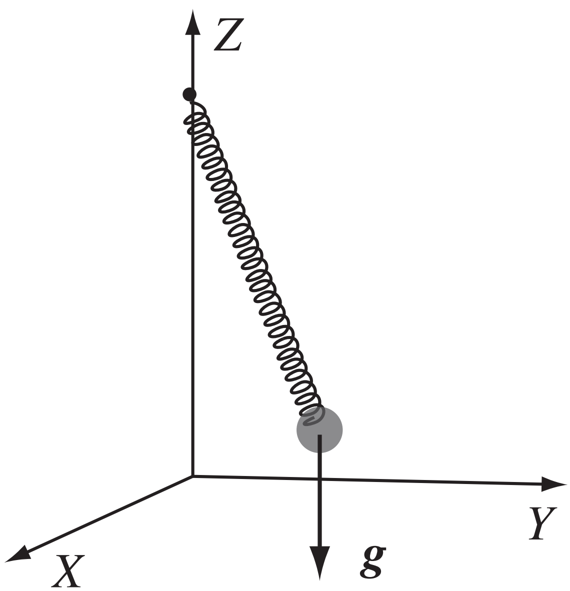


Figure 5.1: Schematic diagram of the elastic pendulum, or swinging spring. Cartesian coordinates $\mathbf{X} = (X, Y, Z) \in \mathbb{R}^3$ centred at the position of equilibrium are shown.

Newton's 2nd Law approach: $m\mathbf{a} = \mathbf{F}$

The sum of the forces on the pendulum bob appear in Newton's Law as

$$m\mathbf{a} = m\ddot{\mathbf{X}} = \underbrace{-\,mg\,\mathbf{e}_3}_{\text{Gravity}} + \underbrace{(-\,k)\Big(1 - \frac{\ell_0}{|\mathbf{X}|}\Big)\mathbf{X}}_{\text{Elastic restoring force}} = \mathbf{F}\,. \qquad (5.2.1)$$

The gravitational force is vertically downward and the restoring force of the spring points along the vector from the centre of support to the position of the mass at any time.

> **Exercise.** At the downward equilibrium, show that the pendulum length is $\ell = \ell_0 + mg/k$. ★

The Lagrangian Approach

The motion equations (5.2.1) for the elastic pendulum also arise from Hamilton's principle $\delta S = 0$ with $S = \int L\, dt$ for the Lagrangian $L : T\mathbb{R}^3 \mapsto \mathbb{R}$ with coordinates $(\mathbf{X},\, \dot{\mathbf{X}})$ given by

$$\begin{aligned} L(\mathbf{X},\, \dot{\mathbf{X}}) &= T - V_{grav} - V_{osc} \\ &= \frac{m}{2}|\dot{\mathbf{X}}|^2 - g\, \mathbf{e}_3 \cdot \mathbf{X} - \frac{k}{2}\big(|\mathbf{X}| - \ell_0\big)^2 , \end{aligned} \tag{5.2.2}$$

where $\mathbf{e}_3 \cdot \mathbf{X} = \langle\, \mathbf{e}_3\, ,\, \mathbf{X}\,\rangle$ is the standard inner product on $\mathbb{R}^3$ between the position vector $\mathbf{X}$ and the vertical unit vector $\mathbf{e}_3$. This Lagrangian has partial derivatives

$$\begin{aligned} \frac{\partial L}{\partial \dot{\mathbf{X}}} &= m\dot{\mathbf{X}} =: \mathbf{P}\, , \\ \frac{\partial L}{\partial \mathbf{X}} &= -\, mg\mathbf{e}_3 - k\Big(1 - \frac{\ell_0}{|\mathbf{X}|}\Big)\mathbf{X}\, . \end{aligned}$$

This Lagrangian is certainly regular: its canonical momentum $\mathbf{P} = \partial L/\partial\dot{\mathbf{X}}$ is linear in the velocity $\dot{\mathbf{X}}$. Newton's equations of motion for the elastic pendulum may now be recovered in Euler-Lagrange form,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{X}}} = \frac{\partial L}{\partial \mathbf{X}}\, . \tag{5.2.3}$$

These may be expressed in Newtonian form as

$$m\ddot{\mathbf{X}} = -\, mg\mathbf{e}_3 - k\Big(1 - \frac{\ell_0}{|\mathbf{X}|}\Big)\mathbf{X}\, , \tag{5.2.4}$$

and they are equivalent to the following equations in canonically conjugate variables:

$$\dot{\mathbf{X}} = \mathbf{P}/m \quad \text{and} \quad \dot{\mathbf{P}} = -\, mg\mathbf{e}_3 - k\Big(1 - \frac{\ell_0}{|\mathbf{X}|}\Big)\mathbf{X}\, . \tag{5.2.5}$$

Thus, as expected for regular Lagrangians, Newton's equations and the Euler-Lagrange equations are equivalent. Alternatively, a direct computation with Hamilton's principle again yields Newton's equations for this problem as,

$$\delta S = 0 = -\int_{t_1}^{t_2} \Big(m\ddot{\mathbf{X}} + mg\mathbf{e}_3 + k\Big(1 - \frac{\ell_0}{|\mathbf{X}|}\Big)\mathbf{X}\Big) \cdot \delta\mathbf{X}\, dt\,, \tag{5.2.6}$$

upon applying integration by parts and using the condition that variation $\delta\mathbf{X}$ vanishes at the endpoints in time.

> **Exercise.** Show using ***Noether's theorem*** that invariance of the Lagrangian in equation (5.2.2) under the infinitesimal action of S^1 rotations about the vertical axis
>
> $$\delta\mathbf{X} = -\,\mathbf{e}_3 \times \mathbf{X}$$
>
> yields conservation of the vertical angular momentum,
>
> $$J_3 = \mathbf{e}_3 \cdot m\dot{\mathbf{X}} \times \mathbf{X}.$$
>
> ★

Conservation of azimuthal angular momentum

As in solving for the motions of the bead sliding on the rotating hoop and the spherical pendulum, it is convenient in this case to use spherical coordinates with azimuthal angle $0 \le \phi < 2\pi$ and polar angle $0 \le \theta < \pi$ measured from the ***downward*** vertical. The Lagrangian (5.2.2) for the elastic spherical pendulum in these coordinates is:

$$\begin{aligned} L(R,\, \dot{R},\, \theta,\, \dot{\theta},\, \dot{\phi}) \quad &= \quad \frac{m}{2}(\dot{R}^2 + R^2\dot{\theta}^2 + R^2\dot{\phi}^2 \sin^2\theta) \\ &\quad - \frac{k}{2}(R - \ell_0)^2 + mgR\cos\theta\,. \end{aligned} \tag{5.2.7}$$

Its corresponding Euler-Lagrange equations are:

$$\begin{aligned}
\ddot{R} &= -\frac{k}{m}(R-\ell_0) + g\cos\theta + R\dot{\theta}^2 + R\dot{\phi}^2\sin^2\theta\,,\\
\frac{d}{dt}(R^2\dot{\theta}) &= R^2\dot{\phi}^2\sin\theta\cos\theta - gR\sin\theta\,,\\
\frac{d}{dt}(R^2\dot{\phi}\sin^2\theta) &= 0\,.
\end{aligned} \tag{5.2.8}$$

Thus, as with the spherical pendulum, azimuthal symmetry of the Lagrangian $L(R,\,\dot{R},\,\theta,\,\dot{\theta},\,\dot{\phi})$ (that is, L being independent of ϕ) implies conservation of azimuthal angular momentum.

Equilibria The stable equilibria are associated with ***conical motion***, with R, θ and $\dot{\phi} = \omega$ all constant. At equilibrium, the previous equations of motion reduce to

$$\begin{aligned}
0 &= -\frac{k}{m}(R-\ell_0) + g\cos\theta + R\omega^2\sin^2\theta\,,\\
0 &= R^2\omega^2\sin\theta\cos\theta - gR\sin\theta\,,\\
R^2\omega\sin^2\theta &= h\,,
\end{aligned} \tag{5.2.9}$$

where h is the conserved azimuthal angular momentum. The first two equilibrium relations determine the spring length in terms of the cone angle,

$$(R-\ell_0)\omega_Z^2\cos\theta = g\,,\quad \text{with}\quad \omega_Z^2 = \frac{k}{m}\,, \tag{5.2.10}$$

which is similar to the equilibrium relation (A.1.23) for the bead on the rotating hoop.

> **Exercise.** For finite spring length, show that the conical equilibrium must always lie below the horizontal plane $\cos\theta = 0$. ★

The Hamiltonian Approach

The Legendre transformation of the Lagrangian in (5.2.2) yields the Hamiltonian $H : T^*\mathbb{R}^3 \mapsto \mathbb{R}$ with canonical coordinates $(\mathbf{P},\mathbf{X})$

given by,

$$
\begin{aligned}
H(\mathbf{P}, \mathbf{X}) &= \mathbf{P} \cdot \dot{\mathbf{X}} - L \\
&= \frac{1}{2m}|\mathbf{P}|^2 + g\mathbf{e}_3 \cdot \mathbf{X} + \frac{k}{2}\big(|\mathbf{X}| - \ell_0\big)^2 . \qquad (5.2.11)
\end{aligned}
$$

The standard symplectic form on $T^*\mathbb{R}^3$ is

$$
\omega = dX^j \wedge dP_j \,.
$$

The corresponding canonical Poisson brackets are

$$
\{F,\, H\} = \frac{\partial F}{\partial \mathbf{X}} \cdot \frac{\partial H}{\partial \mathbf{P}} - \frac{\partial H}{\partial \mathbf{X}} \cdot \frac{\partial F}{\partial \mathbf{P}} \,. \qquad (5.2.12)
$$

These yield Hamilton's equations for the Hamiltonian H in (5.2.11),

$$
\begin{aligned}
\dot{\mathbf{X}} = \{\mathbf{X},\, H\} &= \frac{\partial H}{\partial \mathbf{P}} = \mathbf{P}/m \,, \qquad (5.2.13) \\
\dot{\mathbf{P}} = \{\mathbf{P},\, H\} &= -\frac{\partial H}{\partial \mathbf{X}} \\
&= -mg\mathbf{e}_3 - k\Big(1 - \frac{\ell_0}{|\mathbf{X}|}\Big)\mathbf{X} \,, \qquad (5.2.14)
\end{aligned}
$$

which again are equivalent to Newton's equations (5.2.1) for the elastic pendulum.

Proposition 5.2.1 (Conservation of energy)
Having no explicit time dependence, the Hamiltonian (5.2.11) is conserved.

Proof.

$$
\dot{H} = \{H,\, H\} = 0 \,,
$$

which holds by skew symmetry of the Poisson bracket. ∎

Proposition 5.2.2 (Conserved vertical angular momentum) *The motion equations (5.2.1) preserve the vertical component of angular momentum,*

$$
J_3 = \mathbf{e}_3 \cdot \mathbf{P} \times \mathbf{X} = \mathbf{e}_3 \cdot \mathbf{J} \,. \qquad (5.2.15)
$$

Proof. By direct calculation from Newton's equations

$$\frac{d}{dt}(\mathbf{P}\times\mathbf{X}) = -\,mg\,\mathbf{e}_3\times\mathbf{X} \quad\Longrightarrow\quad \frac{d}{dt}(\mathbf{e}_3\cdot\mathbf{P}\times\mathbf{X}) = 0\,.$$

∎

Proposition 5.2.3 (S^1 symmetry)
The system $(H,\,T^*\mathbb{R}^3,\,\omega)$ possesses an S^1 symmetry given by

$$S^1\times\,T^*\mathbb{R}^3 \mapsto T^*\mathbb{R}^3 : \;(\phi,(\mathbf{Q},\,\mathbf{P})) \rightarrow (R_\phi^{-1}\mathbf{Q},\,R_\phi^{-1}\mathbf{P})\,, \tag{5.2.16}$$

where

$$R_\phi = \begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5.2.17}$$

is the matrix of rotation about axis $\mathbf{e}_3$ by angle ϕ.

Proof. The Poisson bracket and the Hamiltonian are invariant under this rotation. ∎

Proposition 5.2.4 (X_{J_3} generates the S^1 symmetry)
The Hamiltonian vector field X_{J_3} generates the S^1 symmetry of the Hamiltonian.

Proof. By direct computation

$$\begin{aligned}\frac{d}{d\phi}\Big|_{\phi=0}(R_\phi^{-1}\mathbf{X},\,R_\phi^{-1}\mathbf{P}) &= (\mathbf{X}\times\mathbf{e}_3,\,\mathbf{P}\times\mathbf{e}_3) \\ &= (\{\,\mathbf{X}\,,\,J_3\},\,\{\,\mathbf{P}\,,\,J_3\}) = (X_{J_3}\mathbf{X},\,X_{J_3}\mathbf{P})\,.\end{aligned}$$

Here, X_{J_3} is the Hamiltonian vector field for the S^1 symmetry with diagonal action, $(\phi,(\mathbf{Q},\,\mathbf{P}))\rightarrow(R_\phi^{-1}\mathbf{Q},\,R_\phi^{-1}\mathbf{P})$. ∎

Proposition 5.2.5 (The bracket $X_{J_3}H = \{H\,,J_3\,\}$ vanishes)

Proof. The action on H of the Hamiltonian vector field X_{J_3} for $J_3 = \mathbf{e}_3 \cdot \mathbf{J} = \mathbf{e}_3 \cdot \mathbf{P} \times \mathbf{X}$ is obtained from the canonical Poisson bracket by setting

$$
\begin{aligned}
X_{J_3} H = \{H\,, J_3\,\} &= \{\mathbf{X}\,, J_3\,\} \cdot \frac{\partial H}{\partial \mathbf{X}} + \{\mathbf{P}\,, J_3\,\} \cdot \frac{\partial H}{\partial \mathbf{P}} \\
&= \mathbf{X} \times \mathbf{e}_3 \, \cdot \, \frac{\partial H}{\partial \mathbf{X}} + \mathbf{P} \times \mathbf{e}_3 \, \cdot \, \frac{\partial H}{\partial \mathbf{P}} \\
&= \mathbf{X} \times \mathbf{e}_3 \, \cdot \, \Big(mg\, \mathbf{e}_3 + k\Big(1 - \frac{\ell_0}{|\mathbf{X}|}\Big)\mathbf{X}\Big) \\
&\quad + \mathbf{P} \times \mathbf{e}_3 \, \cdot \, \mathbf{P} \\
&= 0\,,
\end{aligned}
$$

so J_3 is conserved, as expected. ■

Remark 5.2.6 *The conservation law $\dot{J}_3 = \{J_3\,,\, H\} = 0$ for $J_3 = \mathbf{e}_3 \cdot \mathbf{P} \times \mathbf{X}$ was expected, since the Hamiltonian vector field X_{J_3} given by*

$$
\begin{aligned}
X_{J_3} = \{\,\cdot\,,\, J_3\,\} &= \frac{\partial J_3}{\partial \mathbf{P}} \cdot \frac{\partial}{\partial \mathbf{X}} - \frac{\partial J_3}{\partial \mathbf{X}} \cdot \frac{\partial}{\partial \mathbf{P}} \\
&= \mathbf{X} \times \mathbf{e}_3 \cdot \frac{\partial}{\partial \mathbf{X}} + \mathbf{P} \times \mathbf{e}_3 \cdot \frac{\partial}{\partial \mathbf{P}}\,,
\end{aligned}
$$

generates rotations of both $\mathbf{X}$ and $\mathbf{P}$ by the same angle ϕ about the vertical axis $\mathbf{e}_3$, and the Hamiltonian in (5.2.11) is invariant under such rotations.

Hamilton's approach for the elastic pendulum in spherical coordinates

The Legendre transform of the Lagrangian in spherical coordinates is

$$
\pi(\dot{R}, R; \dot{\theta}, \theta; \dot{\phi}, \phi) \to \big(P, R\,;\, P_\theta, \theta\,;\, P_\phi, \phi\big) = \Big(\frac{\partial L}{\partial \dot{R}}, R\,;\, \frac{\partial L}{\partial \dot{\theta}}, \theta\,;\, \frac{\partial L}{\partial \dot{\phi}}, \phi\Big)\,.
$$

This transformation yields Hamilton's canonical equations for the elastic spherical pendulum in its natural spherical coordinates. In particular, the canonical momenta are linearly related to their corresponding velocities by

$$
\begin{aligned}
P = \frac{\partial L}{\partial \dot{R}} = m\,\dot{R} \quad &\Longleftrightarrow \quad \dot{R} = P/m\,, \\
P_\theta = \frac{\partial L}{\partial \dot{\theta}} = mR^2\,\dot{\theta} \quad &\Longleftrightarrow \quad \dot{\theta} = \frac{P_\theta}{mR^2}\,, \\
P_\phi = \frac{\partial L}{\partial \dot{\phi}} = mR^2\sin^2\theta\,\dot{\phi} \quad &\Longleftrightarrow \quad \dot{\phi} = \frac{P_\phi}{mR^2\sin^2\theta}\,.
\end{aligned}
\tag{5.2.18}
$$

Being linear, these relations between the momenta and velocities are easily solved. This means the Lagrangian (5.2.2) is non-singular.

Canonical Poisson bracket description

The Hamiltonian for the elastic spherical pendulum is obtained by Legendre transforming its Lagrangian (5.2.7), to produce

$$
\begin{aligned}
H &= P\dot{R} + P_\theta\dot{\theta} + P_\phi\dot{\phi} - L \\
&= \frac{P^2}{2m} + \frac{P_\theta^2}{2mR^2} + \frac{P_\phi^2}{2mR^2\sin^2\theta} \\
&\quad - mgR\cos\theta + \frac{k}{2}(R-\ell_0)^2\,.
\end{aligned}
\tag{5.2.19}
$$

This Hamiltonian yields the following canonical motion equations:

$$
\begin{aligned}
\dot{P} &= \{P,\,H\} = -\frac{\partial H}{\partial R} \\
&= -k(R-\ell_0) + mg\cos\theta + \frac{P_\theta^2}{mR^3} + \frac{P_\phi^2}{mR^3\sin^2\theta}\,, \\
\dot{P}_\theta &= \{P_\theta,\,H\} = -\frac{\partial H}{\partial \theta} \\
&= -mgR\sin\theta + \frac{P_\phi^2}{mR^2\sin^2\theta\tan\theta}\,, \\
\dot{P}_\phi &= \{P_\phi,\,H\} = -\frac{\partial H}{\partial \phi} = 0\,.
\end{aligned}
$$

As expected from Noether's theorem, the azimuthal angular momentum $P_\phi = J_3$ is conserved because the Lagrangian for the elastic spherical pendulum is independent of the azimuthal angle ϕ. This azimuthal symmetry also allows further progress toward characterising its motion. In particular, as we have seen, the equilibrium solutions of the elastic spherical pendulum are azimuthally symmetric.

Substituting velocities for the momenta in (5.2.18) recovers the motion equations (5.2.8). The velocities (5.2.18) may also be recovered in their canonical Hamiltonian forms as

$$
\begin{aligned}
\dot{R} &= \{R,\, H\} = \frac{\partial H}{\partial P} = P/m\,,\\
\dot{\theta} &= \{\theta,\, H\} = \frac{\partial H}{\partial P_\theta} = \frac{P_\theta}{mR^2}\,,\\
\dot{\phi} &= \{\phi,\, H\} = \frac{\partial H}{\partial P_\phi} = \frac{P_\phi}{mR^2 \sin^2\theta}\,.
\end{aligned}
$$

Transformation of coordinates and an approximation

We shift coordinates from the point of support, to the unstretched spring position

$$
\mathbf{x} = \mathbf{X} - \ell_0 \mathbf{e}_3\,,
$$

and assume sufficiently small excursions that $|\mathbf{x}|/\ell_0 = \epsilon \ll 1$. We then approximate the potential energy of oscillation V_{osc} to third order $O(\epsilon^3)$ using $\sqrt{1+\epsilon} = 1 + \epsilon/2 - \epsilon^2/8 + O(\epsilon^3)$ for $\epsilon \ll 1$ as

$$
\begin{aligned}
V_{osc} &= \frac{k}{2}(|\mathbf{X}| - \ell_0)^2 = \frac{k}{2}\Big(|\mathbf{x} + \ell_0\mathbf{e}_3| - \ell_0\Big)^2\\
&= \frac{k}{2}\ell_0^2\bigg(\sqrt{1 + \frac{2\mathbf{e}_3\cdot\mathbf{x}}{\ell_0} + \frac{|\mathbf{x}|^2}{\ell_0^2}} - 1\bigg)^2\\
&= \frac{k\ell_0^2}{2}\Big(\frac{z^2}{\ell_0^2} + \frac{z(x^2+y^2)}{\ell_0^3}\Big) + o(\epsilon^3)\,,
\end{aligned}
$$

in which $\mathbf{x} = (x, y, z)$ and $o(\epsilon^3)$ denotes neglected higher order terms. Likewise, we approximate the gravitational potential as

$$
\begin{aligned}
V_{grav} &= mg\ell(1-\cos\theta) = \frac{mg\ell}{2}\theta^2 + O(\theta^4)\\
&= \frac{mg}{2}\frac{(x^2+y^2)}{\ell} + o(\epsilon^3)\,,
\end{aligned}
$$

with $\theta^2 \simeq (x^2+y^2)/\ell^2 < \epsilon^2 \ll 1$. The corresponding Lagrangian, approximated to cubic order in the amplitudes, is

$$\begin{aligned} L/m = \frac{1}{2}(\dot{x}^2+\dot{y}^2+\dot{z}^2) \quad - \quad & \frac{1}{2}(\omega_R^2(x^2+y^2)+\omega_Z^2 z^2) \\ & + \frac{1}{2}\lambda(x^2+y^2)z\,, \end{aligned} \tag{5.2.20}$$

where x, y and z are Cartesian coordinates centred at the point of equilibrium, $\omega_R = \sqrt{g/\ell}$ is the frequency of linear pendular motion, $\omega_Z = \sqrt{k/m}$ is the frequency of its elastic oscillations and $\lambda = -\omega_Z^2/\ell$.

The Euler-Lagrange equations of motion for the cubically approximated Lagrangian (5.2.20) may be written as

$$\begin{aligned} \ddot{x}+\omega_R^2 x &= \lambda x z\,, \\ \ddot{y}+\omega_R^2 y &= \lambda y z\,, \\ \ddot{z}+\omega_Z^2 z &= \frac{1}{2}\lambda(x^2+y^2)\,. \end{aligned} \tag{5.2.21}$$

Ignoring λ-terms yields the linear modes of oscillation around the downward equilibrium.

Remark 5.2.7 *This system of equations has two constants of motion. These are the total energy E and the vertical angular momentum h given by*

$$\begin{aligned} E &= \frac{1}{2}\left(\dot{x}^2+\dot{y}^2+\dot{z}^2\right)+\frac{1}{2}\left(\omega_R^2(x^2+y^2)+\omega_Z^2 z^2\right)-\frac{1}{2}\lambda(x^2+y^2)z\,, \\ h &= (x\dot{y}-y\dot{x})\,. \end{aligned}$$

Remark 5.2.8 *The system (5.2.21) is not integrable. Its chaotic motions have been studied by many authors (see, e.g., references in [Ly2002b]). Previous studies have considered the two-dimensional case, for which the angular momentum vanishes, $h = 0$.*

5.2.2 Averaged Lagrangian technique

We confine attention to the resonant case $\omega_Z = 2\omega_R$ and apply the averaged Lagrangian technique [Wh1974]. In the averaged Lagrangian technique, the solution of (5.2.21) is assumed to be of the

1:1:2 form

$$
\begin{aligned}
x &= \Re[a(t)\exp(i\omega_R t)]\,, \\
y &= \Re[b(t)\exp(i\omega_R t)]\,, \\
z &= \Re[c(t)\exp(2i\omega_R t)]\,.
\end{aligned}
\tag{5.2.22}
$$

The complex coefficients $a(t)$, $b(t)$ and $c(t)$ are assumed to vary on a time scale which is considerably longer than the period of the oscillations, $\tau = 2\pi/\omega_R$.

Remark 5.2.9 *The representation of the solution of a dynamical system as the product of a slowly varying complex amplitude times a rapidly varying phase factor is often called the* ***slowly varying envelope*** *(SVE) approximation.*

To average the Lagrangian over the rapid time scale τ, one expands the various products in the Lagrangian and keeps only terms that have no rapid phase factors. That is, one argues that under time integration the rapid phase factors will average to zero. This phase averaging process yields the averaged action $\langle S\rangle = \int \langle L\rangle\, dt$ with

$$
\langle L\rangle = \frac{1}{2}\omega_R\Big[\Im\{\dot a a^* + \dot b b^* + 2\dot c c^*\} + \Re\{\kappa(a^2+b^2)c^*\}\Big]\,, \tag{5.2.23}
$$

where $\kappa = \lambda/(4\omega_R)$. Averaging the Lagrangian has injected an S^1 symmetry $(a,b,c) \to (e^{-2i\psi}a, e^{-2i\psi}b, e^{-4i\psi}c)$ which will lead to an additional conservation law.

One regards the quantities $a, b, c \in \mathbb{C}^3$ appearing in $\langle L\rangle$ as generalised coordinates. The Euler-Lagrange equations of motion for the averaged Lagrangian $\langle L\rangle$ are then

$$
\frac{d}{dt}\frac{\partial\langle L\rangle}{\partial \dot a} = \frac{\partial\langle L\rangle}{\partial a}\,, \quad \text{etc.}
$$

Explicitly, these Euler-Lagrange equations are

$$
i\dot a = \kappa a^* c\,, \quad i\dot b = \kappa b^* c\,, \quad i\dot c = \tfrac{1}{4}\kappa(a^2+b^2)\,. \tag{5.2.24}
$$

Remark 5.2.10 *Equations (5.2.24)) are the complex versions of equations (68)–(73) in [Ly2002a]. The latter were derived using the method of* ***multiple time-scale analysis****, where the small parameter ϵ for the analysis was the amplitude of the dependent variables, so that quadratic terms in the unknowns were second order, whereas linear terms were first order in ϵ. Thus, in this case, the averaged Lagrangian technique yields results completely equivalent to those achieved using multiple time-scale analysis.*

Transforming to the three-wave interaction equations

The remaining analysis is facilitated by making the following linear change of variables

$$A = \frac{1}{2}\kappa(a+ib)\,, \quad B = \frac{1}{2}\kappa(a-ib)\,, \quad C = \kappa c\,.$$

Consequently, the three-wave equations of motion (5.2.24) take the symmetric form

$$\begin{aligned} i\dot{A} &= B^*C\,, \\ i\dot{B} &= CA^*\,, \\ i\dot{C} &= AB\,. \end{aligned} \tag{5.2.25}$$

These three complex equations are well-known as the ***three-wave interaction equations***. They appear ubiquitously in nonlinear wave processes. For example, they govern quadratic wave resonance in fluids and plasmas. A brief history of these equations is given below.

Canonical form of three-wave interaction The three-wave interaction equations (5.2.25) may be written in canonical form with Hamiltonian $H = \Re(ABC^*)$ and Poisson brackets

$$\{A, A^*\} = \{B, B^*\} = \{C, C^*\} = -2i\,,$$

as

$$\begin{aligned} i\dot{A} &= i\{A, H\} = 2\partial H/\partial A^*\,, \\ i\dot{B} &= i\{B, H\} = 2\partial H/\partial B^*\,, \\ i\dot{C} &= i\{C, H\} = 2\partial H/\partial C^*\,. \end{aligned} \tag{5.2.26}$$

Remark 5.2.11 (Conservation laws for 3-wave equations)
The three-wave equations conserve the following three quantities:

$$
\begin{aligned}
H &= \frac{1}{2}(ABC^* + A^*B^*C) = \Re(ABC^*)\,, & (5.2.27)\\
J &= |A|^2 - |B|^2\,, & (5.2.28)\\
N &= |A|^2 + |B|^2 + 2|C|^2\,. & (5.2.29)
\end{aligned}
$$

The Hamiltonian vector field $X_H = \{\,\cdot\,,\,H\}$ generates the motion, while $X_J = \{\,\cdot\,,\,J\}$ and $X_N = \{\,\cdot\,,\,N\}$ generate S^1 symmetries $S^1 \times \mathbb{C}^3 \mapsto \mathbb{C}^3$ of the Hamiltonian H. The S^1 symmetries associated to J and N are the following:

$$
J : \begin{pmatrix} A \\ B \\ C \end{pmatrix} \to \begin{pmatrix} e^{-2i\phi}A \\ e^{2i\phi}B \\ C \end{pmatrix} \qquad N : \begin{pmatrix} A \\ B \\ C \end{pmatrix} \to \begin{pmatrix} e^{-2i\psi}A \\ e^{-2i\psi}B \\ e^{-4i\psi}C \end{pmatrix}
$$

The constant of motion J represents the angular momentum about the vertical in the new variables, while N is the new conserved quantity arising from phase-averaging in the Lagrangian L to obtain $\langle L\rangle$.

The following positive-definite combinations of N and J are physically significant:

$$
N_+ \equiv \frac{1}{2}(N + J) = |A|^2 + |C|^2\,, \qquad N_- \equiv \frac{1}{2}(N - J) = |B|^2 + |C|^2\,.
$$

*These combinations are known as the **Manley-Rowe invariants** in the extensive literature about three-wave interactions. The quantities H, N_+ and N_- provide three independent constants of the motion.*

Thus, the modulation equations for the swinging spring are transformed into the three-wave equations, which are known to be completely integrable. See [AlLuMaRo1998] for references to the three wave equations and an extensive elaboration of their properties as a paradigm for Hamiltonian reduction.

5.2.3 A brief history of the three-wave equations

Fluids and plasmas.
The three-wave equations (5.2.25) model the nonlinear dynamics of

the amplitudes of three waves in fluids or plasmas [Br1964]. These include, for example, interactions between planetary Rossby waves in the atmosphere in three-wave resonance [HaMi1977, LoGi1967]. The correspondence between Rossby waves in the atmosphere and drift waves in plasma has been thoroughly explored in [HoHa1994]. Resonant wave-triad interactions also play an essential role in the generation of turbulence and in determining the statistics of its power spectrum.

Laser-matter interaction.
The three-wave equations (5.2.25) are also equivalent to the Maxwell-Schrödinger envelope equations for the interaction between radiation and a two-level resonant medium in a microwave cavity. We shall take up these equations again in this context in (6.1.8) in Chapter 6. As shown in [HoKo1992] perturbations of this system lead to homoclinic chaos, but we shall not explore that issue here. A forced and damped version of the three-wave equations was used in [WeFiOt1980] to study instability saturation by nonlinear mode coupling, and irregular solutions were discovered there indicating the presence of a strange attractor. See also [Ot1993] and [HoKoWe1995] for more detailed studies of the perturbed three-wave system.

Nonlinear optics.
The three-wave system also describes the dynamics of the envelopes of light-waves interacting quadratically in nonlinear material. The system has been examined in a series of papers [AlLuMaRo1998, AlLuMaRo1999, LuAlMaRo2000] using a geometric approach, which allowed the reduced dynamics for the wave intensities to be represented as motion on a closed surface in three dimensions.

5.2.4 A special case of the three-wave equations

In the special case $H = 0$ the system (5.2.25) reduces to three real equations. Let

$$A = iX_1 \exp(i\phi_1)\,, \quad B = iX_2 \exp(i\phi_2)\,, \quad C = iX_3 \exp(i(\phi_1 + \phi_2))\,,$$

where X_1, X_2 and X_3 are real and the phases ϕ_1 and ϕ_2 are constants. The modulation equations become

$$\dot{X}_1 = -X_2 X_3\,, \quad \dot{X}_2 = -X_3 X_1\,, \quad \dot{X}_3 = +X_1 X_2\,. \tag{5.2.30}$$

These equations are re-scaled versions of the Euler equations for the rotation of a free rigid body. The dynamics in this special case is expressible as motion on $\mathbf{R}^3$, namely

$$\dot{\mathbf{X}} = \tfrac{1}{8}\nabla J \times \nabla N = \tfrac{1}{4}(\nabla N_+ \times \nabla N_-)\,, \tag{5.2.31}$$

where

$$N_+ = \frac{1}{2}(X_1^2 + X_3^2) \quad \text{and} \quad N_- = \frac{1}{2}(X_2^2 + X_3^2)\,. \tag{5.2.32}$$

Considering the constancy of J and N, we can describe a trajectory of the motion as an intersection between a hyperbolic cylinder (J constant, see (5.2.28)) and an oblate spheroid (N constant; see (5.2.29)). Equation (1) provides an alternative description. Here we have used the freedom in the $\mathbf{R}^3$ Poisson bracket exploited in [HoMa1991, DaHo1992] and discussed in Section 2.4.4 to represent the equations of motion on the intersection of two orthogonal *circular* cylinders, the level surfaces of the Manley-Rowe quantities, N_+ and N_-. The invariance of the trajectories means that while the level surfaces of J and N differ from those of N_+ and N_-, their intersections are *precisely the same*. For this particular value of $H = 0$, the motion may be further reduced by expressing it in the coordinates lying on one of these two circular cylinders, on which it becomes pendular motion. See Section 2.4.4 for the corresponding transformation of rigid-body motion into pendular motion. See [DaHo1992, AlLuMaRo1998, AlLuMaRo1999] for discussions of geometric phases in this situation.

Exercise.

a. Characterise the equilibrium points of the dynamical system (1) geometrically in terms of the gradients of $N_\pm$. How many are there? Which are stable?

b. Choose cylindrical polar coordinates along the axis of the circular cylinder that represents the level set of N_+ and show that the $\mathbb{R}^3$ Poisson bracket restricted to that level set is canonical.

c. Write the equation of motion on the level set of N_+ as a pendulum equation.

d. For any closed orbits on the level set of N_+, write formulas for its geometric and dynamic phases.

★

5.3 Reduction and reconstruction of solutions

To reduce the system for $H \neq 0$, we employ a further canonical transformation, introduced in [HoKo1992]. The goal is to encapsulate complete information about the Hamiltonian in a single variable Z by using the invariants of the motion. Once Z is found, the Manley-Rowe relations yield the remaining variables. We set:

$$
\begin{aligned}
A &= |A| \exp(i\xi)\,, \\
B &= |B| \exp(i\eta)\,, \\
C &= Z \exp(i(\xi+\eta))\,.
\end{aligned}
\tag{5.3.1}
$$

This transformation is canonical – it preserves the symplectic form

$$dA \wedge dA^* + dB \wedge dB^* + dC \wedge dC^* = dZ \wedge dZ^*\,.$$

In these variables, the Hamiltonian is a function of only Z and Z^*

$$H = \frac{1}{2}(Z+Z^*) \cdot \sqrt{N_+ - |Z|^2} \cdot \sqrt{N_- - |Z|^2}\,.$$

The Poisson bracket is $\{Z, Z^*\} = -2i$ and the canonical equations reduce to

$$i\dot{Z} = i\{Z, H\} = 2\frac{\partial H}{\partial Z^*}\,.$$

This provides the slow dynamics of both the amplitude and phase of $Z = |Z|e^{i\zeta}$.

The amplitude $|Z| = |C|$ is obtained in closed form in terms of Jacobi elliptic functions as the solution of

$$\left(\frac{d\mathcal{Q}}{d\tau}\right)^2 = \left[\mathcal{Q}^3 - 2\mathcal{Q}^2 + (1-\mathcal{J}^2)\mathcal{Q} + 2\mathcal{E}\right], \tag{5.3.2}$$

where $\mathcal{Q}$, $\mathcal{J}$, $\mathcal{E}$ and τ are normalised by the constant of motion N as

$$\mathcal{Q} = \frac{2|Z|^2}{N}, \quad \mathcal{J} = \frac{J}{N}, \quad \mathcal{E} = -\frac{4H^2}{N^3}, \quad \tau = \sqrt{2N}t\,. \tag{5.3.3}$$

Once $|Z|$ is known, $|A|$ and $|B|$ follow immediately from the Manley-Rowe relations,

$$|A| = \sqrt{N_+ - |Z|^2}\,, \qquad |B| = \sqrt{N_- - |Z|^2}\,.$$

The phases ξ and η may now be determined. Using the three-wave equations (5.2.25) together with (5.3.1), one finds

$$\dot{\xi} = -\frac{H}{|A|^2}\,, \qquad \dot{\eta} = -\frac{H}{|B|^2}\,, \tag{5.3.4}$$

so that ξ and η can be integrated by quadratures once $|A|(t)$ and $|B|(t)$ are known. Finally, the phase ζ of Z is determined unambiguously by

$$\frac{d|Z|^2}{dt} = -\,2H\tan\zeta \qquad \text{and} \qquad H = |A||B||Z|\cos\zeta\,. \tag{5.3.5}$$

Hence, we can now reconstruct the full solution as,

$$A = |A|\exp(i\xi)\,, \quad B = |B|\exp(i\eta)\,, \quad C = |Z|\exp\big(i(\xi+\eta+\zeta)\big)\,.$$

5.3.1 Phase portraits

Consider the plane $\mathcal{C}$ in phase space defined by $A = B = 0$. This is a plane of unstable equilibrium points, representing purely vertical oscillations of the spring. The Hamiltonian vanishes identically on this plane, as does the angular momentum J. Each point c_0 in $\mathcal{C}$ has a heteroclinic orbit linking it to its antipodal point $-c_0$. Thus,

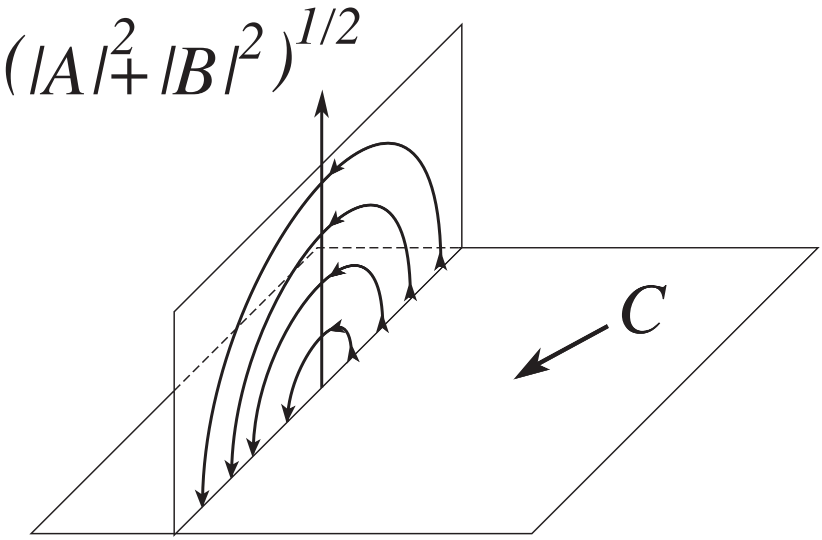


Figure 5.2: C is the plane of critical points, $A = 0 = B$. The vertical axis is $R = |A|^2 + |B|^2$. The vertical plane contains heteroclinic semi-ellipses passing from c_0 to $-c_0$.

the plane $\mathcal{C}$ of critical points is connected to itself by heteroclinic orbits. In Figure 5.2, the horizontal plane is $\mathcal{C}$ and the vertical plane contains heteroclinic orbits from c_0 to $-c_0$.[1]

The vertical axis is $R = \sqrt{|A|^2 + |B|^2}$. Since $N = R^2 + 2|C|^2$ is constant, each heteroclinic orbit is a semi-ellipse. Motion starting on one of these semi-ellipses will move towards an endpoint, taking infinite time to reach it.

In Figure 5.3, taken from [HoKo1992], we present another view of the trajectories for $J = 0$. The Hamiltonian is

$$H = \frac{1}{2}(Z + Z^*)\Big(\frac{1}{2}N - |Z|^2\Big)\,.$$

Accessible points lie on or within the circle $|Z|^2 = N/2$. For $H = 0$ the trajectory is the segment of the imaginary axis within the cir-

[1]Figures in this chapter are from [HoLy2002], with publisher's permission.

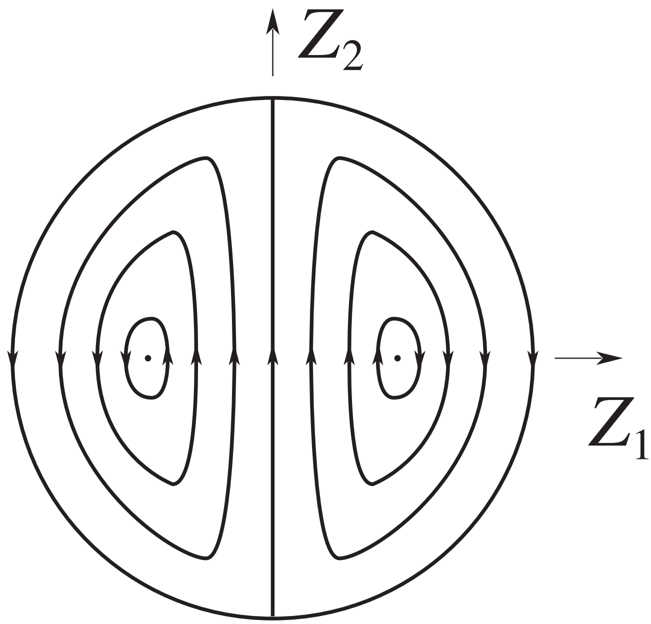


Figure 5.3: Phase portrait in the Z-plane for $J = 0$. The motion is confined within the circle $|Z|^2 = N/2$. The segment of the imaginary axis within this circle is the homoclinic orbit.

cle. This is the homoclinic orbit. For $H \neq 0$, one may solve for the imaginary part of $Z = Z_1 + iZ_2$,

$$Z_2 = \pm\sqrt{-Z_1^2 + \frac{1}{2}N - (H/Z_1)}\,.$$

This relation determines the trajectories in the complex Z plane for the range of H in which real solutions exist. There are two equilibrium points, at $Z = \pm\sqrt{N/6}$, corresponding to solutions in which no energy is exchanged between the vertical and horizontal components. These correspond to the ***cup-like*** and ***cap-like solutions*** first discussed in [ViGo1933].

5.3.2 Geometry of the motion for fixed J

The vertical amplitude is governed by equation (5.3.2), which may be rewritten as

$$\frac{1}{2}\left(\frac{d\mathcal{Q}}{d\tau}\right)^2 + \mathcal{V}(\mathcal{Q}) = \mathcal{E}\,, \tag{5.3.6}$$

with the potential $\mathcal{V}(\mathcal{Q})$ parameterised by $\mathcal{J}$ in (5.3.3) and given by

$$\mathcal{V}(\mathcal{Q}) = -\frac{1}{2}\left[\mathcal{Q}^3 - 2\mathcal{Q}^2 + (1-\mathcal{J}^2)\mathcal{Q}\right]. \tag{5.3.7}$$

The potential $\mathcal{V}(\mathcal{Q})$ has three zeros, $\mathcal{Q} = 0$, $\mathcal{Q} = 1 - \mathcal{J}$ and $\mathcal{Q} = 1 + \mathcal{J}$. Equation (5.3.6) is an energy equation for a particle of unit mass, with position $\mathcal{Q}$ and energy $\mathcal{E}$, moving in a cubic potential field $\mathcal{V}(\mathcal{Q})$. Figure 5.4 plots $\dot{\mathcal{Q}}$, given by (5.3.6), against $\mathcal{Q}$ for the cases $\mathcal{J} = 0$ (left panel) and $\mathcal{J} = 0.25$ (right panel), for a set of values of $\mathcal{E}$ given by

$$\mathcal{E} \in \{-0.0635, -0.0529, -0.0423, -0.0317, -0.0212, -0.0106, 0\}\,.$$

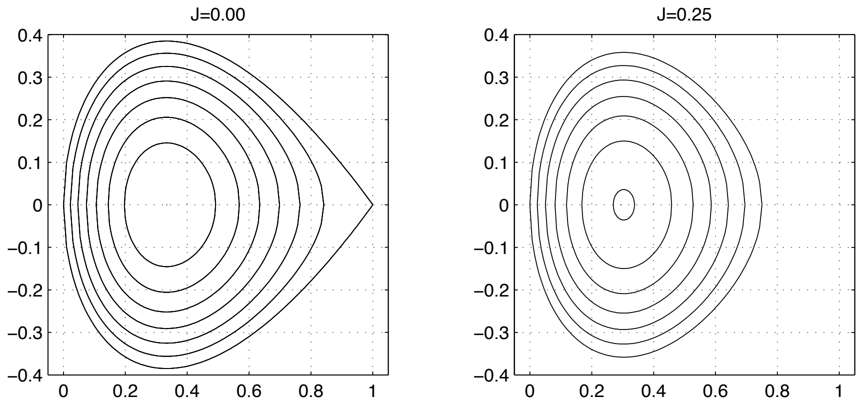


Figure 5.4: Plots of $\dot{Q}$ versus Q for $J = 0$ and $J = 0.25$ for a range of values of $\mathcal{E}$.

Each curve represents the projection onto the reduced phase-space of the trajectory of the modulation envelope. The centres are relative equilibria, corresponding to the elliptic-parabolic solutions in [Ly2002a], which are generalisations of the cup-like and cap-like solutions of [ViGo1933]. The case $\mathcal{J} = 0$ includes the homoclinic trajectory, for which $H = 0$.

5.3.3 Geometry of the motion for $H = 0$

For arbitrary $\mathcal{J}$, the $H = 0$ motions lie on a surface in the space with coordinates $(\mathcal{Q}, \dot{\mathcal{Q}}, \mathcal{J})$. This surface is depicted in Figure 5.5. It has three singular points (i.e., it is equivalent to a sphere with

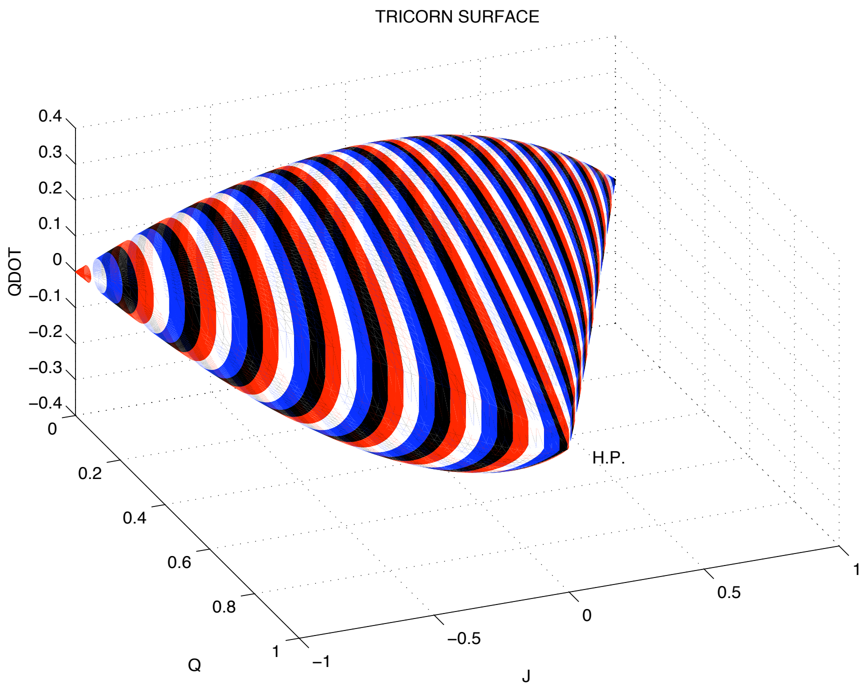


Figure 5.5: Tricorn surface, upon which motion takes place when $H = 0$. The coordinates are $J, Q, \dot{Q}$. The motion takes place on the intersections of this surface with a plane of constant J (such planes are indicated by the stripes). This surface has three singular points. The homoclinic point is marked H.P.

three pinches) and its shape is similar to a tricorn hat. The motion takes place on an intersections of this surface with a plane of constant $\mathcal{J}$. There are three equilibrium solutions: the first corresponds to $\mathcal{J} = 0$. This is marked in Figure 5.5 with the letters H.P. denoting the *homoclinic point* at the extremity of the homoclinic trajectory. This homoclinic point corresponds to purely vertical oscillatory motion. The other two equilibrium solutions with $\mathcal{J} = \pm 1$ correspond to purely horizontal motion, clockwise or counterclockwise, with the spring tracing out a cone. The purely vertical motion is unstable; while the two conical motions are stable. (The linear evolution equations for perturbations about conical motion were investigated in [Ly2002a].) The dynamics on the tricorn surface is similar to the motion of a free rigid body. The instability at the singular homoclinic point is responsible for the step-wise switching of the azimuthal angle when it undergoes nearly vertical oscillations.

5.3.4 Three-wave surfaces

There is yet another way to depict the motion in a reduced phase-space. Let us consider a reduced phase-space with x and y axes $X = \Re\{ABC^*\}$ and $Y = \Im\{ABC^*\}$ and z-axis $\mathcal{Q} = 2|Z|^2/N$. We note that $X \equiv H$. It follows from (5.2.27) and (5.2.28) that

$$X^2 + Y^2 = |A|^2|B|^2|C|^2 = \tfrac{1}{4}|Z|^2\Big[(2|Z|^2 - N)^2 - J^2\Big] .$$

We define $\mathcal{X} = (2/N^{3/2})X$ and $\mathcal{Y} = (2/N^{3/2})Y$, then write

$$\mathcal{X}^2 + \mathcal{Y}^2 = \frac{1}{2}\left[\mathcal{Q}^3 - 2\mathcal{Q}^2 + (1 - \mathcal{J}^2)\mathcal{Q}\right] = -\mathcal{V}(\mathcal{Q})\,, \tag{5.3.8}$$

where $\mathcal{V}$ is defined in (5.3.7). We note that $\mathcal{X}^2 = -\mathcal{E}$ and $\mathcal{Y}^2 = \frac{1}{2}(d\mathcal{Q}/d\tau)^2$. Equation (5.3.8) implies that the motion takes place on a surface of revolution about the $\mathcal{Q}$-axis. The radius for a given value of $\mathcal{Q}$ is the square-root of the cubic $-\mathcal{V}(\mathcal{Q})$. The physically accessible region is $0 \le \mathcal{Q} \le 1 - |\mathcal{J}|$. Several such surfaces (for $\mathcal{J} \in \{0.0, 0.1, 0.2, 0.3\}$) are shown in Figure 5.6.

Since $\mathcal{X}^2 = \mathcal{H}^2 = 4H^2/N^3$, the motion in $\mathbb{R}^3$ with coordinates $\mathcal{X}, \mathcal{Y}, \mathcal{Q}$ for given $\mathcal{J}$ takes place on the intersection of the corresponding three-wave surface of revolution in equation (5.3.8) with a plane of constant $\mathcal{X}$. These intersections recover the closed orbits shown in Figure 5.4 .

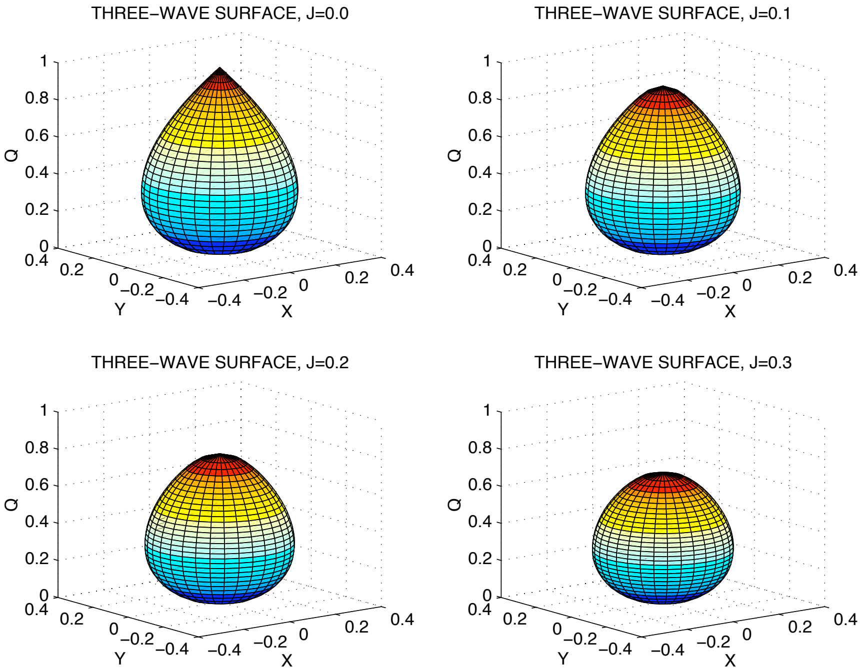


Figure 5.6: Surfaces of revolution about the Q-axis for $J \in \{0.0, 0.1, 0.2, 0.3\}$. The radius for given Q is determined by the square-root of the cubic $-V(Q)$. For given J, the motion takes place on the intersection of the corresponding surface with a plane of constant X.

Motion on the tricorn surface is related to motion on the three-wave surfaces of revolution, as follows. The tricorn surface refers to $H = 0$; the reduced $1:1:2$ resonant motion for $H \neq 0$ is represented by trajectories *inside* this surface. Slicing the tricorn surface in a plane of fixed $\mathcal{J}$ produces a set of closed trajectories, the outermost of these is for $H = 0$ the others arise for $H \neq 0$ The cases $\mathcal{J} = 0$ and $\mathcal{J} = 0.25$ are plotted in Figure 5.4 above. If the $\mathcal{J}$-section is distorted into a cup-like surface, by taking H as a vertical coordinate and plotting each trajectory at a height depending on its H value, one finds half of a closed surface. Each trajectory is selected by a H-plane section. Alteration of the sign of H corresponds to reversal of time. Completing the surface by reflection in the plane $H = 0$ gives the surface generated by rotating the root-cubic graph

$\sqrt{-\mathcal{V}(\mathcal{Q})}$ about the $\mathcal{Q}$-axis, i.e., the surface given by (5.3.8). These surfaces are called ***three-wave surfaces*** in [AlLuMaRo1998]. They foliate the volume contained within the surface for $\mathcal{J} = 0$.

5.3.5 Precession of the swing plane

The characteristic feature of the behaviour of the physical spring is its step-wise precession, which we shall now discuss. As the motion changes cyclically from horizontal swinging to vertical springing oscillations, one observes that each successive horizontal swinging excursion departs in a direction rotated from the previous swing-plane by a constant angle about the vertical. The horizontal projection of the motion appears as an ellipse of high eccentricity and the swinging motion is primarily confined to a vertical plane. The vertical plane through the major axis of this ellipse is called the *swing plane*. Starting from nearly vertical initial oscillations, the motion gradually develops into an essentially horizontal swinging motion. This horizontal swinging does not persist, but soon passes back again into nearly vertical springing oscillations similar to the initial motion. Subsequently, a horizontal swing again develops, but now in the rotated swing plane as shown in Figure 5.7. The step-wise precession of this exchange between springing and swinging motion continues indefinitely in the absence of dissipation and is the characteristic experimental feature of the swinging spring. An expression for the characteristic step-wise change in direction of the swing plane from one horizontal excursion was derived by using a method called pattern evocation in [HoLy2002]. In general, this method yields only approximate results. However, in this case, the result yielded by pattern evocation turns out to be exact, as was shown by using the monodromy method in [DuGiCu2004].

The pattern evocation method and its prediction for the step-wise dynamics of the precession angle in Figure 5.7 may be explained by returning to the solution (5.2.22) of (5.2.21) in the $1:1:2$

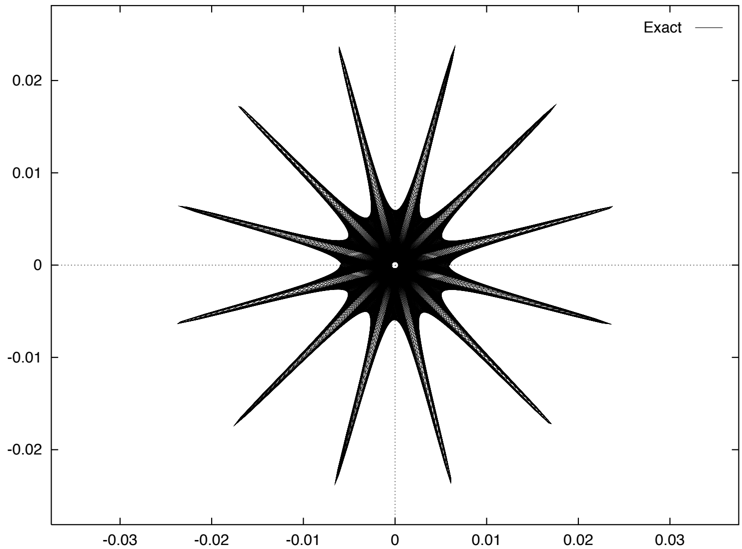


Figure 5.7: A projection of the regular step-wise precession of the swing plane of the pendulum after each return to its springing motion makes a star-shaped pattern. The magnitude of the precession angle is determined from the initial conditions.

resonant form,

$$\begin{aligned} x &= \Re[a(t)\exp(i\omega_R t)]\,, \\ y &= \Re[b(t)\exp(i\omega_R t)]\,, \\ z &= \Re[c(t)\exp(2i\omega_R t)]\,. \end{aligned} \tag{5.3.9}$$

One defines the precession angle in terms of the horizontal (x, y) projection of the trajectory of the pendulum. Recall that the full solution for the horizontal components is

$$\begin{aligned} x &= \Re\{a\exp(i\omega_R t)\} = |a|\cos(\omega_R t + \alpha)\,, \\ y &= \Re\{b\exp(i\omega_R t)\} = |b|\cos(\omega_R t + \beta)\,, \end{aligned}$$

where α and β are the phases of a and b. The amplitudes and phases are assumed to vary slowly. If they are regarded as constant over a period $\tau = 1/\omega_R$ of the fast motion, these equations describe a central ellipse,

$$Px^2 + 2Qxy + Ry^2 = S\,, \tag{5.3.10}$$

where $P = |b|^2$, $Q = -|ab|\cos(\alpha - \beta)$, $R = |a|^2$ and $S = J^2$. The area of the ellipse is easily calculated and is found to have the constant value πJ. Its orientation is determined by eliminating the cross-term in (5.3.10) by rotating the axes through an angle θ, given by

$$\tan 2\theta = \frac{2Q}{P-R} = \frac{2|ab|\cos(\alpha-\beta)}{|a|^2 - |b|^2}\,. \tag{5.3.11}$$

The semi-axes of the ellipse are given by

$$\begin{aligned} A_1 &= \frac{J}{\sqrt{P\cos^2\theta + Q\sin 2\theta + R\sin^2\theta}}\,, \\ A_2 &= \frac{J}{\sqrt{P\sin^2\theta - Q\sin 2\theta + R\cos^2\theta}}\,. \end{aligned}$$

The area of the ellipse is $\pi A_1 A_2 = \pi J$ and its eccentricity may be calculated immediately. In the case of unmodulated motion, the instantaneous ellipse corresponds to the trajectory. The amplitudes in the modulated motion are only approximations to the trajectory, but one may define the orientation or azimuth of the swing plane at any time to be the angle θ given by (5.3.11). This is the basis of the pattern evocation method.

The approximate and exact solutions of the swing plane azimuthal angle θ in (5.3.11) were found by [HoLy2002] in numerical simulations to track each other essentially to machine error. The reason for this was later explained by [DuGiCu2004] who showed that the step-wise precession angle of the swing plane of the resonant spring pendulum is a topological winding number depending only on initial conditions whose value was unaffected by the integrable approximation introduced here. This explanation is beyond our present scope, but its lesson is clear. Namely, implementing structure-preserving approximations may sometimes preserve exactly what is needed to understand the global behaviour of a nonlinear dynamical system.

Chapter 6

Maxwell-Bloch laser-matter equations

This chapter brings us to the threshold of modern applications of geometric mechanics. It investigates the real-valued Maxwell-Bloch equations on $\mathbb{R}^3$ as a Hamiltonian dynamical system obtained by applying an S^1 symmetry reduction to an invariant subsystem of a dynamical system on $\mathbb{C}^3$. These equations on $\mathbb{R}^3$ are bi-Hamiltonian and possess several inequivalent Lie-Poisson structures parameterised by classes of orbits in the group $SL(2, \mathbb{R})$, as discussed in [DaHo1992]. Each Lie-Poisson structure possesses an associated Casimir function. When reduced to level sets of these functions, the motion takes various symplectic forms, ranging from that of the pendulum to that of the Duffing oscillator. The values of the geometric and dynamic phases obtained in reconstructing the solutions are found to depend upon the choice of Casimir function, that is, upon the parameterisation of the reduced symplectic space.

6.1 Self-induced transparency

In ***self-induced transparency***, radiation energy leaves the leading edge of an optical laser pulse, coherently excites the atoms of a resonant dielectric medium, and then returns to the trailing edge of the pulse with no loss, but with a delay caused by temporary storage of pulse energy in the atoms. (This delay shows up as an anomalously slow pulse, whose velocity may be one-thousandth, or less, than the speed of light in a vacuum.) The physics of self-induced transparency is reviewed in [AlEb1975]. To the extent that resonant interaction of coherent light with a medium calls into play only a single atomic transition and the laser may be taken to be monochromatic, the medium has effectively only two levels.

For sufficiently short pulse duration, the coherent interaction between the pulse and the medium leading to self-induced transparency may be taken to be lossless. For most lasers and most atoms, this two-level lossless model is an excellent approximation and is quite adequate for an understanding of the basic physics behind many coherent transient phenomena. Self-induced transparency equations based upon this model are derived from the Maxwell-Schrödinger equations in [HoKo1991] by averaging over fast phases in the variational principle for the Maxwell-Schrödinger equations. A sketch of that derivation is given now in order to facilitate the considerations of the rest of the chapter.

6.1.1 The Maxwell-Schrödinger Lagrangian

The dimensionless Maxwell-Schrödinger equations on space-time parameterised by $(z,t) \in \mathbb{R} \times \mathbb{R}$ are,

$$\begin{aligned} E_{zz} - \ddot{E} &= 2\kappa \ddot{P}\,, \\ \dot{a}_+ &= \frac{1}{2\kappa} a_+ - E a_-\,, \\ \dot{a}_- &= \frac{1}{2\kappa} a_- - E a_+\,, \end{aligned} \tag{6.1.1}$$

where $E(z,t)$ denotes the electric field and $P(z,t)$ is the polarisability. The latter may be written in terms of the two atomic-level amplitudes a_+ and a_- as

$$P = a_+ a_-^* + a_+^* a_-\,. \tag{6.1.2}$$

The ratio of frequencies, $\kappa = \omega_c/\omega_0 << 1$, is a small parameter. Here ω_0 is the atomic transition frequency, and $\omega_c = (2\pi r n d^2 \omega_0/\hbar)^{1/2}$ is the cooperative frequency of the medium with dipole density n and atomic dipole moment d. The wave equation for the linearly polarised electric field E follows from Maxwell's equations

$$\dot{D} = B_z, \qquad \dot{B} = E_z\,, \tag{6.1.3}$$

where $D = E + 2\kappa P$ is the electric displacement and B is the magnetic field. Introducing the magnetic vector potential A that satisfies the relations $\dot{A} = E$ and $A_z = B$ allows the Maxwell-Schrödinger equations to be written as stationarity conditions for Hamilton's principle, $\delta S = 0$, with action S given by

$$S = \int \Big[\frac{1}{2}\dot{A}^2 - \frac{1}{2}A_z^2 + 2\kappa\dot{A}\left(a_+a_-^* + a_+^*a_-\right) - \left(|a_+|^2 - |a_-|^2\right) \\ + i\kappa\left(a_+^*\dot{a}_+ - a_+\dot{a}_+^* - a_-^*\dot{a}_- - a_-\dot{a}_-^*\right)\Big]\,dzdt\,.$$

The third term in the integrand of the action S is the interaction term, which couples the electromagnetic field to the matter fields. Stationary variations with respect to A, a_-^*, and a_+^* give

$$\begin{aligned}
\delta A : &\qquad \ddot{A} + 2\kappa\dot{P} - A_{zz} = 0\,,\\
\delta a_+^* : &\qquad i\dot{a}_+ - \frac{1}{2\kappa}a_+ + Ea_- = 0\,,\\
\delta a_-^* : &\qquad i\dot{a}_- + \frac{1}{2\kappa}a_- + Ea_+ = 0\,.
\end{aligned}$$

6.1.2 Envelope approximation

The Maxwell-Schrödinger equations may be simplified by writing the atomic amplitudes as ***modulated travelling waves***,

$$a_+ = ue^{-i(t-z)/2\kappa}, \qquad a_- = ve^{+i(t-z)/2\kappa}\,, \tag{6.1.4}$$

with complex envelope functions $u, v \in \mathbb{C}$. We also take the vector potential to be a modulated rightward moving wave, also in the envelope form,

$$A = i\kappa we^{-i(t-z)/\kappa} - i\kappa w^* e^{i(t-z)/\kappa}\,, \tag{6.1.5}$$

with the complex envelope function w. In these expressions, the ***complex envelope functions*** u, v and w are assumed to depend only on time, t. Now the electric field is given by

$$E = \dot{A} = we^{-i(t-z)/\kappa} + w^* e^{i(t-z)/\kappa} + O(\kappa)\,. \tag{6.1.6}$$

Thus, to first order in $\kappa \ll 1$, the quantity w is the complex electric field envelope.

6.1.3 Averaged Lagrangian for envelope equations

Averaging the action S over the fast phases, performing the integration in z, dividing by 4κ, and dropping terms of higher order in κ yields the ***phase-averaged action***

$$\begin{aligned}\bar{S} \;&:=\; \int \Big[i(w^*\dot{w} - w\dot{w}^*) + i(u^*\dot{u} - u\dot{u}^*) \\ &\qquad + i(v^*\dot{v} - v\dot{v}^*) + wu^*v + w^*uv^* \Big] dt\,. \end{aligned} \tag{6.1.7}$$

Varying the phase-averaged action $\bar{S}$ yields

$$\begin{aligned}\delta\bar{S} \;&=\; \int \Big[\delta w^* \big(i\dot{w} + uv^*\big) + \delta w\big(- i\dot{w}^* + u^*v\big) \\ &\qquad + \delta u^* \big(i\dot{u} + wv^*\big) + \delta u\big(- i\dot{u}^* + v^*w^*\big) \\ &\qquad + \delta v^* \big(i\dot{v} + uw^*\big) + \delta v\big(- i\dot{v}^* + u^*w\big) \Big] dt\,. \end{aligned}$$

Thus, stationarity of the phase-averaged action $\bar{S}$ implies the following ***Maxwell-Schrödinger envelope (MSE) equations***

$$\dot{u} = ivw\,, \qquad \dot{v} = iuw^*\,, \qquad \dot{w} = iuv^*\,. \tag{6.1.8}$$

These equations for the wave envelopes (u, v, w) recover the three-wave equations (5.2.25) for the $1 : 1 : 2$ resonance of the swinging spring (or elastic spherical pendulum) treated in Chapter 5 under the map $(u, v, w) \to (-C, -A, -B)$. Thus, the analysis of the swinging spring from Chapter 5 will also apply here, and the further analysis in this chapter for the wave envelope equations will also apply to the swinging spring.

The Maxwell-Schrödinger envelope equations (6.1.8) are canonically Hamiltonian on $\mathbb{C}^3$ with symplectic form,

$$\Omega = \frac{1}{-2i}\left[dw \wedge dw^* + \, du \wedge du^* + \, dv \wedge dv^*\right]. \tag{6.1.9}$$

The Hamiltonian function in these variables is given by,

$$H\,(u, v, w) = -\,\frac{1}{2}\,(u^* vw + uv^* w^*) = -\,\Re(u^* vw)\,. \tag{6.1.10}$$

The MSE equations (6.1.8) also conserve the ***Manley-Rowe invariants***,

$$C = |u|^2 + |v|^2 \qquad \text{and} \qquad C' = |u|^2 + |w|^2\,. \tag{6.1.11}$$

These two conservation laws arise because of invariance of the Hamiltonian (6.1.11) under the two S^1 phase shifts,

$$u \longrightarrow e^{i\theta} u, \quad v \longrightarrow e^{i\theta} v, \quad \text{and} \quad u \longrightarrow e^{i\varphi} u \quad w \longrightarrow e^{i\varphi} w, \tag{6.1.12}$$

generated by C and C', respectively. The Hamiltonian vector field for the MSE Hamiltonian $H(u, v, w, u^*, v^*, w^*)$ is given by

$$\begin{aligned} X_H \quad &= \quad 2i\left(\frac{\partial H}{\partial u}\frac{\partial}{\partial u^*} - \frac{\partial H}{\partial u^*}\frac{\partial}{\partial u}\right) \\ &\qquad + 2i\left(\frac{\partial H}{\partial v}\frac{\partial}{\partial v^*} - \frac{\partial H}{\partial v^*}\frac{\partial}{\partial v}\right) \\ &\qquad + 2i\left(\frac{\partial H}{\partial w}\frac{\partial}{\partial w^*} - \frac{\partial H}{\partial w^*}\frac{\partial}{\partial w}\right). \end{aligned} \tag{6.1.13}$$

6.1.4 Complex Maxwell-Bloch equations

The five-dimensional ***Maxwell-Bloch system*** [AlEb1975] is obtained by reducing the system (6.1.8) on $\mathbb{C}^3$ by using the S^1 group action in θ generated by the constant C. We introduce the following transformation to coordinates which are invariant under the S^1 action generated by C,

$$ix = 2w, \quad y = 2uv^*, \quad z = |u|^2 - |v|^2\,, \quad C = |u|^2 + |v|^2\,. \tag{6.1.14}$$

This is the ***Hopf fibration*** in the (u, v) coordinates for $\mathbb{C}^2$. Physically, the variables x, y, and z represent the electric field, the polarisation, and the population inversion, respectively. Transformation (6.1.14) gives us a five-dimensional system on the space $\mathbb{C}^2 \times \mathbb{R}$, coordinatised by (x, y, z). The Hamiltonian function, the Hamiltonian vector field for H, the equations of motion, and the new constants formed from those in (6.1.11) become

$$H = \frac{i}{2}\left(x^* y - x y^*\right) = -\Im(x^* y), \qquad (6.1.15)$$

$$\begin{aligned} X_H = \; & 2i\left(\frac{\partial H}{\partial x}\frac{\partial}{\partial x^*} - \frac{\partial H}{\partial x^*}\frac{\partial}{\partial x}\right) + 2iz\left(\frac{\partial H}{\partial y}\frac{\partial}{\partial y^*} - \frac{\partial H}{\partial y^*}\frac{\partial}{\partial y}\right) \\ & + iy\left(\frac{\partial H}{\partial y}\frac{\partial}{\partial z^*} - \frac{\partial H}{\partial z}\frac{\partial}{\partial y}\right) + iy^*\left(\frac{\partial H}{\partial z}\frac{\partial}{\partial y^*} - \frac{\partial H}{\partial y^*}\frac{\partial}{\partial z}\right), \end{aligned}$$

$$\dot{x} = y, \quad \dot{y} = xz, \quad \dot{z} = \frac{1}{2}\left(x^* y + x y^*\right), \qquad (6.1.16)$$

$$K = z + \frac{1}{2}|x|^2 \quad \text{and} \quad L = |y|^2 + z^2 .$$

Physically, K is the sum of the atomic excitation energy and the electric field energy, while $L = 1$, for unitarity. The phase space geometry of the solutions of the Maxwell-Bloch system (6.1.16) on $\mathbb{C}^2 \times \mathbb{R}$ and its *three* Hamiltonian structures are discussed in [FoHo1991], in the context of Lax pairs for soliton theory.

6.1.5 Real Maxwell-Bloch equations

In the remainder of this chapter, we shall discuss the phase space geometry and Hamiltonian structure of the invariant subsystem of (6.1.16) obtained by restricting to real-valued x and y. Thus, the dynamics of this invariant subsystem lie on the zero-level surface of the Hamiltonian function H in (6.1.15), with coordinates $x_1 =$

$\Re(x)$, $x_2 = \Re(y)$, and $x_3 = z$. The equations of motion (6.1.16) then become the three-dimensional ***real-valued Maxwell-Bloch system,***

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = x_1 x_3, \quad \dot{x}_3 = -x_1 x_2\,, \tag{6.1.17}$$

which is amenable to the methods of geometric mechanics developed in the present text. Equations (6.1.17) also appear as the large Rayleigh number limit of the famous Lorenz system (see [Sp1982]).

6.2 Classifying Lie-Poisson Hamiltonian structures for real-valued Maxwell-Bloch

The real-valued Maxwell-Bloch system (6.1.17)

$$\dot{x}_1 = x_2\,, \quad \dot{x}_2 = x_1 x_3\,, \quad \dot{x}_3 = -x_1 x_2\,, \tag{6.2.1}$$

is expressible in three-dimensional vector notation as

$$\dot{\mathbf{x}} = \nabla H_1 \times \nabla H_2\,, \tag{6.2.2}$$

where H_1 and H_2 are the two conserved functions

$$H_1 = \frac{1}{2}(x_2^2 + x_3^2) \quad \text{and} \quad H_2 = x_3 + \frac{1}{2}x_1^2\,. \tag{6.2.3}$$

Geometrically, equation (6.2.2) implies that the motion takes place on intersections of level sets of the functions H_1 and H_2 in the space $\mathbb{R}^3$ with coordinates (x_1, x_2, x_3). In fact, this geometic characterisation is not unique.

Proposition 6.2.1 *Equations (6.2.2) may be re-expressed equivalently as*

$$\dot{\mathbf{x}} = \nabla H \times \nabla C\,, \tag{6.2.4}$$

in which H and C are the following $SL(2,\mathbb{R})$ combinations of the functions H_1 and H_2,

$$\begin{bmatrix} H \\ C \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \mu & \nu \end{bmatrix} \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} \quad \textit{with} \quad \alpha\nu - \beta\mu = 1\,. \tag{6.2.5}$$

Proof. For the proof, it suffices to verify

$$\begin{aligned}\dot{\mathbf{x}} &= \nabla H \times \nabla C \\ &= \nabla(\alpha H_1 + \beta H_2) \times \nabla(\mu H_1 + \nu H_2) \\ &= (\alpha\nu - \beta\mu)\nabla H_1 \times \nabla H_2 \\ &= \nabla H_1 \times \nabla H_2 \,.\end{aligned}$$

■

Remark 6.2.2 *Thus, the real-valued Maxwell-Bloch equations (6.2.1) are unchanged when the conserved functions H_1 and H_2 are replaced by the $SL(2,\mathbb{R})$ combinations H and C. Consequently, the solutions, which represent motion in $\mathbb{R}^3$, will remain the same under reparameterisation of the Hamiltonians by the $SL(2,\mathbb{R})$ group action (6.2.6). Geometrically, the invariance of the trajectories in $\mathbb{R}^3$ means that while the level surfaces of H and C may differ from those of H_1 and H_2, their intersections are exactly the same.*

6.2.1 Lie-Poisson structures

Let us now examine the Lie-Poisson structure of our system. Because of the invariance of (6.2.6) under the above action of $SL(2,\mathbb{R})$, this structure will not be unique. We shall adopt Hamiltonian vector fields as our basic working objects. The correspondence of Hamiltonian vector fields with Poisson brackets is given by the anti-isomorphism (1.11.9) in Lemma 1.11.4. Namely,

$$X_H F = -X_F H = \big\{F,\, H\big\}\,, \tag{6.2.6}$$

which implies by (1.11.9) that

$$X_{\{F,H\}} = -\left[X_F,\, X_H\right]. \tag{6.2.7}$$

In view of (6.2.4) the Hamiltonian vector field for any function $G : \mathbb{R}^3 \to \mathbb{R}$ is expressed as

$$X_G = (\nabla G \times \nabla C)\cdot\nabla\,, \quad \text{with} \quad C = \mu H_1 + \nu H_2\,. \tag{6.2.8}$$

In component form, this is expressed as

$$\begin{aligned} X_G &= (\nu + \mu x_3)\left(\frac{\partial G}{\partial x_2}\frac{\partial}{\partial x_1} - \frac{\partial G}{\partial x_1}\frac{\partial}{\partial x_2}\right) \\ &\quad + \nu x_1\left(\frac{\partial G}{\partial x_3}\frac{\partial}{\partial x_2} - \frac{\partial G}{\partial x_2}\frac{\partial}{\partial x_3}\right) \\ &\quad + \mu x_2\left(\frac{\partial G}{\partial x_1}\frac{\partial}{\partial x_3} - \frac{\partial G}{\partial x_3}\frac{\partial}{\partial x_1}\right), \end{aligned} \tag{6.2.9}$$

and any dynamical quantity Q evolves with time according to

$$\dot{Q} = X_H Q\,, \quad \text{with} \quad X_H = (\nabla H \times \nabla C)\cdot\nabla\,. \tag{6.2.10}$$

Expression (6.2.9) explicitly depends on the parameters μ and ν, through its dependence on the distinguished function, or Casimir C, as prescribed in Proposition 6.2.1. The Hamiltonian function $H = \alpha H_1 + \beta H_2$ also contains real parameters α and β, with the proviso that $\alpha\nu - \beta\mu = 1$. The Lie algebra structure underlying the Poisson structure is determined from the Lie brackets among the Hamiltonian vector fields associated with the coordinate functions x_i, with $i = 1, 2, 3$. These are,

$$\begin{aligned} X_1 &:= X_{x_1} = \mu x_2\partial_3 - (\nu + \mu x_3)\partial_2\,, \\ X_2 &:= X_{x_2} = (\nu + \mu x_3)\partial_1 - \nu x_1\partial_3\,, \\ X_3 &:= X_{x_3} = \nu x_1\partial_2 - \mu x_2\partial_1\,, \end{aligned} \tag{6.2.11}$$

where $\partial_i = \partial/\partial x_i$. The commutators among the Hamiltonian vector fields for the coordinate functions are found to be

$$[X_1, X_2] = -\mu X_3,\ [X_2, X_3] = -\nu X_1,\ [X_3, X_1] = -\mu X_2\,. \tag{6.2.12}$$

The Lie algebra spanned by the divergence-free vector fields X_i with $i = 1, 2, 3$, depends on the values of μ and ν. This dependence is correlated with the type of orbits in the group $SL(2,\mathbb{R})$. Thus, three cases can arise.

Case 1: $\mu = 0$, $\nu \neq 0$. Define

$$Y_1 := -\nu X_1\,, \qquad Y_2 := X_2\,, \qquad Y_3 := X_3\,.$$

The commutation relations of the Lie algebra for this case are

$$[Y_1, Y_2] = 0, \qquad [Y_2, Y_3] = Y_1, \qquad [Y_3, Y_1] = 0\,. \tag{6.2.13}$$

These are the commutation relations of the ***Heisenberg algebra***.

Case 2: $\mu \neq 0$, $\nu = 0$. Define

$$Y_1 := -X_1/\mu\,, \qquad Y_2 := X_2\,, \qquad Y_3 := X_3\,.$$

The commutation relations among the Hamiltonian vector fields for the coordinate functions in this case are

$$[Y_1, Y_2] = Y_3, \qquad [Y_2, Y_3] = 0, \qquad [Y_3, Y_1] = Y_2\,. \tag{6.2.14}$$

These are the commutation relations of the ***Euclidean algebra*** of the plane.

Case 3: $\mu = \epsilon_\mu \neq 0$, $\nu = \epsilon_\nu \neq 0$, where $\epsilon_\sigma = \mathrm{Sign}(\sigma)$. Define the re-scaled Hamiltonian vector fields,

$$Y_1 = -\epsilon_\nu X_1/\left|\mu\right|\,, \quad Y_2 = \frac{X_2}{(|\mu|\,|\nu|)^{1/2}}\,, \quad Y_3 = \frac{X_3}{(|\mu|\,|\nu|)^{1/2}}\,.$$

The commutation relations among the Hamiltonian vector fields in this case are

$$[Y_1, Y_2] = \epsilon Y_3, \qquad [Y_2, Y_3] = Y_1, \qquad [Y_3, Y_1] = \epsilon Y_2\,, \tag{6.2.15}$$

in which $\epsilon = \epsilon_{\mu\nu} = \mathrm{Sign}(\mu\nu)$. Two subcases arise.

Subcase 3a: $\epsilon = 1$. The Lie algebra is isomorphic to $so(3)$.

Subcase 3b: $\epsilon = -1$. The Lie algebra is isomorphic to $so(2,1)$ and $so(1,2)$.

6.2.2 Classes of Casimir functions

Cases 1, 2 and 3 above are each associated with a particular class of Casimir function C.

Case 1: $\mu = 0,\ \nu \neq 0.$

The level sets of C are ***parabolic cylinders*** oriented along the x_2-axis (see Figure 6.1),

$$C = \nu\left(x_3 + \frac{1}{2}x_1^2\right) .$$

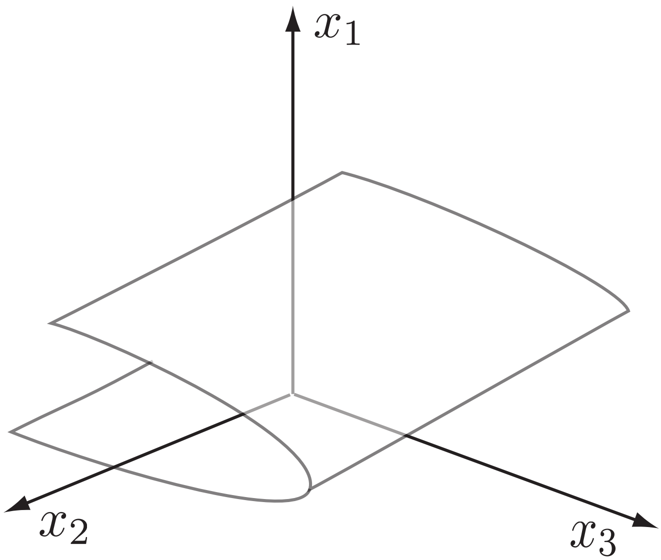


Figure 6.1: Parabolic cylinder level sets for Case 1.

Case 2: $\mu \neq 0,\ \nu = 0.$

For this second case, the level sets are ***circular cylinders*** oriented along the x_1-axis (see Figure 6.2), defined whenever $C/\mu > 0$,

$$C = \frac{1}{2}\mu\left(x_2^2 + x_3^2\right) .$$

Subcase 3a: $\mu \neq 0,\ \nu \neq 0$, with $\mu\nu > 0$. $\mu = \epsilon_\mu \neq 0,\ \nu = \epsilon_\nu \neq 0$, where $\epsilon_\sigma = \mathrm{Sign}(\sigma)$. The level sets are ***ellipsoids of revolution*** (see Figure 6.3), with semimajor axis $r_1 = r$, $r_2 = r_3 =$

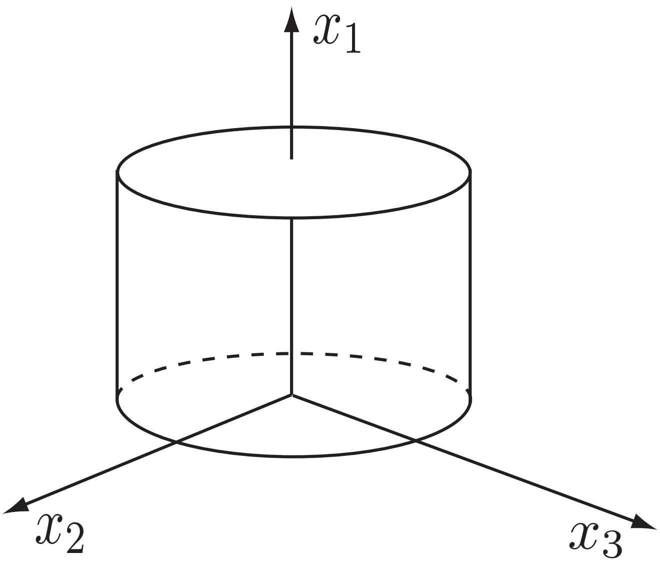


Figure 6.2: Circular cylinder level sets for Case 2.

$(\nu/\mu)^{1/2}r$, centred at $(0, 0, -\nu/\mu)$; they are defined whenever $4\mu C + \nu^2 > 0$.

$$x_1^2 + \frac{\mu}{\nu}\left[x_2^2 + \left(x_3 + \frac{\nu}{\mu}\right)^2\right] = \frac{2C}{\nu} + \frac{\nu}{2\mu} =: r^2 .$$

Subcase 3b: $\mu \neq 0$, $\nu \neq 0$, with $\mu\nu < 0$. These level sets are non-compact surfaces. They are two-sheeted ***hyperboloids of revolution*** if $4\mu C + \nu^2 < 0$, one-sheeted hyperboloids if $4\mu C + \nu^2 > 0$, and a cone whenever $4\mu C + \nu^2 = 0$ (see Figure 6.4). The two varieties of hyperboloids correspond to the two choices of the algebra, either $so(2, 1)$, or $so(1, 2)$.

Remark 6.2.3 *Each of the above classes, in addition to giving us a type for the function C, also gives a prescription for the Hamiltonian function*

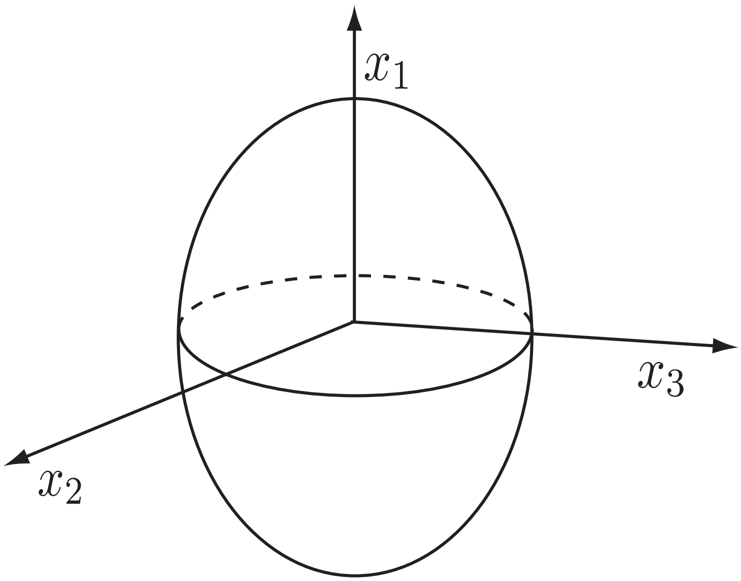


Figure 6.3: Ellipsoidal level sets for Case 3a.

H. In fact, the admissible pairs (H, C) are prescribed by (6.2.6) where the corresponding $SL(2, \mathbb{R})$ matrices are given as follows:

$$
\begin{aligned}
&\textbf{\textit{Case 1}}: \ g_1 = \begin{bmatrix} 1/\nu & \beta \\ 0 & \nu \end{bmatrix}, && H = \frac{H_1}{\nu} + \beta H_2\,, && C = \nu H_2\,, \\
&\textbf{\textit{Case 2}}: \ g_2 = \begin{bmatrix} \alpha & -1/\mu \\ \mu & 0 \end{bmatrix}, && H = \alpha H_1 - \frac{H_2}{\mu}\,, && C = \mu H_1\,, \\
&\textbf{\textit{Case 3}}: \ g_3 = \begin{bmatrix} \alpha & \beta \\ \mu & \nu \end{bmatrix}, && H = \alpha H_1 + \beta H_2, && \mu\nu \neq 0\,,
\end{aligned}
$$

with $\alpha\nu - \beta\mu = 1$ in the last case. In any of these three cases, the locus of the Hamiltonian function H depends on the values of the parameters α or β. Therefore, it is not unique. Indeed, bifurcations may occur as we vary the parameters. For instance, a change in sign may change a level surface of energy from an ellipsoid to a hyperboloid. Nonetheless, the intersections of the level surfaces of C and H do not see these bifurcations. In fact, they

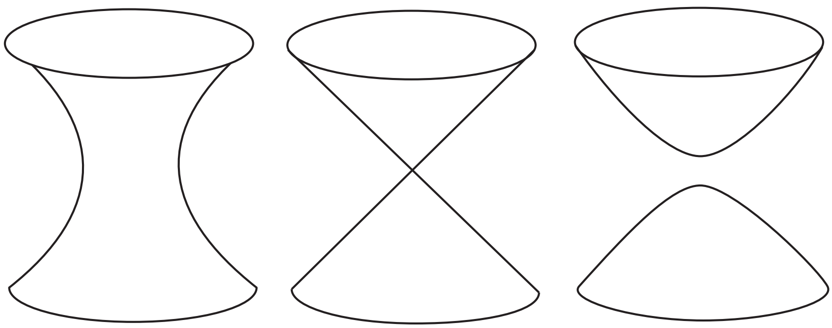


Figure 6.4: Hyperbolic level sets for Case 3b.

do not depend on the parameters at all. However, the representation *of the dynamics does depend on the choice of parameters. For example, as we shall see in the next section, Case 1 yields Duffing oscillator dynamics, while Case 2 yields pendulum dynamics.*

6.3 Reductions to the two-dimensional level sets of the distinguished functions

Each of the three cases presented in Section 6.2 yields a distinct reduction of the real-valued Maxwell-Bloch system (6.1.17) in $\mathbb{R}^3$ to a symplectic system on a two-dimensional manifold specified by a level set of the corresponding Casimir function C. The three reductions give different coordinate representations of the same solutions in $\mathbb{R}^3$. This section illustrates the process of reduction by deriving the reduced phase spaces for Cases 1 and 2.

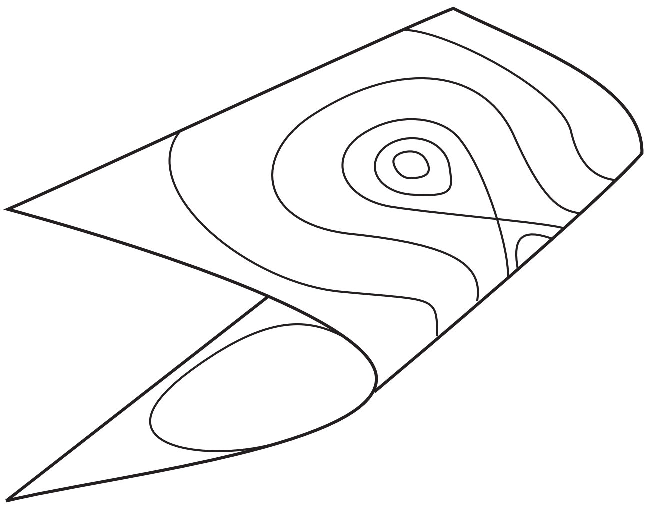


Figure 6.5: Phase portrait for Case 1.

For more details about the nature of reduction and techniques used to perform it, the reader is referred, for example, to [AbMa1978, Ol2000]. For the explicit reductions of the real-valued Maxwell-Bloch system (6.1.17) to symplectic systems on two-dimensional manifolds for the remaining Cases 3a and 3b, see [DaHo1992].

Case 1: $\mu = 0,\, \nu \neq 0.$

A level set of $H_2 = x_3 + \frac{1}{2}x_1^2$ is a ***parabolic cylinder*** oriented along the x_2-direction. On a level set of H_2, one has

$$H_1 = \frac{1}{2}x_2^2 + \frac{1}{2}\left(H_2 - \frac{1}{2}x_1^2\right)^2,$$

so that

$$d^3x = dx_1 \wedge dx_2 \wedge dx_3 = dx_1 \wedge dx_2 \wedge dH_2\,.$$

The $\mathbb{R}^3$ bracket restricts to such a level set as

$$\big\{F,\, H\big\}d^3x = \frac{1}{\nu}dH_2 \wedge \big\{F,\, H_1\big\}_{p-cyl}dx_1 \wedge dx_2\,,$$

where the Poisson bracket on the parabolic cylinder $H_2 = const$ is symplectic,

$$\{F,\, H_1\}_{p-cyl} = \frac{\partial F}{\partial x_1}\frac{\partial H_1}{\partial x_2} - \frac{\partial H_1}{\partial x_1}\frac{\partial F}{\partial x_2}\,.$$

Hence, the equations of motion are

$$\dot{x}_1 \;=\; \frac{\partial H_1}{\partial x_2} = x_2\,, \tag{6.3.1}$$

$$\dot{x}_2 \;=\; -\frac{\partial H_1}{\partial x_1} = x_1\left(H_2 - \frac{1}{2}x_1^2\right). \tag{6.3.2}$$

Therefore, an equation of motion for x_1 emerges, which may be expressed in the form of Newton's Law,

$$\ddot{x}_1 = x_1\left(H_2 - \frac{1}{2}x_1^2\right). \tag{6.3.3}$$

This is the ***Duffing oscillator***. It has critical points at

$$(x_1,\, x_2) = \big(0,\, 0\big) \quad \text{and} \quad \big(\pm\sqrt{2H_2}, 0\big)\,.$$

The first of these critical points is unstable (a saddle point) and the other two are stable (centres). See Figure 6.5.

Case 2: $\mu \neq 0,\, \nu = 0$.
A level set of $H_1 = \frac{1}{2}(x_2^2 + x_3^2)$ is a ***circular cylinder*** oriented along the x_1-direction. On a level set of H_1, one has polar coordinates $(x_2,\, x_3) = (r\cos\varphi,\, r\sin\varphi)$, where $r^2 = 2H_1$. so that

$$d^{\,3}x = dx_1\wedge dx_2\wedge dx_3 = dx_1\wedge dH_1\wedge d\varphi = dH_1\wedge d\varphi\wedge dx_1\,.$$

The $\mathbb{R}^3$ bracket restricts to such a level set as

$$\{F,\, H\}d^{\,3}x = -\frac{1}{\mu}dH_1\wedge\{F,\, H_2\}_{c-cyl}d\varphi\wedge dx_1\,,$$

where the Poisson bracket on the circular cylinder $H_1 = const$ is symplectic,

$$\{F,\, H_2\}_{c-cyl} = \frac{\partial F}{\partial\varphi}\frac{\partial H_1}{\partial x_1} - \frac{\partial H_2}{\partial\varphi}\frac{\partial F}{\partial x_1}\,.$$

Hence, the equations of motion are

$$\dot{\varphi} = \big\{\varphi,\, H_2\big\}_{p-cyl} = \frac{\partial H_2}{\partial x_1}\,, \tag{6.3.4}$$

$$\dot{x}_1 = \big\{x_1,\, H_2\big\}_{p-cyl} = -\frac{\partial H_2}{\partial \varphi}\,, \tag{6.3.5}$$

where

$$H_2 = \frac{1}{2}x_1^2 + x_3 = \frac{1}{2}x_1^2 + \sqrt{2H_1}\sin\varphi\,. \tag{6.3.6}$$

Therefore, an equation of motion for φ results, which may be expressed in the form of Newton's Law,

$$\ddot{\varphi} = -\sqrt{2H_1}\cos\varphi = -\sqrt{2H_1}\sin(\varphi + \pi/2)\,. \tag{6.3.7}$$

This is the equation for a ***simple pendulum*** in the angle $\varphi+\pi/2$. It has critical points at $\varphi + \pi/2 = (0,\, \pi)$. The first of these is stable and the other one is unstable. See Figure 6.6.

6.4 Remarks on geometric phases

The reduced motion takes place on symplectic manifolds, which are the level sets of the constants of motion. The choice of these constants of motion allows considerable freedom in phase space parameterisation. When reconstructing the solutions from the reduced system, especially for periodic solutions, phases arise as a result of travelling over one period in the reduced space (in our case, on a level surface of C). These phases are associated with the group action of the reduction. Let M be a Poisson manifold on which a Lie group G acts as a Hamiltonian Lie group of symmetry transformations with corresponding Lie algebra $\mathfrak{g}$, and let $C : M \to \mathfrak{g}^*$ be the associated momentum map. For the Nambu bracket reductions considered in this chapter, the canonical 1-form is

$$\Theta := p_i dq_i = pdq + Cd\varphi\,, \tag{6.4.1}$$

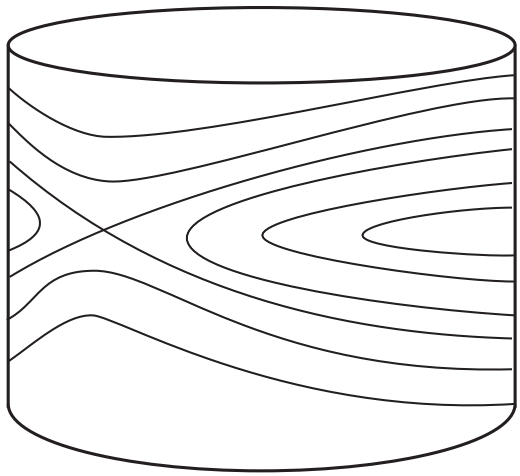


Figure 6.6: Phase portrait for Case 2.

where p and q are the symplectic coordinates for the level surface of C on which the reduced motion takes place. Rearranging gives

$$C d\varphi = -\, pdq + p_i dq_i \,. \tag{6.4.2}$$

The total phase change around a particular periodic orbit on a level set of C may now be obtained by integrating this equation around the orbit and writing the result as the sum of the following two parts:

$$\oint C d\varphi = \underbrace{-\oint pdq}_{\text{Geometric}} + \underbrace{\oint p_i dq_i}_{\text{Dynamic}} = C\Delta\varphi_{\text{geom}} + C\Delta\varphi_{\text{dyn}}. \tag{6.4.3}$$

Hence, we identify

$$\Delta\varphi_{\text{geom}} = -\frac{1}{C}\oint_{\partial S} p\, dq = -\frac{1}{C}\int_S dp \wedge dq \,, \tag{6.4.4}$$

where the second equality uses ***Green's theorem***. (In this formula, ∂S denotes a periodic orbit on the reduced phase space, and S is the surface enclosed by this orbit.) For each of the reductions to symplectic motion on level surfaces of the Casimir functions C described in Sections 6.2 and 6.3, there is a corresponding geometric phase given by (6.4.4), which is proportional to the area enclosed by any periodic orbit. In addition, the dynamical phase is defined by

$$\Delta\varphi_{\text{dyn}} = \frac{1}{C}\oint_{\partial S} p_i\, dq_i = \frac{1}{C}\int_{\partial S} (p\dot{q} + C\dot{\varphi})\, dt\,. \tag{6.4.5}$$

Naturally, the total phase $\Delta\varphi_{\text{tot}} = \Delta\varphi_{\text{dyn}} + \Delta\varphi_{\text{geom}}$ is given by the sum of expressions (6.4.4) and (6.4.5). For Hamiltonian functions that are quadratic in each of the momenta, expression (6.4.5) adopts a particularly simple form,

$$\begin{aligned}
\Delta\varphi_{\text{dyn}} &= \frac{1}{C}\oint_{\partial S}\left(p\frac{\partial H}{\partial p} + C\frac{\partial H}{\partial C}\right) dt \\
&= \frac{1}{C}\oint_{\partial S} 2(H - V)\, dt \\
&= \frac{2T}{C}\left[H - \langle V\rangle\right],
\end{aligned} \tag{6.4.6}$$

where T is the period of the orbit on which the integration is performed and $\langle V\rangle$ denotes the average of the potential energy over the orbit [Mo1991].

Remark 6.4.1 *Compare the geometric phase relation (6.4.6) for quadratic Hamiltonians with formula (4.3.5) for the geometric phase of the optical travelling-wave pulse.*

Clearly, the values of the geometric and dynamics phases, $\Delta\varphi_{\text{geom}}$ and $\Delta\varphi_{\text{dyn}}$, respectively, depend upon the values of the functions H and C. However, the ***total phase*** is an intrinsic property of the orbit and, thus, is independent of any phase space reparameterisations. However, for a given orbit ∂S, one could choose the values of H and C such that $\Delta\varphi_{\text{dyn}} = 0$, and the total phase (for that orbit) would then be completely geometrical. This means the value of the geometric phase is not intrinsic. Rather, its value depends upon the choice of parameterisation of the reduced phase space.

Exercise. Examples of geometric phases for two reductions
Compute the geometric phases for the two reductions by symmetry of the Maxwell-Bloch equations in Section 6.3, one onto a parabolic cylinder and the other onto a circular cylinder. ★

Appendix A

Enhanced coursework

A.1 Problem formulations & selected solutions

Exercise. Formulate the following simple mechanical problems using **(a)** Newton's approach, **(b)** Lagrange's approach and **(c)** Hamilton's approach.

- Bead sliding on a rotating hoop
- Spherical pendulum in polar coordinates
- Charged particle in a magnetic field
- Kaluza-Klein variational principle
- Rigid body with flywheel
- Canonical transformations
- Complex phase space
- A single harmonic oscillator
- Resonant coupled oscillators
- Kepler problem
- Lie derivatives and differential forms
- A cubically nonlinear oscillator

★

A.1.1 The bead sliding on a rotating hoop

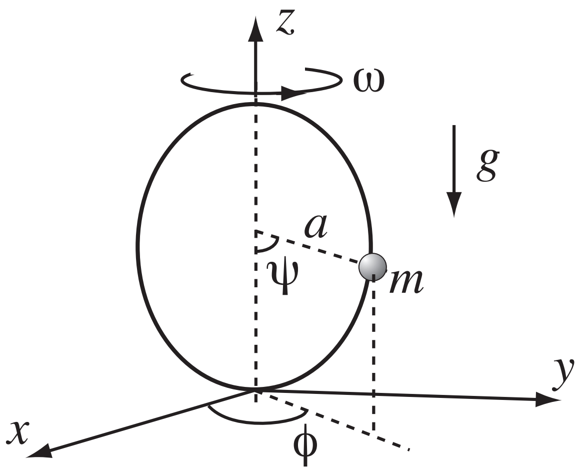


Figure A.1: A bead sliding without friction on a rotating hoop.

Newton's 2nd Law approach $\mathbf{F} = m\mathbf{a}$

Forces

- Gravity: $-mg\hat{\mathbf{k}}$;
- Constraint Forces:
  In directions $\hat{\mathbf{e}}_r$ and $\hat{\mathbf{e}}_\theta$ that keep the particle on the hoop.

In a fixed inertial frame at the centre of the hoop, one chooses spherical coordinates with azimuthal angle $0 \leq \phi < 2\pi$ and polar angle $0 \leq \theta < \pi$ measured from the ***downward*** vertical as in Figure A.1,

$$\begin{aligned} x &= R \sin\theta \cos\phi \,, \\ y &= R \sin\theta \sin\phi \,, \\ z &= -\, R \cos\theta \,. \end{aligned} \tag{A.1.1}$$

The components of the accelerations are

$$
\begin{aligned}
\ddot{x} &= -\omega^2 x - \dot{\theta}^2 x + (R\cos\theta\cos\phi)\ddot{\theta} - 2R\omega\dot{\theta}\cos\theta\sin\phi\,,\\
\ddot{y} &= -\omega^2 y - \dot{\theta}^2 y + (R\cos\theta\sin\phi)\ddot{\theta} + 2R\omega\dot{\theta}\cos\theta\cos\phi\,,\\
\ddot{z} &= -z\dot{\theta}^2 + (R\sin\theta)\ddot{\theta}\,,
\end{aligned}
\tag{A.1.2}
$$

and the unit vector in the positive θ-direction is

$$
\hat{\mathbf{e}}_\theta = \cos\theta\cos\phi\,\hat{\mathbf{i}} + \cos\theta\sin\phi\,\hat{\mathbf{j}} + \sin\theta\,\hat{\mathbf{k}}\,. \tag{A.1.3}
$$

Newton's Law $\mathbf{F} = m\mathbf{a}$ is written in these coordinates, as follows:

- $\hat{\mathbf{e}}_\phi$ and $\hat{\mathbf{e}}_r$ components give constraint forces;
- The equations of motion comes from the $\hat{\mathbf{e}}_\theta$ component of the force, $\mathbf{F}\cdot\hat{\mathbf{e}}_\theta = m\mathbf{a}\cdot\hat{\mathbf{e}}_\theta$.

Using the expression for $\hat{\mathbf{e}}_\theta$ yields:

$$
\begin{aligned}
\mathbf{F}\cdot\hat{\mathbf{e}}_\theta &= -mg\sin\theta\,,\\
m\mathbf{a}\cdot\hat{\mathbf{e}}_\theta &= m(\ddot{x}\cos\theta\cos\phi + \ddot{y}\cos\theta\sin\phi + \ddot{z}\sin\theta)\\
&= mR(\ddot{\theta} - \omega^2\sin\theta\cos\theta)\,.
\end{aligned}
\tag{A.1.4}
$$

Consequently, one finds the Newtonian equation of motion,

$$
R\ddot{\theta} - \omega^2 R\sin\theta\cos\theta + g\sin\theta = 0\,. \tag{A.1.5}
$$

For the case when the rotation vanishes ($\omega = 0$) this simplifies to

$$
R\ddot{\theta} = -\,g\sin\theta\,, \tag{A.1.6}
$$

which not unexpectedly is the equation for the pendulum.

However, the presence of $\omega \neq 0$ acts to reduce the gravitational restoring force as

$$
R\ddot{\theta} = -\,(g - \omega^2 R\cos\theta)\sin\theta\,. \tag{A.1.7}
$$

The Lagrangian Approach

In terms of the velocity,

$$\mathbf{v} = R\dot{\theta}\,\hat{\mathbf{e}}_\theta + (\omega R \sin\theta)\,\hat{\mathbf{e}}_\phi\,, \tag{A.1.8}$$

the kinetic energy of the bead sliding on the hoop is expressed as

$$T = \frac{m}{2}|\mathbf{v}|^2 = \frac{mR^2}{2}(\dot{\theta}^2 + \omega^2 \sin^2\theta)\,, \tag{A.1.9}$$

and its potential energy is

$$V = -\,mgR\cos\theta\,. \tag{A.1.10}$$

The corresponding Lagrangian for this mechanical system is

$$L = T - V = \frac{mR^2}{2}(\dot{\theta}^2 + \omega^2 \sin^2\theta) + mgR\cos\theta\,. \tag{A.1.11}$$

The Euler-Lagrange equation in these coordinates is

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}} = \frac{\partial L}{\partial \theta}\,. \tag{A.1.12}$$

Evaluating each of the partial derivatives yields

$$\begin{aligned} \frac{\partial L}{\partial \dot{\theta}} &= mR^2\dot{\theta}, \\ \frac{\partial L}{\partial \theta} &= mR^2\omega^2\sin\theta\cos\theta - mgR\sin\theta. \end{aligned} \tag{A.1.13}$$

So, on dividing through by mR^2 the motion equation (A.1.5) is recovered,

$$mR^2\ddot{\theta} - mR^2\omega^2\sin\theta\cos\theta + mgR\sin\theta = 0. \tag{A.1.14}$$

Hamilton's Approach

The Legendre transform of the Lagrangian coordinates

$$\pi(\dot{\theta}, \theta) \to (P, \theta) = \Big(\frac{\partial L}{\partial \dot{\theta}}, \theta\Big) \tag{A.1.15}$$

gives the phase space coordinates in which Hamilton's equation is expressed. The canonical angular momentum is

$$P = \frac{\partial L}{\partial \dot{\theta}} = mR^2\dot{\theta}\,, \tag{A.1.16}$$

is related to the angular velocity by

$$\dot{\theta} = \frac{P}{mR^2}\,. \tag{A.1.17}$$

Legendre transforming the Lagrangian to the Hamiltonian gives

$$\begin{aligned} H &= P\dot{\theta} - L \\ &= \frac{P^2}{mR^2} - \frac{P^2}{2mR^2} - \frac{mR^2\omega^2}{2}\sin^2(\theta) - mgR\cos\theta \\ &= \frac{P^2}{mR^2} - mgR\cos\theta - \frac{mR^2\omega^2}{2}\sin^2\theta. \end{aligned} \tag{A.1.18}$$

Remark A.1.1 *Note: This Hamiltonian is not the sum of the kinetic and potential energies of the particle.*

Hamilton's equations are:

$$\begin{aligned} \dot{P} &= -\frac{\partial H}{\partial \theta} = -mgR\sin\theta + mR^2\omega^2\sin\theta\cos\theta\,, \\ \dot{\theta} &= \frac{\partial H}{\partial P} = \frac{P}{mR^2}\,. \end{aligned} \tag{A.1.19}$$

Substituting for the momentum recovers the motion equations (A.1.5) and (A.1.14).

Notice that the motion equation may be rewritten as Newton's Law for an effective potential $V_{eff}(\theta)$ as

$$\ddot{\theta} = -\frac{\partial V_{eff}(\theta)}{\partial \theta}\,, \tag{A.1.20}$$

with effective potential

$$V_{eff}(\theta) = -(g/R)\cos\theta - (\omega^2/2)\sin^2\theta\,. \tag{A.1.21}$$

Equilibrium solutions

In the stationary solutions (equilibria) of these equations, the bead does not move. These are solutions for which $\ddot{\theta} = \dot{\theta} = 0$.

Equilibria

- $\dot{\theta} = 0, \;\; \ddot{\theta} = 0$ gives the equilibrium condition as
$$R\omega^2 \sin\theta\cos\theta = g\sin\theta\,. \tag{A.1.22}$$
- $\theta = 0$ and $\theta = \pi$ are equilibrium solutions corresponding to the particle position at the top and bottom of the hoop.
- If $\theta \neq 0$ or $\theta \neq \pi$ then the equilibrium condition is
$$R\omega^2 \cos\theta = g\,. \tag{A.1.23}$$
- This has two solutions if $g/(R\omega^2) < 1$.
- $\omega_c = \sqrt{g/R}$ is the critical rotation frequency.
- ω_c is also the frequency of linearised oscillations for the simple pendulum
$$R\ddot{\theta} + g\theta = 0\,. \tag{A.1.24}$$
- When $\omega < \omega_c$ there are two solutions $\theta = 0$ and $\theta = \pi$.
- When $\omega > \omega_c$ the solutions are $\theta = 0, \pi \pm \cos^{-1}(g/(R\omega^2))$.
- By plotting the effective potential $V_{eff}(\theta)$ in (A.1.21) one sees that the solution for the bead at the bottom of the hoop ($\theta = 0$) changes stability at a critical frequency, given by
$$\omega^2 < g/R\,.$$
The solution for the bead at the bottom of the hoop is stable (a minimum in potential) for slow angular rotation velocity, below critical. It is unstable (a maximum in potential) when the hoop rotates faster than critical.

- The loss of stability of an equilibrium solution accompanied by the development of of two new stable equilibria is called a **pitchfork bifurcation**. See [MaRa1994] for more discussions and examples of pitchfork bifurcations. Of course, the solution with the bead balancing at the top of the hoop ($\theta = \pi$) is always unstable.

A.1.2 Spherical pendulum in polar coordinates

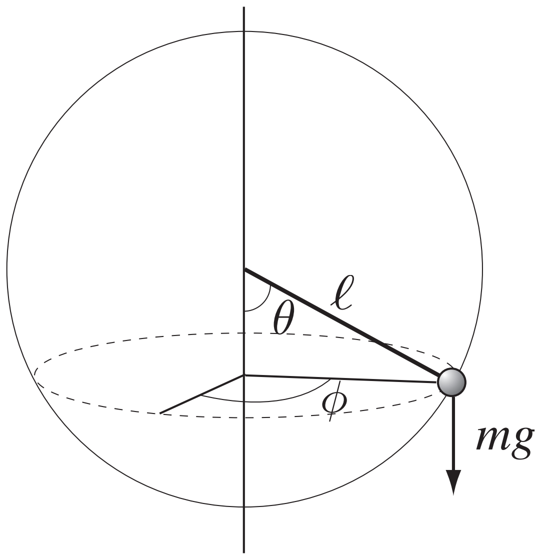


Figure A.2: Spherical pendulum.

Lagrangian approach

As for the bead on the rotating hoop, we use spherical coordinates with azimuthal angle $0 \leq \phi < 2\pi$ and polar angle $0 \leq \theta < \pi$ measured from the ***downward*** vertical defined in terms of Cartesian coordinates by (note minus sign in z)

$$
\begin{aligned}
x &= R\sin\theta\cos\phi\,,\\
y &= R\sin\theta\sin\theta\,,\\
z &= -\,R\cos\theta\,.
\end{aligned}
\tag{A.1.25}
$$

Kinetic Energy

In Cartesian coordinates, the kinetic energy is

$$T = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)\,. \tag{A.1.26}$$

Upon translating into spherical polar coordinates, the velocity components become

$$
\begin{aligned}
\dot{x} &= R\dot{\theta}\cos\theta\cos\phi - R\dot{\phi}\sin\theta\sin\phi\,,\\
\dot{y} &= R\dot{\theta}\cos\theta\sin\phi - R\dot{\phi}\sin\theta\cos\phi\,,\\
\dot{z} &= R\dot{\theta}\sin\theta\,,
\end{aligned}
\tag{A.1.27}
$$

and the kinetic energy becomes

$$T = \frac{mR^2}{2}(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta)\,. \tag{A.1.28}$$

This is the standard form for the particle kinetic energy in spherical coordinates.

Potential Energy

The potential energy of the spherical pendulum is

$$V = mgz = -\,mgR\cos\theta\,. \tag{A.1.29}$$

Lagrangian

Its Lagrangian is similar to equation (A.1.11) for the rotating hoop

$$L = T - V = \frac{mR^2}{2}(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta) + mgR\cos\theta\,. \tag{A.1.30}$$

- θ **equation**

  The Euler-Lagrange equation is

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} = 0\,, \tag{A.1.31}$$

  in which for the spherical pendulum Lagrangian,

$$\begin{aligned} \frac{\partial L}{\partial \dot{\theta}} &= mR^2\dot{\theta}\,, \\ \frac{\partial L}{\partial \theta} &= \dot{\phi}^2 mR^2 \sin\theta\cos\theta - mgR\sin\theta\,. \end{aligned} \tag{A.1.32}$$

  Consequently, one finds an equation similar in form to the motion equation (A.1.5) or (A.1.14) for the bead on the rotating hoop,

$$mR^2\ddot{\theta} - \dot{\phi}^2 mR^2 \sin\theta\cos\theta + mgR\sin\theta = 0\,. \tag{A.1.33}$$

- ϕ **equation** The Euler-Lagrange equation in ϕ is

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} - \frac{\partial L}{\partial \phi} = 0\,. \tag{A.1.34}$$

  Consequently, one computes that

$$\begin{aligned} \frac{\partial L}{\partial \dot{\phi}} &= mR^2\dot{\phi}\sin^2\theta\,, \\ \frac{\partial L}{\partial \phi} &= 0\,, \\ \frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} &= \frac{d}{dt}(mR^2\dot{\phi}\sin^2\theta) = 0\,. \end{aligned} \tag{A.1.35}$$

  Thus, as guaranteed by Noether's theorem, azimuthal symmetry of the Lagrangian (that is, L being independent of ϕ) implies conservation of azimuthal angular momentum.

Hamiltonian approach

One computes the canonical momenta and solves for velocities in terms of momenta and coordinates as:

$$\begin{aligned} P_\theta &= \frac{\partial L}{\partial \dot\theta} = mR^2\dot\theta\,, \qquad \text{so} \qquad \dot\theta = \frac{P_\theta}{mR^2}\,, \\ P_\phi &= \frac{\partial L}{\partial \dot\phi} = mR^2\dot\phi\sin^2\theta\,, \qquad \text{so} \qquad \dot\phi = \frac{P_\phi}{mR^2\sin^2\theta}\,. \end{aligned} \tag{A.1.36}$$

The Hamiltonian is obtained by the Legendre transforming the Lagrangian, as

$$\begin{aligned} H &= P_\theta\dot\theta + P_\phi\dot\phi - L \\ &= \frac{P_\theta^2}{2mR^2} + \frac{P_\phi^2}{2mR^2\sin^2\theta} - mgl\cos\theta\,. \end{aligned}$$

This Hamiltonian has canonical motion equations,

$$\begin{aligned} \dot P_\theta &= -\frac{\partial H}{\partial\theta} = -\,mgl\sin\theta + \frac{P_\phi^2}{mR^2}\,\frac{\cos\theta}{\sin^3\theta}\,, \\ \dot P_\phi &= -\frac{\partial H}{\partial\phi} = 0\,. \end{aligned}$$

The angular frequencies are recovered in their canonical form from the Hamiltonian as

$$\begin{aligned} \dot\theta &= \frac{\partial H}{\partial P_\theta} = \frac{P_\theta}{mR^2}\,, \\ \dot\phi &= \frac{\partial H}{\partial P_\phi} = \frac{P_\phi}{mR^2\sin^2\theta}\,. \end{aligned}$$

By Noether's theorem, the azimuthal angular momentum P_ϕ is conserved because the Lagrangian (and hence the Hamiltonian) of the spherical pendulum are independent of ϕ. This symmetry also allows further progress toward characterising the spherical pendular motion. In particular, the equilibria are azimuthally symmetric.

Substituting for $\dot\phi^2$ in equation (A.1.33) from (A.1.36) yields

$$mR^2\ddot\theta = -\,mgR\sin\theta + \Big(\frac{P_\phi}{mR^2\sin^2\theta}\Big)^2 mR^2\sin\theta\cos\theta\,. \tag{A.1.37}$$

This may be rewritten in terms of an effective potential $V_{eff}(\theta)$ as

$$\ddot{\theta} = -\,\frac{\partial V_{eff}(\theta)}{\partial \theta}\,, \tag{A.1.38}$$

with

$$V_{eff}(\theta) = -\,(g/R)\cos\theta + \frac{P_\phi^2}{2mR^2\sin^2\theta}\,. \tag{A.1.39}$$

This approach enables a phase plane analysis in $(\theta,\ P_\theta)$. Combining this with conservation of energy defined as

$$E/(mR^2) = \dot{\theta}^2/2 + V_{eff}(\theta)\,,$$

leads easily to a solution for $\theta(t)$ as a quadrature integral,

$$t - t_0 = \frac{1}{\sqrt{2}}\int \frac{d\theta}{\sqrt{E/(mR^2) - V_{eff}(\theta)}}\,.$$

In principle, a formula for the solution for the azimuthal angle $\phi(t)$ would follow via another quadrature obtained from the conservation of angular momentum P_ϕ. However, in practice, one may wish to resort to a more direct approach, such as numerical integration of (A.1.39).

A.1.3 Charged particle in a given magnetic field

The problem is to formulate the equations of motion for an electrically charged particle in a given magnetic field, first by using the traditional *minimal coupling method* in physics, and then by using the *Kaluza-Klein method*. In the latter approach, the coupling constant between the field and the particle (its charge) emerges as a conserved momentum that is canonically conjugate to an *internal* S^1 *dimension* associated with the particle. In this approach the equations are reinterpreted as geodesic motion with respect to the Kaluza-Klein metric defined on the tangent bundle $TQ_{KK} = \mathbb{R}^3 \times S^1$.

Lagrangian approach: minimal coupling

Consider a particle of charge e and mass m moving in a magnetic field $\mathbf{B}$, where $\mathbf{B} = \nabla \times \mathbf{A}$ is a given magnetic field on $\mathbb{R}^3$. The Lagrangian for the motion is given by the *minimal coupling prescription*

$$L(\mathbf{q}, \dot{\mathbf{q}}) = \frac{m}{2}\|\dot{\mathbf{q}}\|^2 + \frac{e}{c}\dot{\mathbf{q}} \cdot \mathbf{A}(\mathbf{q})\,, \tag{A.1.40}$$

in which the constant c is the speed of light and $\|\dot{\mathbf{q}}\|^2 = \dot{\mathbf{q}} \cdot \dot{\mathbf{q}}$ in $\mathbb{R}^3$. This is the $\mathbf{J} \cdot \mathbf{A}$ (jay-dot-ay) prescription with current $\mathbf{J} = e\dot{\mathbf{q}}/c$. The derivatives of this Lagrangian are

$$\frac{\partial L}{\partial \dot{\mathbf{q}}} = m\dot{\mathbf{q}} + \frac{e}{c}\mathbf{A} =: \mathbf{p} \quad \text{and} \quad \frac{\partial L}{\partial \mathbf{q}} = \frac{e}{c}\nabla \mathbf{A}^T \cdot \dot{\mathbf{q}}\,,$$

in which $\nabla \mathbf{A}^T \cdot \dot{\mathbf{q}} = \sum_j \dot{q}_j \nabla A^j$. Hence, the Euler-Lagrange equation for this system is, with notation $\nabla \mathbf{A} \cdot \dot{\mathbf{q}} = \dot{q}^j \partial \mathbf{A}/\partial x^j$,

$$m\,\ddot{\mathbf{q}} = \frac{e}{c}(\nabla \mathbf{A}^T \cdot \dot{\mathbf{q}} - \nabla \mathbf{A} \cdot \dot{\mathbf{q}}) = \frac{e}{c}\,\dot{\mathbf{q}} \times \mathbf{B}\,, \tag{A.1.41}$$

which is ***Newton's equation for the Lorentz force***.

The Lagrangian L in equation (A.1.40) is hyperregular, because the fibre derivative

$$\mathbf{p} = \frac{\partial L(\mathbf{q}, \dot{\mathbf{q}})}{\partial \dot{\mathbf{q}}} = m\dot{\mathbf{q}} + \frac{e}{c}\mathbf{A}(\mathbf{q})$$

is linearly invertible

$$\dot{\mathbf{q}} = \frac{\partial H(\mathbf{q}, \mathbf{p})}{\partial \mathbf{p}} = \frac{1}{m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}(\mathbf{q})\right).$$

The corresponding Hamiltonian is given by the invertible change of variables,

$$H(\mathbf{q}, \mathbf{p}) = \mathbf{p} \cdot \dot{\mathbf{q}} - L(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2m}\left\|\mathbf{p} - \frac{e}{c}\mathbf{A}(\mathbf{q})\right\|^2. \tag{A.1.42}$$

Finally, the canonical equations for this Hamiltonian recover Newton's equations (A.1.41) for the Lorentz force law.

Charged particle, magnetic field, Kaluza-Klein construction

Although the minimal-coupling Lagrangian in equation (A.1.40) is not expressed as the kinetic energy of a metric, Newton's equations for the Lorentz force law (A.1.41) may still be obtained as geodesic equations. This is accomplished by suspending them in a higher dimensional space by using the ***Kaluza-Klein construction***, which proceeds as follows.

Let Q_{KK} be the manifold $\mathbb{R}^3 \times S^1$ with variables $(\mathbf{q}, \theta)$. On Q_{KK} introduce the ***Kaluza-Klein Lagrangian*** $L_{KK} : TQ_{KK} \simeq T\mathbb{R}^3 \times TS^1 \mapsto \mathbb{R}$ as

$$L_{KK}(\mathbf{q}, \theta, \dot{\mathbf{q}}, \dot{\theta}) = \frac{1}{2}m\|\dot{\mathbf{q}}\|^2 + \frac{1}{2}\left(\mathbf{A} \cdot \dot{\mathbf{q}} + \dot{\theta}\right)^2 . \tag{A.1.43}$$

The Lagrangian L_{KK} is positive definite in $(\dot{\mathbf{q}}, \dot{\theta})$; so it may be regarded as the kinetic energy of a metric, the ***Kaluza-Klein metric*** on TQ_{KK}. This means the Euler-Lagrange equation for L_{KK} will be the geodesic equation of this metric for integral curves on $Q_{KK} = \mathbb{R}^3 \times S^1$. (This construction fits the idea of $U(1)$ gauge symmetry for electromagnetic fields in $\mathbb{R}^3$. It could be generalized substantially, but this would take us beyond our present scope.) The Legendre transformation for L_{KK} gives the momenta

$$\mathbf{p} = m\dot{\mathbf{q}} + (\mathbf{A} \cdot \dot{\mathbf{q}} + \dot{\theta})\mathbf{A} \qquad \text{and} \qquad \pi = \mathbf{A} \cdot \dot{\mathbf{q}} + \dot{\theta}. \tag{A.1.44}$$

Since L_{KK} does not depend on θ, the Euler-Lagrange equation

$$\frac{d}{dt}\frac{\partial L_{KK}}{\partial \dot{\theta}} = \frac{\partial L_{KK}}{\partial \theta} = 0 \,,$$

shows that $\pi = \partial L_{KK}/\partial\dot{\theta}$ is conserved. The ***charge*** is now defined by $e := c\pi$. The Hamiltonian H_{KK} associated to L_{KK} by the Legendre transformation (A.1.44) is

$$\begin{aligned}
H_{KK}(\mathbf{q},\theta,\mathbf{p},\pi) &= \mathbf{p}\cdot\dot{\mathbf{q}} + \pi\dot{\theta} - L_{KK}(\mathbf{q},\dot{\mathbf{q}},\theta,\dot{\theta}) \\
&= \mathbf{p}\cdot\frac{1}{m}(\mathbf{p}-\pi\mathbf{A}) + \pi(\pi - \mathbf{A}\cdot\dot{\mathbf{q}}) \\
&\quad - \frac{1}{2}m\|\dot{\mathbf{q}}\|^2 - \frac{1}{2}\pi^2 \\
&= \mathbf{p}\cdot\frac{1}{m}(\mathbf{p}-\pi\mathbf{A}) + \frac{1}{2}\pi^2 \\
&\quad - \pi\mathbf{A}\cdot\frac{1}{m}(\mathbf{p}-\pi\mathbf{A}) - \frac{1}{2m}\|\mathbf{p}-\pi\mathbf{A}\|^2 \\
&= \frac{1}{2m}\|\mathbf{p}-\pi\mathbf{A}\|^2 + \frac{1}{2}\pi^2. \qquad \text{(A.1.45)}
\end{aligned}$$

On the constant level set $\pi = e/c$, the Kaluza-Klein Hamiltonian H_{KK} is a function of only the variables $(\mathbf{q},\mathbf{p})$ and is equal to the Hamiltonian (A.1.42) for charged particle motion under the Lorentz force up to an additive constant. This example provides an easy but fundamental illustration of the geometry of (Lagrangian) reduction by symmetry. The canonical equations for the Kaluza-Klein Hamiltonian H_{KK} now reproduce Newton's equations for the Lorentz force law, reinterpreted as geodesic motion with respect to the Kaluza-Klein metric defined on the tangent bundle TQ_{KK} in (A.1.43).

A.1.4 Rigid body with flywheel: Lagrange gyrostat

Formulation of the problem

Rigid body with flywheel

Formulate and analyse the equations of motion for a rigid body that has a flywheel attached and aligned with its intermediate principal axis. The kinetic energy for this system is given as

$$KE = \frac{1}{2}\lambda_1\Omega_1^2 + \frac{1}{2}I_2\Omega_2^2 + \frac{1}{2}\lambda_3\Omega_3^2 + \frac{1}{2}J_2(\dot{\alpha}+\Omega_2)^2\,,$$

where $\boldsymbol{\Omega} = (\Omega_1,\Omega_2,\Omega_3)$ is the angular velocity vector, $\dot{\alpha}$ is the rotational frequency of the flywheel about the intermediate principal axis, and λ_1, I_2, J_2, λ_3 are positive constants.

- Write the Lagrangian for this system.
- Find the angular momenta $\mathbf{\Pi} \in \mathbb{R}^3$ and $\ell_2 \in \mathbb{R}^1$.
- Legendre transform to obtain the Hamiltonian and Poisson bracket in the variables $\mathbf{\Pi}, \ell_2, \alpha$.
- Write the equations of motion for this system.
- Consider the case when $I_2 = \lambda_1$, write the equation for Π_3 in the Newtonian form,

$$\frac{d^2\Pi_3}{dt^2} = -\frac{\partial}{\partial\Pi_3}V(\Pi_3)\,,$$

  and perform the phase plane analysis required to understand the bifurcation of the solutions of this system. What sorts of bifurcations are available for it?

Analysis

This problem integrates the various approaches of Newton, Lagrange and Hamilton, and takes insight from each approach.

Just as for the isolated rigid body, the energy is purely kinetic; so one may define the kinetic energy Lagrangian for this system $L : TSO(3)/SO(3) \times TS^1 \to \mathbb{R}^3$ as

$$L(\mathbf{\Omega},\, \dot\alpha) = \frac{1}{2}\lambda_1\Omega_1^2 + \frac{1}{2}I_2\Omega_2^2 + \frac{1}{2}\lambda_3\Omega_3^2 + \frac{1}{2}J_2(\dot\alpha + \Omega_2)^2\,,$$

where $\mathbf{\Omega} = (\Omega_1, \Omega_2, \Omega_3)$ is the angular velocity vector, $\dot\alpha$ is the rotational frequency of the flywheel about the intermediate principal axis, and λ_1, I_2, J_2, λ_3 are positive constants corresponding to the principle moments of inertia, including the presence of the flywheel. Because the Lagrangian is independent of the angle α, its canonically conjugate angular momentum, $\ell_2 := \partial L/\partial\dot\alpha$ will be conserved. This suggests a move into the Hamiltonian picture, where the conserved ℓ_2 will become a constant parameter.

- Legendre transforming this Lagrangian allows us to express its Hamiltonian in terms of the angular momenta $\mathbf{\Pi} = \partial L/\partial \mathbf{\Omega} \in \mathbb{R}^3$ and $\ell_2 = \partial L/\partial \dot{\alpha} \in \mathbb{R}^1$ of the rigid body and flywheel, respectively,

$$\begin{aligned} H(\mathbf{\Pi}, \ell_2) &= \mathbf{\Pi} \cdot \mathbf{\Omega} + \ell_2 \dot{\alpha} - L(\mathbf{\Omega},\, \dot{\alpha}) \qquad \text{(A.1.46)} \\ &= \frac{\Pi_1^2}{2\lambda_1} + \frac{\Pi_3^2}{2\lambda_3} + \underbrace{\frac{1}{2I_2}\big(\Pi_2 - \ell_2\big)^2}_{\text{Offset along } \Pi_2} + \frac{\ell_2^2}{2J_2} \,. \end{aligned}$$

This Hamiltonian is an ellipsoid in coordinates $\mathbf{\Pi} \in \mathbb{R}^3$, whose centre is *offset* in the Π_2-direction by an amount equal to the conserved angular momentum ℓ_2 of the flywheel.

The offset of the energy ellipsoid by ℓ_2 along the Π_2-axis radically alters its intersections with the angular momentum sphere $|\mathbf{\Pi}| = const.$ Its dynamical behaviour governed as motion along these altered intersections is quite different from the rigid body, which has no offset of its energy ellipsoid.

In fact, the dynamics of the rigid body and flywheel system is almost identical to the dynamics of the optical travelling-wave pulses in Section 4.5. Figures 4.6 and 4.7 are particularly useful for envisoning the choreography of bifurcations available to the rigid body and flywheel system.

- The Poisson bracket in the variables

$$\mathbf{\Pi}, \ell_2, \alpha \in so(3)^* \times T^*S^1 \,,$$

is a direct sum of the rigid body bracket for $\mathbf{\Pi} \in so(3)^* \simeq \mathbb{R}^3$ and the canonical bracket for the flywheel phase space coordinates $(\ell_2, \alpha) \in T^*S^1$.

$$\begin{aligned} \{F\,,\,H\} &= -\,\mathbf{\Pi} \cdot \left(\frac{\partial F}{\partial \mathbf{\Pi}} \times \frac{\partial H}{\partial \mathbf{\Pi}} \right) + \frac{\partial F}{\partial \alpha}\frac{\partial H}{\partial \ell_2} - \frac{\partial H}{\partial \alpha}\frac{\partial F}{\partial \ell_2} \\ &= -\,\epsilon_{klm}\,\Pi_k\,\frac{\partial F}{\partial \Pi_l}\,\frac{\partial H}{\partial \Pi_m} + \frac{\partial F}{\partial \alpha}\frac{\partial H}{\partial \ell_2} - \frac{\partial H}{\partial \alpha}\frac{\partial F}{\partial \ell_2} \,. \end{aligned}$$

- The Hamiltonian equations of motion for this system are, thus,

$$\begin{aligned}
\dot{\Pi}_1 = \{\dot{\Pi}_1,\, H\} &= \Big(\frac{1}{\lambda_3} - \frac{1}{I_2}\Big)\Pi_2\Pi_3 + \frac{\ell_2}{I_2}\Pi_3\,,\\
\dot{\Pi}_2 = \{\dot{\Pi}_2,\, H\} &= \Big(\frac{1}{\lambda_1} - \frac{1}{\lambda_3}\Big)\Pi_1\Pi_3\,,\\
\dot{\Pi}_3 = \{\dot{\Pi}_3,\, H\} &= \Big(\frac{1}{I_2} - \frac{1}{\lambda_1}\Big)\Pi_1\Pi_2 - \frac{\ell_2}{I_2}\Pi_1\,,\\
\dot{\ell}_2 = \{\ell_2,\, H\} &= 0\,,\\
\dot{\alpha} = \{\alpha,\, H\} &= \frac{\partial H}{\partial \ell_2} = \frac{\ell_2}{J_2} - \frac{1}{I_2}\big(\Pi_2 - \ell_2\big)\\
&= \frac{\ell_2}{J_2} - \Omega_2\,.
\end{aligned}$$

- The constants of motion for this system are ℓ_2, $|\mathbf{\Pi}|$ and H. The system is integrable in the Π_i variables. The dynamics of the flywheel angle α decouples from the rest and may be found separately, after solving for Π_i, $i = 1, 2, 3$.

- We consider the case when $I_2 = \lambda_1$, for which the offset energy ellipsoid in (A.1.46) is cylindrically symmetric about an axis parallel to the Π_3-axis. In this case, the motion equation for Π_3 will simplify to Newtonian form.

  When $I_2 = \lambda_1$, the quadratic rigid body nonlinearity in the Π_3-equation vanishes and its dynamics simplifies to

$$\ddot{\Pi}_3 = -\frac{\ell_2}{I_2}\dot{\Pi}_1 = \bigg(-\frac{\ell_2^2}{\lambda_1^2} + \Big(\frac{1}{\lambda_1} - \frac{1}{\lambda_3}\Big)\frac{\ell_2\Pi_2}{\lambda_1}\bigg)\Pi_3\,.$$

  The Hamiltonian in this case produces a linear relationship between Π_2 and Π_3^2, whose coefficients involve only system constants and constants of the motion, namely,

$$\begin{aligned}
H &= \frac{1}{2\lambda_1}|\mathbf{\Pi}|^2 + \frac{1}{2}\Big(\frac{1}{\lambda_3} - \frac{1}{\lambda_1}\Big)\Pi_3^2\\
&\quad - \frac{\Pi_2\ell_2}{\lambda_1} + \frac{1}{2}\Big(\frac{1}{\lambda_1} + \frac{1}{J_2}\Big)\ell_2^2\,.
\end{aligned}$$

One may rearrange this relationship into the following form, that will conveniently allow Π_2 to be eliminated in the equation for Π_3. This form is,

$$\Big(\frac{1}{\lambda_1}-\frac{1}{\lambda_3}\Big)\frac{\ell_2\Pi_2}{\lambda_1}=A-\frac{1}{2}\Big(\frac{1}{\lambda_3}-\frac{1}{\lambda_1}\Big)^2\Pi_3^2\,,$$

where A is a constant of motion given by

$$A=\Big(\frac{1}{\lambda_1}-\frac{1}{\lambda_3}\Big)\bigg(\frac{1}{2\lambda_1}|\mathbf{\Pi}|^2-H+\frac{1}{2}\Big(\frac{1}{\lambda_1}+\frac{1}{J_2}\Big)\ell_2^2\bigg)\,.$$

Upon eliminating Π_2 from the equation above for Π_3, one finds the following second order differential equation involving Π_3 alone,

$$\ddot{\Pi}_3=\bigg(A-\frac{\ell_2^2}{\lambda_1^2}-\frac{1}{2}\Big(\frac{1}{\lambda_3}-\frac{1}{\lambda_1}\Big)^2\Pi_3^2\bigg)\Pi_3\,. \qquad \text{(A.1.47)}$$

- ***Discussion of equation (A.1.47)***
  Equation (A.1.47) may be expressed in Newtonian form,

$$\ddot{\Pi}_3=-\frac{\partial}{\partial\Pi_3}V(\Pi_3)\,,$$

with a quartic potential

$$V(\Pi_3)=-\Big(A-\frac{\ell_2^2}{\lambda_1^2}\Big)\Pi_3^2\;+\;\frac{1}{8}\Big(\frac{1}{\lambda_3}-\frac{1}{\lambda_1}\Big)^2\Pi_3^4\,,$$

and a conserved energy

$$E=\frac{1}{2}\dot{\Pi}_3^2+V(\Pi_3)\,.$$

Equation (A.1.47) for Π_3 has equilibria at

$$\Pi_3=0\quad\text{and}\quad\Pi_3=\frac{\pm\lambda_1\lambda_3\sqrt{2(A\lambda_1^2-\ell_2^2)}}{\lambda_1-\lambda_3}\,.$$

It has one unstable equilibrium at $\Pi_3=0$ and two stable equilibria away from $\Pi_3=0$ when A is sufficiently large and positive. Otherwise, it has only one stable equilibrium at $\Pi_3=0$.

Equation (A.1.47) has a pitchfork bifurcation when $A\lambda_1^2 - \ell_2^2$ passes through zero. At this point, the equilibrium at $\Pi_3 = 0$ loses its stability as two new stable equilibria are created on either side of it. This pitchfork bifurcation may be studied and illustrated using classical methods of phase plane analysis.

- Equation (A.1.47) is an example of the famous ***Duffing oscillator equation***, which has been very well studied.

  For example, he text [GuHo1983] provides a wide-ranging treatment of various aspects of the Duffing equation, including the chaotic response of its solutions to periodic perturbations.

- See [Ko1984] for further analysis of this problem, including its development of Hamiltonian chaos under periodic perturbations. Homoclinic chaos certainly involves geometric mechanics, but this topic is beyond the scope of the present text.

A.2 Introduction to oscillatory motion

A.2.1 Criteria for canonical transformations

(A) Let $\mathbb{Z} = (\mathbb{P}, \mathbb{Q})$ and $z = (p, q)$. The canonical Poisson bracket is

$$\{F, H\}_{\mathbb{Z}} = \left(\frac{\partial F}{\partial z}\right)^T \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \frac{\partial H}{\partial z}\,. \qquad \text{(A.2.1)}$$

Under a phase space transformation $z \to \mathbb{Z}(z)$, the chain rule is used to transform Poisson brackets as

$$\begin{aligned} \{F, H\}_{\mathbb{Z}} &= \left(\frac{\partial F}{\partial \mathbb{Z}}\right)^T \left(\frac{\partial \mathbb{Z}}{\partial z}\right)^T \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \frac{\partial \mathbb{Z}}{\partial z}\frac{\partial H}{\partial \mathbb{Z}} \\ &= \left(\frac{\partial F}{\partial \mathbb{Z}}\right)^T \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \frac{\partial H}{\partial \mathbb{Z}}\,. \end{aligned} \qquad \text{(A.2.2)}$$

Therefore, the canonical Poisson bracket will be preserved, provided

$$\left(\frac{\partial \mathbb{Z}}{\partial z}\right)^T \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \frac{\partial \mathbb{Z}}{\partial z} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} . \tag{A.2.3}$$

$2N \times 2N$ matrices S satisfying

$$S^T J S = J, \quad \text{with } J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{A.2.4}$$

are said to be ***symplectic*** and are denoted $S \in Sp(2N)$.

Remark A.2.1 *According to (A.2.3) a phase space transformation $T^*\mathbb{R}^N \to T^*\mathbb{R}^N$ given by $z \to \mathbb{Z}(z)$ is canonical, if its Jacobian is symplectic; that is, if $\partial\mathbb{Z}/\partial z \in Sp(2N)$.*

(B) If the transformed one-form satisfies $\mathbb{P}d\mathbb{Q} - pdq = dF$, then

$$\oint_c \mathbb{P}d\mathbb{Q} = \oint_c pdq \,, \tag{A.2.5}$$

for every closed contour which is contractible to a point. By Stokes theorem, this implies

$$\int_A d\mathbb{Q} \wedge d\mathbb{P} = \int_A dq \wedge dp \,, \tag{A.2.6}$$

for a surface whose boundary forms the closed contour, $\partial A = c$. Consequently, a transformation that changes the canonical one-form by an exact form preserves the area in phase space defined by the symplectic two-form.

Remark A.2.2 *A transformation that multiplies the canonical one-form by a constant may also be regarded as canonical, because the constant may be simply absorbed, e.g., into the units of time.*

Exercise. Show that the following transformations of coordinates are each canonical.

1. $T^*\mathbb{R}^2/\{0\} \to T^*\mathbb{R}_+ \times T^*S^1$, given by

$$x + iy = re^{i\theta}, \quad p_x + ip_y = (p_r + ip_\theta/r)e^{i\theta}.$$

2. $T^*\mathbb{R}^2 \to \mathbb{C}^2$, given with $k = 1, 2$ by

$$a_k = q_k + ip_k, \quad a_k^* = q_k - ip_k.$$

3. $T^*\mathbb{R}^2/\{0\} \to \mathbb{R}_+^2 \times T^2$, given with $k = 1, 2$ by

$$I_k = \frac{1}{2}(q_k^2 + p_k^2), \quad \phi_k = \tan^{-1}(p_k/q_k).$$

★

Answer.

1. Evaluating the real part, $\Re\big((p_x + ip_y)(dx - idy)\big) = \Re\big((p_r + ip_\theta/r)(dr - ird\theta)\big)$ yields

$$p_x dx + p_y dy = p_r dr + p_\theta d\theta.$$

Taking exterior derivative gives

$$dp_x \wedge dx + dp_y \wedge dy = dp_r \wedge dr + dp_\theta \wedge d\theta.$$

2. $da \wedge da^* = (dq_k + idp_k) \wedge (dq_k - idp_k) = -2idq \wedge dp$

3. For each k (no sum) define

$$a_k = q_k + ip_k, \quad r_k = |a_k| = (q_k^2 + p_k^2)^{1/2}, \quad \phi_k = \tan(p_k/q_k)$$

Then for each k (no sum) with $I_k = \frac{1}{2}r_k^2$ and

$$\begin{aligned} dI_k \wedge d\phi_k &= \tfrac{1}{2}dr_k^2 \wedge d\phi_k \\ &= (q_k dq_k + p_k dp_k) \wedge \frac{q_k dp_k - p_k dq_k}{q_k^2 + p_k^2} \\ &= \frac{q_k^2 dq_k \wedge dp_k - p_k^2 dp_k \wedge dq_k}{q_k^2 + p_k^2} \\ &= dq_k \wedge dp_k \end{aligned}$$

▲

Summary

The symplectic form on an even-dimensional manifold M is a closed nondegenerate 2-form $\omega = dq \wedge dp$. Canonical transformations preserve the symplectic structure.

Theorem 1 (Darboux) *Local coordinates always exist in which ω may be written in the canonical form*

$$\omega = \sum_{j=1}^{N} dq_j \wedge dp_j \, . \tag{A.2.7}$$

Here ω provides a linear isomorphism from $TM \to T^*M$

$$\omega(X_H, \cdot) = dH \, , \tag{A.2.8}$$

where X_H is a vector field generated by H, the Hamiltonian. The inverse mapping $T^*M \to TM$ is

$$\begin{aligned} X_H &= \{\, \cdot \, , \, H\} \\ &= H_p \frac{\partial}{\partial q} - H_q \frac{\partial}{\partial p} \\ &= \dot{q} \frac{\partial}{\partial q} + \dot{p} \frac{\partial}{\partial p} \, . \end{aligned} \tag{A.2.9}$$

Check:

$$\begin{aligned} \omega(X_H, \cdot) &= X_H \,\lrcorner\, (dq \wedge dp) \\ &= \dot{q} dp - \dot{p} dq \\ &= H_p dp + H_q dq \\ &= dH \, . \end{aligned} \tag{A.2.10}$$

Theorem A.2.3 (Poincaré's theorem) *Flow by Hamilton's canonical equations occurs as a canonical transformation.*

Proof.

$$\begin{aligned}\frac{d}{dt}dq \wedge dp &= d\dot{q} \wedge dp + dq \wedge d\dot{p} \\ &= dH_p \wedge dp - dq \wedge dH_q \\ &= d(H_p dp + H_q dq) \\ &= d^2 H = 0\,.\end{aligned} \tag{A.2.11}$$

Hence, we have proved Poincaré's theorem,

$$dp(t) \wedge dq(t) = dp(0) \wedge dq(0)\,, \tag{A.2.12}$$

for Hamiltonian flows. ■

Definition A.2.4 (Phase Space Volume) *The phase space volume is*

$$dq_1 \wedge dp_1 \wedge \cdots \wedge dq_n \wedge dp_n = dVol\,. \tag{A.2.13}$$

For Hamiltonian flows, each $dq_i \wedge dp_i$ is preserved. Hence, so is $dVol$. This is known as ***Liouville's theorem***.

A.2.2 Complex phase space for a single oscillator

A single oscillator

The Hamiltonian for a single **linear** oscillator is given by

$$H = \frac{\omega}{2}(q^2 + p^2)\,. \tag{A.2.14}$$

The transformation to ***oscillator variables*** $(q,p) \to (a, a^*)$ for is defined by

$$a := q + ip \quad \text{and} \quad a^* := q - ip\,, \tag{A.2.15}$$

or

$$\begin{bmatrix} a^* \\ a \end{bmatrix} = \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix} \begin{bmatrix} q \\ p \end{bmatrix}, \tag{A.2.16}$$

for a single oscillator degree of freedom. This transformation is canonical,

$$\frac{-1}{2}\,\mathrm{Im}\,(da \wedge da^*) = \frac{-1}{2i} da \wedge da^* = dq \wedge dp\,. \tag{A.2.17}$$

As a check, the difference of canonical one-forms in these two sets of variables is an exact differential,

$$\frac{1}{2}\,\mathrm{Im}\,(a\,da^*) - pdq = \frac{1}{2}\,d(pq)\,. \tag{A.2.18}$$

The Hamiltonian (A.2.14) for the linear oscillator is given in these variables by

$$H = \frac{\omega}{2}(q^2+p^2) = \frac{\omega}{2}|a|^2, \tag{A.2.19}$$

with $a = q + ip$ and $a^* = q - ip$. For this particular Hamiltonian, the oscillator variables a and a^* evolve as

$$\begin{aligned} \dot{a} &= -i\omega a\,, \qquad & a(t) &= \exp(-i\omega t)\,, \\ \dot{a}^* &= i\omega a^*, \qquad & a^*(t) &= \exp(i\omega t)\,. \end{aligned} \tag{A.2.20}$$

Thus, for a linear oscillator, the variables a and a^* evolve on the unit circle in the complex plane with the same constant phase speed (linear phase shift) but in opposite directions.

For an arbitrary Hamiltonian, the evolution of the complex oscillator variable $a \in \mathbb{C}^1$ is governed by

$$\begin{aligned} \dot{a} = \dot{q} + i\dot{p} &= \frac{\partial H}{\partial p} - i\frac{\partial H}{\partial q} \\ &= \Big(\frac{\partial H}{\partial a}\frac{\partial a}{\partial p} + \frac{\partial H}{\partial a^*}\frac{\partial a^*}{\partial p}\Big) - i\Big(\frac{\partial H}{\partial a}\frac{\partial a}{\partial q} + \frac{\partial H}{\partial a^*}\frac{\partial a^*}{\partial q}\Big) \\ &= -2i\frac{\partial H}{\partial a^*} = \{a, H\}\,. \end{aligned} \tag{A.2.21}$$

An identical calculation gives the complex conjugate equation for the evolution of a^*

$$\dot{a}^* = 2i\frac{\partial H}{\partial a} = \{a^*, H\}\,. \tag{A.2.22}$$

Note that the Poisson bracket of $\{a, a^*\}$ satisfies

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

Now consider the Hamiltonian vector field

$$\begin{aligned} X_H &= \dot{a}\frac{\partial}{\partial a} + \dot{a}^*\frac{\partial}{\partial a^*} \\ &= -2i\Big(\frac{\partial H}{\partial a^*}\frac{\partial}{\partial a} - \frac{\partial H}{\partial a}\frac{\partial}{\partial a^*}\Big) = \{\cdot, H\}\,. \end{aligned} \tag{A.2.23}$$

Contraction of the Hamiltonian vector field with the symplectic two-form in oscillator variables yields

$$X_H \rfloor \frac{-1}{2i} da \wedge da^* = \frac{-1}{2i}(-2i)\Big(\frac{\partial H}{\partial a^*} da^* + \frac{\partial H}{\partial a} da\Big) = dH\,. \quad \text{(A.2.24)}$$

Remark A.2.5 (Summary for one degree of freedom)
The natural configuration space of variables for a single degree of freedom is the real line $\mathbb{R}$ with coordinate q. Making the identification $T^*\mathbb{R} \simeq \mathbb{R}^2 \simeq \mathbb{C}$ yields a canonical transformation which is particularly apt for oscillators.

A.3 Planar Isotropic Simple Harmonic Oscillator (PISHO)

A.3.1 Formulations of PISHO equations

Newton's Law ($m\mathbf{a} = \mathbf{F}$) for PISHO with displacement $\mathbf{X} \in \mathbb{R}^2$,

$$\ddot{\mathbf{X}} = -\mathbf{X} \quad \text{on} \quad TT\mathbb{R}^2 \quad \text{(A.3.1)}$$

is equivalent to the ***Euler-Lagrange equation,***

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{X}}} = \frac{\partial L}{\partial \mathbf{X}}\,, \quad \text{(A.3.2)}$$

which follows from stationarity $\delta S = 0$ of the ***action function***

$$S = \int_a^b L(\mathbf{X}, \dot{\mathbf{X}})\, dt\,,$$

for the ***Lagrangian*** $L : T\mathbb{R}^2 \to \mathbb{R}$ given by the *difference* between the kinetic and potential energies for PISHO

$$L(\mathbf{X}, \dot{\mathbf{X}}) = \frac{1}{2}\Big(|\dot{\mathbf{X}}|^2 - |\mathbf{X}|^2\Big) = KE - PE\,. \quad \text{(A.3.3)}$$

Note:

1. Newton's Law for PISHO (A.3.1) conserves the *sum* of energies, $KE + PE$. (Prove this statement!)
2. The Lagrangian for PISHO (A.3.3) only depends on the *magnitudes* of the vectors $\mathbf{X}$ and $\dot{\mathbf{X}}$.

By defining the ***momentum*** $\mathbf{P}$ as the ***fibre derivative*** of the Lagrangian,

$$\mathbf{P} := \frac{\partial L}{\partial \dot{\mathbf{X}}}(\mathbf{X}, \dot{\mathbf{X}}) \in T^*\mathbb{R}^2 \,, \tag{A.3.4}$$

one transforms $(\mathbf{X}, \dot{\mathbf{X}}) \in T\mathbb{R}^2$ to new variables $(\mathbf{Q}, \mathbf{P}) \in T^*\mathbb{R}^2$ with $\mathbf{Q} := \mathbf{X}$ and $\mathbf{P} := \partial L/\partial \dot{\mathbf{X}}$. Then the ***Legendre transformation*** to the ***Hamiltonian*** $H(\mathbf{Q}, \mathbf{P})$

$$H(\mathbf{Q}, \mathbf{P}) := \mathbf{P} \cdot \dot{\mathbf{Q}} - L(\mathbf{Q}, \dot{\mathbf{Q}})$$

allows the Euler-Lagrange equation (A.3.2) to be rewritten as ***Hamilton's canonical equations***

$$\dot{\mathbf{Q}} = \frac{\partial H}{\partial \mathbf{P}} \quad \text{and} \quad \dot{\mathbf{P}} = -\frac{\partial H}{\partial \mathbf{Q}} \,. \tag{A.3.5}$$

In general, Hamilton's canonical equations are equivalent to the corresponding Euler-Lagrange equation, provided the fibre derivative in (A.3.4) can be solved for the velocity $\dot{\mathbf{X}}$ as a function of $(\mathbf{Q}, \mathbf{P})$. This is always possible if the fibre derivative is a diffeomorphism (smooth invertible map, whose inverse is also smooth), in which case the Lagrangian is said to be *hyper-regular*. For PISHO, the Legendre transformation yields momentum $\mathbf{P} = \dot{\mathbf{Q}}$ and the Hamiltonian is,

$$H = \frac{1}{2}\Big(|\mathbf{P}|^2 + |\mathbf{Q}|^2\Big), \tag{A.3.6}$$

which is the conserved total energy, written as a function of the Legendre-transformed variables $(\mathbf{Q}, \mathbf{P})$.

Remark A.3.1

1. *Stationarity $\delta S = 0$ of the action function*

$$S = \int_a^b L(\mathbf{X}, \dot{\mathbf{X}})\, dt\,, \tag{A.3.7}$$

implies the Euler-Lagrange equation in (A.3.2),

$$\begin{aligned} 0 = \delta S &= \delta \int_a^b L(\mathbf{X}(t), \dot{\mathbf{X}}(t))\, dt \\ &=: \frac{d}{ds}\Big|_{s=0} \int_a^b L(\mathbf{X}(s,t), \dot{\mathbf{X}}(s,t))\, dt \\ &= \int_a^b \Big(- \frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{X}}} + \frac{\partial L}{\partial \mathbf{X}} \Big) \cdot \delta\mathbf{X}\, dt + \Big[\frac{\partial L}{\partial \dot{\mathbf{X}}} \cdot \delta \mathbf{X} \Big]_a^b \end{aligned}$$

after integrating by parts in time t for any smoothly parameterized variation

$$\delta \mathbf{X}(t) = \frac{d}{ds}\Big|_{s=0} \mathbf{X}(s,t)$$

that vanishes at the endpoints in time. This derivation of the Euler-Lagrange equation (A.3.2) by stationarity $\delta S = 0$ under variations of the action function (A.3.7) is called ***Hamilton's principle****.*

2. *If δS vanishes when the Euler-Lagrange equation is satisfied for a variation $\delta\mathbf{X}(t)$ that is a symmetry of the Lagrangian and does* not *vanish at the endpoints, then the endpoint term arising in the integration by parts must be a constant of the Euler-Lagrange motion.*

   This is ***Noether's theorem****: Every continuous symmetry of the action function S implies a conservation law under the motion obtained from the Euler-Lagrange equation.*

   For example, the PISHO Lagrangian in (A.3.3) is invariant under rotations around the vertical axis, for which $\delta\mathbf{X}(t) = \hat{\mathbf{Z}} \times \mathbf{X}$. In this case, the endpoint term is

$$L_Z = \dot{\mathbf{X}} \cdot \hat{\mathbf{Z}} \times \mathbf{X} = \hat{\mathbf{Z}} \cdot \mathbf{X} \times \dot{\mathbf{X}}\,,$$

   which is recognised as the vertical component of angular momentum.

3. *Likewise, substitution of the Legendre transformation into Hamilton's principle yields Hamilton's canonical equations (A.3.5)*

$$
\begin{aligned}
0 = \delta S &= \delta \int_a^b \Big(\mathbf{P} \cdot \dot{\mathbf{Q}} - H(\mathbf{Q}, \mathbf{P}) \Big)\, dt \\
&= \int_a^b \bigg(\Big(\dot{\mathbf{Q}} - \frac{\partial H}{\partial \mathbf{P}} \Big) \cdot \delta \mathbf{P} - \Big(\dot{\mathbf{P}} + \frac{\partial H}{\partial \mathbf{P}} \Big) \cdot \delta \mathbf{Q} \bigg)\, dt \\
&\quad + \Big[\mathbf{P} \cdot \delta \mathbf{Q} \Big]_a^b
\end{aligned}
$$

for any smoothly parameterized variation $\delta\mathbf{Q}(t)$ that vanishes at the endpoints in time.

A.3.2 The solution of PISHO in polar coordinates

Since the Lagrangian (A.3.3) for PISHO depends only on the magnitudes of the displacement and velocity in the plane, it is invariant under S^1 rotations about the origin of coordinates in $\mathbb{R}^2$. Consequently, one may transform the motion in Euclidean coordinates $\mathbf{X} = (x, y)$ to polar coordinates (r, θ) defined by $x + iy = re^{i\theta}$, to find the Lagrangian

$$
L(r, \dot{r}, \dot{\theta}) = \frac{1}{2}\big(\dot{r}^2 + r^2\dot{\theta}^2 - r^2\big)\,.
$$

Since the transformed Lagrangian has no explicit θ dependence; it is invariant under $\theta \to \theta + \epsilon$ for any constant $\epsilon \in S^1$. Just as in Cartesian coordinates, in polar coordinates Noether's Theorem for this S^1-invariance of the Lagrangian implies a conservation law. In this case, the symmetry under translations in θ implies conservation of the fibre derivative,

$$
\frac{\partial L}{\partial \dot{\theta}} = r^2\dot{\theta} =: p_\theta,
$$

which we may again identify as the ***angular momentum*** about the axis of rotation. For a given p_θ, the evolution of the angle $\theta(t)$ may be obtained from the motion for $r(t)$ by integrating the ***reconstruction equation***,

$$
\dot{\theta}(t) = \frac{p_\theta}{r^2(t)}\,. \tag{A.3.8}
$$

The Hamiltonian in these polar coordinates is given by

$$H(r, p_r, p_\theta) = \frac{1}{2}\Big(p_r^2 + \frac{p_\theta^2}{r^2} + r^2\Big)\,,$$

and Hamilton's canonical equations may be written as

$$\begin{aligned}
\dot{r} &= \frac{\partial H}{\partial p_r} = p_r\,, \quad \dot{p}_r = -\frac{\partial H}{\partial r} = -\,r + \frac{p_\theta^2}{r^3}\,,\\
\dot{\theta} &= \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{r^2}\,, \quad \dot{p}_\theta = -\frac{\partial H}{\partial \theta} = 0\,.
\end{aligned}$$

However, after writing Hamilton's canonical equations, instead of setting about solving them for their solutions (which are well known in this familiar case), we may notice something about their geometrical meaning. Namely, the level sets of the Hamiltonian $H = const$ and angular momentum $p_\theta = const$ may be regarded as two surfaces in $\mathbb{R}^3$ with coordinates $\boldsymbol{\chi} = (r, p_r, p_\theta)$. In these $\mathbb{R}^3$ coordinates, Hamilton's equations may be written as

$$\dot{\boldsymbol{\chi}}(t) = \nabla H \times \nabla p_\theta\,, \quad \text{where} \quad \nabla := \frac{\partial}{\partial \boldsymbol{\chi}}\,. \tag{A.3.9}$$

This means that the evolution in $\mathbb{R}^3$ with coordinates $\boldsymbol{\chi} = (r, p_r, p_\theta)$ takes place along the intersections of the level sets of the constants of motion p_θ and H. This realisation is an opportunity to start thinking geometrically about mechanics.

Remark A.3.2

1. *The process of obtaining the geometrical picture of the solution set did not depend on the particular choice of the Hamiltonian, only its on its θ-independence.*

2. *Instead of transforming to polar coordinates, one could have chosen to transform directly to S^1-invariant variables, such as $|\mathbf{P}|^2$, $|\mathbf{Q}|^2$, $\mathbf{P}\cdot\mathbf{Q}$ and $p_\theta^2 = |\mathbf{P}\times\mathbf{Q}|^2$. We will consider this option later.*

A.3.3 Geometric and dynamic phases for PISHO

The phase θ may be reconstructed from the solution for $r^2(t)$ by integrating the reconstruction equation (A.3.8), once an expression for the radial motion $r(t)$ has been found. However, more geometric insight into the solution may be found by considering the relation

$$(p_x + ip_y)(dx - idy) = (p_r + ip_\theta/r)(dr - ird\theta)$$

whose real part yields the ***action 1-form***,

$$\mathbf{p} \cdot d\mathbf{q} := p_x dx + p_y dy = p_\theta d\theta + p_r dr \,. \tag{A.3.10}$$

Hence, integration around a periodic solution in the (r, p_r) plane on a level surface of p_θ yields the phase relation,

$$\oint p_\theta \, d\theta = p_\theta \Delta\theta = \underbrace{-\oint p_r \, dr}_{\text{Geometric}} + \underbrace{\oint \mathbf{p} \cdot d\mathbf{q}}_{\text{Dynamic}} \,. \tag{A.3.11}$$

In this formula, the total phase change around a closed periodic orbit of period T in (r, p_r) plane decomposes into the sum of two parts. On writing this decomposition of the phase as

$$\Delta\theta = \Delta\theta_{geom} + \Delta\theta_{dyn} \,, \tag{A.3.12}$$

one sees that the geometric part, given by Stokes theorem as

$$\begin{aligned} p_\theta \Delta\theta_{geom} &= -\oint p_r \, dr \\ &= -\iint dp_r \wedge dr = \text{Orbital area} \,, \end{aligned} \tag{A.3.13}$$

is the area enclosed by the periodic orbit in the radial phase plane, written in terms of the antisymmetric wedge product of differential forms, $\wedge$. Thus, the name: ***geometric phase*** for $\Delta\theta_{geom}$, because this part of the phase only depends on the geometric area of the periodic orbit. The rest of the phase is given by

$$
\begin{aligned}
p_\theta \Delta\theta_{dyn} &= \oint \mathbf{p}\cdot d\mathbf{q} = \int_0^T \mathbf{p}\cdot\dot{\mathbf{q}}\,dt \\
&= \int_0^T \left(p_r \frac{\partial H}{\partial p_r} + p_\theta \frac{\partial H}{\partial p_\theta} \right) dt \\
&= \int_0^T \left(p_r^2 + \frac{p_\theta^2}{r^2} \right) dt \\
&= \int_0^T \left(2H - r^2 \right) dt \\
&= T\Big(2H - \big\langle r^2 \big\rangle \Big)\,,
\end{aligned}
\qquad (A.3.14)
$$

where the integral $\int_0^T r^2 dt = T\langle r^2\rangle$ defines the time average $\langle r^2\rangle$ over the orbit of period T of the squared orbital radius. This part of the phase depends on the Hamiltonian, orbital period and the time average over the orbit of the squared radius. Thus, the name: ***dynamic phase*** for $\Delta\theta_{dyn}$, because this part of the phase depends on the Hamiltonian responsible for the dynamics of the orbit, not just its area.

Exercise. Write the Hamiltonian forms of the PISHO equations in terms of S^1-invariant quantities, for the following two cases:

1. $x = |\mathbf{Q}|^2$, $y = |\mathbf{P}|^2$, $z = \mathbf{P}\cdot\mathbf{Q}$ and $p_\theta^2 = |\mathbf{P}\times\mathbf{Q}|^2$
2. $R = |a_1|^2 + |a_2|^2$, $X_1 + iX_2 = 2a_1^* a_2$ and $X_3 = |a_1|^2 - |a_2|^2$, with $a_k := q_k + ip_k$.

★

Exercise. Write the Poisson brackets among the variables in the two cases in the previous exercise. ★

A.4 Complex phase space for two oscillators

Two uncoupled oscillators

The configuration space of variables for two oscillators is $\mathbb{R}\times\mathbb{R} \simeq \mathbb{R}^2$ and the cotangent bundle is $T^*\mathbb{R}^2 \simeq \mathbb{C}^2$. In the transformation to ***oscillator variables***, $a_j = q_j + ip_j$, $j = 1, 2$, the symplectic two-form is computed from

$$\begin{aligned} \sum_{j=1}^{2} dq_j \wedge dp_j &= \frac{-1}{2i} \sum_{j=1}^{2} d(q_j + ip_j) \wedge d(q_j - ip_j) \\ &= \frac{1}{-2i} \sum_{j=1}^{2} da_j \wedge da_j^* \,. \end{aligned} \tag{A.4.1}$$

So the transformation to oscillator variables is canonical. The corresponding Poisson bracket is,

$$\{a_j, a_k^*\} = \{q_j + ip_j \,,\, q_k - ip_k\} = -2i\,\{q_j \,,\, p_k\} = -2i\,\delta_{jk} \,.$$

Consequently, the canonical motion equations in oscillator variables become

$$\begin{aligned} &\dot{a}_j = \{a_j, H\} = -\,2i\,\frac{\partial H}{\partial a_j^*} \\ &\text{and} \\ &\dot{a}_j^* = \{a_j^*, H\} = 2i\,\frac{\partial H}{\partial a_j} \,. \end{aligned} \tag{A.4.2}$$

The Hamiltonian for two uncoupled linear oscillators with angular frequencies ω_1 and ω_2 is expressed in the variables $(a_1, a_2) \in \mathbb{C}^2$ as

$$H = \sum_{j=1}^{2} \frac{\omega_j}{2}(q_j^2 + p_j^2) = \sum_{j=1}^{2} \frac{\omega_j}{2}|a_j|^2, \quad \text{with} \quad a_j = q_j + ip_j \,. \tag{A.4.3}$$

This Hamiltonian is invariant under two independent phase shifts, or ***circle actions*** $SO(2) \times SO(2) \simeq S^1 \times S^1$, which are also the solutions of the equations of motion in (A.4.2).

$$a_j(t) = a_j(0)e^{-i\omega_j t} \quad \text{and} \quad a_j^*(t) = a_j^*(0)e^{i\omega_j t} \,, \tag{A.4.4}$$

for $j = 1, 2$, with no sum on j implied.

A.5 2D resonant oscillators

Many problems in physics involve oscillator Hamiltonians that are invariant under the $n:m$ circle action in which $\omega_1 : \omega_2 = n : m$, for integers $n, m \in \mathbb{Z}$. In that case, the two oscillators are said to be in $n:m$ resonance. The corresponding circle action is

$$\begin{aligned} \phi_{n:m} : \mathbb{C}^2 \to \mathbb{C}^2 \quad \text{as} \quad & (a_1, a_2) \to (e^{in\phi}a_1, e^{im\phi}a_2) \\ \text{and} \quad & (a_1^*, a_2^*) \to (e^{-in\phi}a_1^*, e^{-im\phi}a_2^*)\,. \end{aligned} \tag{A.5.1}$$

The simplest of these modifications is to break directional isotropy,

Consider the 2D oscillator Hamiltonian $H : \mathbb{C}^2 \to \mathbb{R}$,

$$\begin{aligned} H &= \frac{1}{2}\sum_{j=1}^{2} \omega_j |a_j|^2 \\ &= \frac{1}{4}(\omega_1+\omega_2)\big(|a_1|^2+|a_2|^2\big) + \frac{1}{4}(\omega_1-\omega_2)\big(|a_1|^2-|a_2|^2\big)\,. \end{aligned} \tag{A.5.2}$$

Definition A.5.1 *When $\omega_1 = \omega_2$ in (A.5.2) the 2D oscillator is **isotropic**, otherwise it is **anisotropic**.*

Remark A.5.2

- *H in (A.5.2) is invariant under $a_j \to a_j' = exp(i\phi_j)a_j$ and the solution of*

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

  is $a_j(t) = \exp(-i\omega_j t)a_j(0)$. This is motion at constant speed on $S^1 \times S^1 \simeq T^2$ (the 2-torus).

- *2D positive resonance occurs when $\omega_1 : \omega_2 = m : n$ with $m, n \in \mathbb{Z}^+$.*

- *The relation $\omega_1 : \omega_2 = n : m$ with $n : m \in \mathbb{Z}$ is called a **resonance** because the phase shifts match for n oscillations of a_1 and m oscillations of a_2.*

- *H in (A.5.2) is the sum of a $1:1$ resonant oscillator and a $1:-1$ oscillator.*

A.5.1 $1:1$ resonance

Exercise. What quadratic monomials in $(a_1,\, a_2) \in \mathbb{C}^2$ are invariant under $\omega_1 : \omega_2 = 1 : 1$ diagonal S^1 phase changes? ★

Answer. The following quadratic monomials in $(a_1,\, a_2) \in \mathbb{C}^2$ are invariant under $1 : 1$ phase shifts: $\{a_1 a_1^*, a_2 a_2^*, a_1 a_2^*, a_1^* a_2\}$. Choose

$$\begin{aligned} |a_1|^2 + |a_2|^2 &= R\,, \\ |a_1|^2 - |a_2|^2 &= Z\,, \\ 2a_1 a_2^* &= X - iY\,. \end{aligned} \tag{A.5.3}$$

▲

Exercise. Find the linear transformations generated by X, Y, Z, R on a_1, a_2. Express them as matrix operations. ★

Answer. The infinitesimal transformations are given by

$$\begin{aligned} X_R a_j &= \{a_j, R\} = -\,2i\,\frac{\partial R}{\partial a_j^*} = -\,2i\,a_j\,, \\ X_Z a_1 &= \{a_1, Z\} = -\,2i\,a_1\,, \\ X_Z a_2 &= \{a_2, Z\} = +\,2i\,a_2\,, \\ X_X a_1 &= \{a_1, X\} = -\,2i\,a_2\,, \\ X_X a_2 &= \{a_2, X\} = -\,2i\,a_1\,, \\ X_Y a_1 &= \{a_1, Y\} = -\,2a_2\,, \\ X_Y a_2 &= \{a_2, Y\} = +\,2a_1\,. \end{aligned}$$

These infinitesimal transformations may be expressed as the matrix operations,

$$
\begin{aligned}
X_Z \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} &= -2i \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \\
X_X \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} &= -2i \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \\
X_Y \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} &= -2i \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix},
\end{aligned}
$$

or, in vector notation, as

$$X_Z\mathbf{a} = -\,2i\,\sigma_3\mathbf{a}\,, \quad X_X\mathbf{a} = -\,2i\,\sigma_1\mathbf{a}\,, \quad X_Y\mathbf{a} = -\,2i\,\sigma_2\mathbf{a}\,.$$

From these expressions, one recognises that the finite transformations, or flows, of the Hamiltonian vector fields for (X, Y, Z) are rotations about the (X, Y, Z) axes, respectively. ▲

> **Exercise.** Define the Poincaré sphere.
>
> How are the S^1 invariant quadratic monomials (X, Y, Z) for $\omega_1 : \omega_2 = 1 : 1$ related to the Poincaré sphere? ★

Answer.

Definition A.5.3 *The Poincaré sphere is the **orbit manifold** for invariant* $1 : 1$ *resonance dynamics. Being invariant under the flow of the Hamiltonian vector field* $X_R = \{\,\cdot\,,\ R\}$*, each point on the Poincaré sphere consists of a resonant orbit under the* $1 : 1$ *circle action*

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

The S^1 invariant quadratic monomials (X, Y, Z) for $\omega_1 : \omega_2 = 1 : 1$ satisfy

$$
\begin{aligned}
R^2 &= Z^2 + X^2 + Y^2 \\
&= (|a_1|^2 - |a_2|^2)^2 + 4|a_1|^2|a_2|^2 \\
&= (|a_1|^2 + |a_2|^2)^2.
\end{aligned}
\tag{A.5.4}
$$

The zero level set $R^2 - (X^2 + Y^2 + Z^2) = 0$ is the Poincaré sphere.
▲

Exercise. Given $\{a_j, a_k^*\} = -2i\delta_{jk}$ compute the Poisson brackets among the S^1 invariants X, Y, Z, R for the $1:1$ resonance.

★

Answer. $\{R, Q\} = 0$, for $Q \in \{X, Y, Z\}$, and $\{X, Y\} = Z$, plus cyclic permutations of (X, Y, Z). ▲

Exercise. For the Hamiltonian,

$$
\begin{aligned}
H &= \frac{\omega_1}{2}\,(R + Z) + \frac{\omega_2}{2}\,(R - Z) \\
&= \frac{1}{2}(\omega_1 + \omega_2)R + \frac{1}{2}(\omega_1 - \omega_2)Z\,,
\end{aligned}
\tag{A.5.5}
$$

write the equations $\dot{X}, \dot{Y}, \dot{Z}, \dot{R}$ for the S^1 invariants X, Y, Z, R of the $1:1$ resonance. Also write these equations in vector form, with $\mathbf{X} = (X, Y, Z)^T$. Describe this motion in terms of level sets of the Poincaré sphere and the Hamiltonian H. ★

Answer. The dynamics is given by the Poisson bracket relation,

$$
\begin{aligned}
\dot{F} = \{F, H\} &= -\nabla\frac{R^2}{2}\cdot\nabla F\times\nabla H(X, Y, Z) \\
&= -\frac{1}{2}(\omega_1 - \omega_2)\,\nabla\frac{R^2}{2}\cdot\nabla F\times\nabla Z\,.
\end{aligned}
$$

Then $\dot{R} = 0 = \dot{Z}$ and

$$
\dot{X} = \frac{1}{2}\,(\omega_1 - \omega_2)Y\,, \qquad \dot{Y} = -\,\frac{1}{2}\,(\omega_1 - \omega_2)X\,.
$$

In vector form, with $\mathbf{X} = (X, Y, Z)^T$, this is

$$
\dot{\mathbf{X}} = \frac{1}{2}\,(\omega_1 - \omega_2)\mathbf{X}\times\widehat{\mathbf{Z}}\,,
$$

where $\widehat{\mathbf{Z}}$ is the unit vector in the Z-direction ($\cos\theta = 0$). This motion is uniform rotation in the positive direction along a latitude of the Poincaré sphere $R = const$. This azimuthal rotation on a latitude at fixed polar angle on the sphere occurs along the intersections of level sets of the Poincaré sphere $R = const$ and the planes $Z = const$, which are level sets of the Hamiltonian for a fixed value of R.

▲

Exercise. How is angular momentum $L := \epsilon_{3jk} p_j q_k = p_1 q_2 - p_2 q_1$ related to the $1:1$ resonance invariants (X, Y, Z)? What transformation does L apply to these variables? Check your answer explicitly. ★

Answer. The transformation of the $1:1$ resonant variables is found by noticing that

$$L = \epsilon_{3jk} p_j q_k = p_1 q_2 - p_2 q_1 = \mathrm{Im}(a_1 a_2^*) = -\frac{Y}{2}\,,$$

and computing the Poisson bracket,

$$\frac{dF}{d\phi} = -\frac{1}{2}\Big\{F, Y\Big\} = \frac{1}{2}\nabla\frac{R^2}{2}\cdot\nabla F \times \nabla Y = -\frac{1}{2}\nabla F \cdot \mathbf{X} \times \widehat{\mathbf{Y}}\,.$$

This is a counterclockwise rotation of the (X, Y, Z) coordinate frame about the Y-axis. The flow of the Hamiltonian vector field $X_L = \{\,\cdot\,,\, L\}$ is given by

$$\phi(\exp tX_L) = \begin{pmatrix} \cos t & 0 & \sin t \\ 0 & 1 & 0 \\ -\sin t & 0 & \cos t \end{pmatrix}.$$

To check this result explicitly, compute the skew-symmetric matrix,

$$\Big(\frac{d}{dt}\phi(\exp tX_L)\phi(\exp -tX_L)\Big)\Big|_{t=0} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix},$$

and apply it to confirm that

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} Z \\ 0 \\ -X \end{pmatrix} = -\mathbf{X} \times \widehat{\mathbf{Y}} \,.$$

▲

A.5.2 $1 : -1$ resonance

> **Exercise.** What quadratic monomials in a_1 and a_2 that are invariant under resonant S^1 phase changes, $\omega_1 : \omega_2 = 1 : -1$. What is the orbital manifold? How are the solutions related to ray optics? ★

Answer. The quadratic monomials of a_1 and a_2 that are left invariant under $1 : -1$ resonant S^1 phase shifts are:

$$\Big\{ |a_1|^2, |a_2|^2, a_1 a_2, a_1^* a_2^* \Big\} \,.$$

Consequently, the following are also invariant:

$$\begin{aligned} S &= |a_1|^2 - |a_2|^2, \\ Y_1 &= |a_1|^2 + |a_2|^2, \\ Y_2 + iY_3 &= 2a_1 a_2 \,. \end{aligned} \tag{A.5.6}$$

These satisfy

$$Y_2^2 + Y_3^2 = 4|a_1|^2|a_2|^2 = Y_1^2 - S^2 \,,$$

so the level sets of the orbital manifold are the hyperboloids of revolution around the Y_1-axis parameterised by S. That is,

$$S^2 = Y_1^2 - Y_2^2 - Y_3^2 \,. \tag{A.5.7}$$

We remark that:
$S = |a_1|^2 - |a_2|^2 = const$ is an ***hyperboloid*** in both $\mathbb{C}^2$ and $\mathbb{R}^3$.
$Y_1 = |a_1|^2 + |a_2|^2 = const$ is a ***sphere*** $S^3 \in \mathbb{C}^2$, and it is a ***plane*** in $\mathbb{R}^3$.

For the $1 : -1$ resonance Hamiltonian $H = Y_1$, the evolution of S, Y_2, Y_3 is described by

$$\begin{aligned} \dot{\mathbf{Y}} &= \nabla S^2 \times \widehat{\mathbf{Y}}_\mathbf{1} = 2\widehat{\mathbf{Y}}_\mathbf{1} \times \mathbf{Y} \\ \dot{S} &= \{S, H\} = 0, \\ \dot{Y}_1 &= \{Y_1, H\} = \{Y_1, Y_1\} = 0, \\ \dot{Y}_2 &= \{Y_2, H\} = \{Y_2, Y_1\} = -2Y_3, \\ \dot{Y}_3 &= \{Y_3, H\} = \{Y_3, Y_1\} = 2Y_2\,. \end{aligned} \tag{A.5.8}$$

Thus, S is invariant under the flow generated by the Hamiltonian $H = Y_1$, while Y_2 and Y_3 rotate clockwise around the Y_1-axis in a plane at $Y_1 = const$. This is the same motion as in equation (1.12.8) for the paraxial harmonic guide. Looking more closely, one sees that the Lie-Poisson bracket for the paraxial rays is identical to that for the $1 : -1$ resonance. This is a coincidence that occurs because the Lie algebras $sp(2, \mathbb{R})$ and $su(1, 1)$ happen to be identical.

▲

Exercise. How do a_1 and a_2 evolve under the Hamiltonian $H = Y_1$?

Is this evolution consistent with the $\mathbb{R}^3$-evolution?

If not, why not? ★

Answer. The motion under the Hamiltonian $H = Y_1$ is given by

$$a_1(t) = \exp(-2it)a_1(0),\ a_2(t) = \exp(-2it)a_2(0)\,.$$

This is demonstrated by computing

$$\{Y_2 + iY_3, Y_1\} = -2Y_3 + 2iY_2 = 2i(Y_2 + iY_3)\,. \tag{A.5.9}$$

Hence, we have the evolution,

$$Y_2(t) + iY_3(t) = \exp(2it)\Big(Y_2(0) + iY_3(0)\Big)\,. \tag{A.5.10}$$

This is a clockwise rotation. However, the previous solution for the oscillator variables would have given a counterclockwise rotation and at rate 4. This is no real surprise, because the map from oscillator variables to $\mathbb{R}^3$-variables is defined modulo an arbitrary constant factor, corresponding to a multiple of the units of time. It can be rectified by introducing a constant, for example, by taking either $t \to -2t$, or $S^2 \to -2S^2$. ▲

Exercise. Find the transformations generated by Y_2 and Y_3 on a_1, a_2. ★

Answer. The infinitesimal transformations are given by

$$\begin{aligned}\{a_1, Y_2\} &= -2i\, a_2^*\,, \qquad \{a_2, Y_2\} = -2i\, a_1^*\,,\\ \{a_1, Y_3\} &= 2\, a_2^*\,, \qquad \{a_2, Y_3\} = 2\, a_1^*\,.\end{aligned}$$

▲

Exercise. For the $1:-1$ resonance, consider the following hyperbolic analogue of the Riemann stereographic projection obtained from the mapping

$$\zeta = \frac{Y_2 + iY_3}{Y_1 + 1}\,. \tag{A.5.11}$$

Using hyperbolic polar coordinates as in (1.12.2) with $S = 1$,

$$Y_1 = \cosh u\,, \quad Y_2 = \sinh u \cos\psi\,, \quad Y_3 = \sinh u \sin\psi\,, \tag{A.5.12}$$

find that $\zeta = e^{i\psi}\coth u$. Show that

$$Y_1 = \frac{1 + |\zeta|^2}{1 - |\zeta|^2} \quad \text{and} \quad Y_2 + iY_3 = \frac{2\zeta}{1 - |\zeta|^2}\,. \tag{A.5.13}$$

Compute the Poisson brackets for $\{\zeta, \zeta^*\}$ on the complex plane. ★

A.5.3 Hamiltonian flow for $m_1 : m_2$ resonance

A convenient canonical transformation for the $m_1 : m_2$ resonance is given by

$$a_j = q_j + ip_j \quad \text{for} \quad j = 1, 2,$$

for which

$$\{q_j, p_k\} = \delta_{jk} \quad \text{implies} \quad \{a_j, a_k^*\} = -2i\,\delta_{jk}\,.$$

The $m_1 : m_2$ resonance is the flow of the Hamiltonian vector field

$$X_R = \{\,\cdot\,, R\} = -\frac{2i}{m_j} a_j \frac{\partial}{\partial a_j} + \frac{2i}{m_j} a_j^* \frac{\partial}{\partial a_j^*}\,,$$

which is generated by the Poisson bracket with

$$R = \frac{1}{m_1}\,|a_1|^2 + \frac{1}{m_2}\,|a_2|^2\ .$$

The characteristic equations for this Hamiltonian vector field are,

$$\dot{a}_j = \{a_j, R\} = -2i\frac{\partial R}{\partial a_j^*} = -\frac{2i}{m_j} a_j \qquad \text{(no sum)} \quad \text{for} \quad j = 1, 2,$$

whose flow solutions are

$$a_j\,(t) = e^{-2it/m_j} a_j\,(0) \quad \text{(no sum)} \quad \text{for} \quad j = 1, 2\,.$$

This flow leaves invariant the quantity,

$$a_1^{m_1}\,(t)\, a_2^{*\,m_2}\,(t) = a_1^{m_1}\,(0)\, a_2^{*\,m_2}\,(0)\ .$$

A.5.4 Multi-sheeted polar coordinates

The symplectic 2-form in canonically conjugate coordinates satisfying the Poisson bracket relation $\{q,\, p\} = 1$ is

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

Under the canonical transformation to oscillator variables, $a = q + ip \in \mathbb{C}$, this becomes

$$\begin{aligned} da \wedge da^* &= (dq + idp) \wedge (dq + idp) \\ &= -2i\,(dq \wedge dp) = -\,2i\,\omega\,, \end{aligned}$$

and the corresponding Poisson bracket is $\{a,\, a^*\} = -2i$.

One transforms to multi-sheeted polar coordinates in $b \in \mathbb{C}$ by introducing another complex variable, defined by

$$\begin{aligned} b &= \frac{1}{\sqrt{m}}\,|a|^{1-m}\,a^m \\ &= \frac{1}{\sqrt{m}}\,|a|\,e^{im\varphi}\,, \end{aligned}$$

with $\varphi \in [0, 2\pi)$. In ordinary polar coordinates, this is

$$\begin{aligned} a &= |a|\,e^{i\varphi}\,, \\ \varphi &= \tan^{-1}\frac{p}{q}\,, \end{aligned}$$

and

$$a = re^{i\varphi} = \sqrt{q^2+p^2}\,e^{i\varphi}\,.$$

The corresponding symplectic forms are given by

$$\begin{aligned} da \wedge da^* &= d\left(re^{i\varphi}\right) \wedge d\left(re^{-i\varphi}\right) \\ &= e^{i\varphi}\,(dr + ird\varphi) \wedge e^{-i\varphi}\,(dr - ird\varphi) \\ &= -2i\,rdr \wedge d\varphi\,, \end{aligned}$$

and

$$\begin{aligned} db \wedge db^* &= \frac{1}{m} d\left(re^{im\varphi}\right) \wedge d\left(re^{-im\varphi}\right) \\ &= \frac{1}{m} e^{im\varphi}\,(dr + imrd\varphi) \wedge e^{-im\varphi}\,(dr - imrd\varphi) \\ &= \frac{1}{m}\,(-2imrdr \wedge d\varphi) \\ &= -\,2i\,rdr \wedge d\varphi\,. \end{aligned}$$

The associated Poisson bracket is $\{b,\, b^*\} = -2i$.

A.5.5 Resonant $m_1 : m_2$ torus

Define the quotient map $\pi : \mathbb{C}^2 \to \mathbb{R}^3$ for the $m_1 : m_2$ resonant torus as

$$\mathbf{n} = b_k \boldsymbol{\sigma}_{kl} b_l^* = \pi\left(\mathbf{b}\right),$$

where $\boldsymbol{\sigma}$ is th 3-vector of 2×2 Pauli matrices. Explicitly, this is given by

$$\begin{aligned} n_0 &= |b_1|^2 + |b_2|^2 = \frac{1}{m_1}|a_1|^2 + \frac{1}{m_2}|a_2|^2 , \\ n_3 &= |b_1|^2 - |b_2|^2 = \frac{1}{m_1}|a_1|^2 - \frac{1}{m_2}|a_2|^2 , \end{aligned}$$

and

$$n_1 + in_2 = 2b_1b_2^* = \frac{2}{\sqrt{m_1m_2}}\frac{a_1^{m_1}}{|a_1|^{m_1-1}}\frac{a_2^{*m_2}}{|a_2|^{m_2-1}} .$$

One computes the corresponding Poisson brackets on $\mathbb{R}^3$ as,

$$\begin{aligned} \{n_j, n_k\} &= -\varepsilon_{jkl}n_l , \\ \{n_0, n_k\} &= 0 . \end{aligned}$$

The orbit manifold is defined by the algebraic relation,

$$\begin{aligned} n_1^2 + n_2^2 + n_3^2 &= |n_1 + in_2|^2 + n_3^2 \\ &= 4|b_1|^2|b_2|^2 + \left(|b_1|^2 - |b_2|^2\right)^2 \\ &= \left(|b_1|^2 + |b_2|^2\right)^2 = n_0^2 , \end{aligned}$$

whose locus is a sphere centred at the origin in $\mathbb{R}^3$.

For $\mathbb{R}^3 \in \operatorname{image}\pi$, the inverse of the quotient map

$$\pi^{-1} : \mathbb{R}^3 \to \mathbb{C}^2 ,$$

satisfies the formulas,

$$\begin{aligned} n_0 + n_3 &= \frac{2}{m_1}|a_1|^2 , \qquad |a_1|^2 = \frac{m_1}{2}\left(n_0 + n_3\right) , \\ n_0 - n_3 &= \frac{2}{m_2}|a_2|^2 , \qquad |a_2|^2 = \frac{m_2}{2}\left(n_0 - n_3\right) . \end{aligned}$$

Thus, each point on a sphere given by a level set of

$$n_0 \quad = \quad |b_1|^2 + |b_2|^2 = \frac{1}{m_1}\,|a_1|^2 + \frac{1}{m_2}\,|a_2|^2\,,$$

corresponds to an $m_1 : m_2$ resonant torus.

A.6 A quadratically nonlinear oscillator

Consider the Hamiltonian dynamics on a symplectic manifold of a system comprising two real degrees of freedom, with real phase space variables (x, y, θ, z), symplectic form

$$\omega = dx \wedge dy + d\theta \wedge dz$$

and Hamiltonian

$$H = \tfrac{1}{2}y^2 + x\left(\tfrac{1}{3}x^2 - z\right) - \tfrac{2}{3}z^{3/2} \tag{A.6.1}$$

whose potential energy

$$V(x, z) = x\left(\tfrac{1}{3}x^2 - z\right) \tag{A.6.2}$$

is a cubic in the coordinate x.

1. Write the canonical Poisson bracket for this system.

   Answer.

   $$\{F\,,\,H\} = H_yF_x - H_xF_y + H_zF_\theta - H_\theta F_z$$

   ▲

2. Write Hamilton's canonical equations for this system. Explain how to keep $z \geq 0$, so that H and θ remain real.

   Answer. Hamilton's canonical equations for this system are

   $$\begin{aligned}\dot{x} &= \{x, H\} = H_y = y\,,\\ \dot{y} &= \{y, H\} = -\,H_x = -(x^2 - z)\,,\end{aligned}$$

and

$$\dot{\theta} = \{\theta, H\} = H_z = -(x + \sqrt{z}\,)\,,$$
$$\dot{z} = \{z, H\} = -\,H_\theta = 0\,.$$

For H and θ to remain real, one need only choose the initial value of the constant of motion $z \geq 0$. ▲

3. At what values of x, y and H does the system have stationary points in the (x, y) plane?

   Answer. The system has (x, y) stationary points when its time derivatives vanish: at $y = 0$, $x = \pm\sqrt{z}$ and $H = -\frac{4}{3}z^{3/2}$. ▲

4. Propose a strategy for solving these equations. In what order should they be solved?

   Answer. Since z is a constant of motion, the equation for its conjugate variable $\theta(t)$ decouples from the others and may be solved as a quadrature *after* first solving for $x(t)$ and $y(t)$ on a level set of z. ▲

5. Identify the constants of motion of this system and explain why they are conserved.

   Answer. There are two constants of motion:
   (i) The Hamiltonian H for the canonical equations is conserved, because the Poisson bracket in $\dot{H} = \{H, H\}$ is antisymmetric.
   (ii) The momentum z conjugate to θ is conserved, because $H_\theta = 0$. ▲

6. Compute the associated Hamiltonian vector field X_H and show that it satisfies

$$X_H \,\lrcorner\, \omega = dH$$

   Answer.

$$\begin{aligned} X_H = \{\,\cdot\,,\,H\} &= H_y\partial_x - H_x\partial_y + H_z\partial_\theta - H_\theta\partial_z \\ &= y\partial_x - (x^2 - z)\partial_y - (x + \sqrt{z})\partial_\theta, \end{aligned}$$

so that

$$X_H \mathbin{\lrcorner} \omega = ydy + (x^2 - z)dx - (x + \sqrt{z}\,)dz = dH$$

▲

7. Write the Poisson bracket that expresses the Hamiltonian vector field X_H as a divergenceless vector field in $\mathbb{R}^3$ with coordinates $\mathbf{x} = (x, y, z) \in \mathbb{R}^3$. Explain why this Poisson bracket satisfies the Jacobi identity.

   Answer. Write the evolution equations for $\mathbf{x} = (x, y, z)^T \in \mathbb{R}^3$ as

$$\begin{aligned} \dot{\mathbf{x}} = \{\mathbf{x}, H\} = \nabla H \times \nabla z &= (H_y, -H_x, 0)^T \\ &= (y, z - x^2, 0)^T \\ &= (\dot{x}, \dot{y}, \dot{z})^T . \end{aligned}$$

   Hence, for any smooth function $F(\mathbf{x})$,

$$\{F, H\} = \nabla z \cdot \nabla F \times \nabla H = F_x H_y - H_x F_y \,.$$

   This is the canonical Poisson bracket for one degree of freedom, which is known to satisfy the Jacobi identity. ▲

8. Identify the Casimir function for this $\mathbb{R}^3$ bracket. Show explicitly that it satisfies the definition of a Casimir function.

   Answer. Substituting $F = \Phi(z)$ for a smooth function Φ into the bracket expression yields

$$\{\Phi(z), H\} = \nabla z \cdot \nabla \Phi(z) \times \nabla H = \nabla H \cdot \nabla z \times \nabla \Phi(z) = 0,$$

   for all H. This proves that $F = \Phi(z)$ is a Casimir function for any smooth Φ. ▲

9. Sketch a graph of the intersections of the level surfaces in $\mathbb{R}^3$ of the Hamiltonian and Casimir function. Determine the directions of flow along these intersections. Identify the locations and types of any relative equilibria at the tangent points of these surfaces.

Answer. The sketch should show a saddle-centre fish shape pointing rightward in the (x, y) plane with elliptic equilibrium at $(x, y) = (\sqrt{z}, 0)$, hyperbolic equilibrium at $(x, y) = (-\sqrt{z}, 0)$. The directions of flow have sign$(\dot{x})$=sign(y). The saddle-centre shape is sketched in the lower panel of Figure A.3. ▲

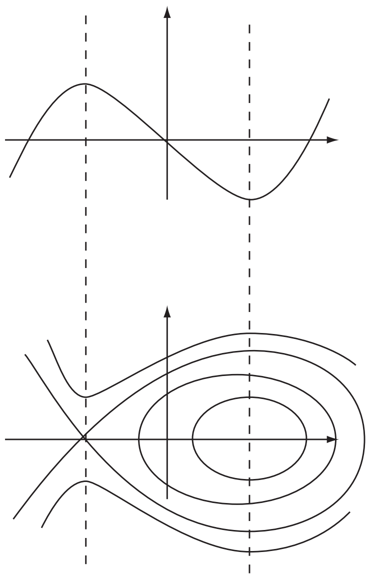


Figure A.3: Upper panel: sketch of the cubic potential $V(x, z)$ in equation (A.6.1) at constant z. Lower panel: sketch of its (x, y) phase plane, comprising several level sets of $H(x, y, z)$ at constant z. This is the saddle-centre fish shape.

10. Linearise around the relative equilibria on a level set of the Casimir (z) and compute its eigenvalues.

Answer. On a level surface of z the (x, y) coordinates satisfy $\dot{x} = y$ and $\dot{y} = z - x^2$. Linearising around $(x_e, y_e) = (\pm\sqrt{z}, 0)$ yields with $(x, y) = (x_e + \xi(t), y_e + \eta(t))$

$$\begin{bmatrix} \dot{\xi} \\ \dot{\eta} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2x_e & 0 \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix}.$$

Its characteristic equation,

$$\det \begin{bmatrix} \lambda & -1 \\ 2x_e & \lambda \end{bmatrix} = \lambda^2 + 2x_e = 0,$$

yields $\lambda^2 = -2x_e = \mp 2\sqrt{z}$.
Hence, the eigenvalues are,
$\lambda = \pm i\sqrt{2}z^{1/4}$ at the elliptic equilibrium $(x_e, y_e) = (\sqrt{z}, 0)$, and
$\lambda = \pm\sqrt{2}z^{1/4}$ at the hyperbolic equilibrium $(x_e, y_e) = (-\sqrt{z}, 0)$. ▲

11. As shown in Figure A.3, the hyperbolic equilibrium point is connected to itself by a homoclinic orbit. Reduce the equation for the homoclinic orbit to an indefinite integral expression.

Answer. On the homoclinic orbit the Hamiltonian vanishes, so that

$$H = \frac{1}{2}y^2 + x\left(\frac{1}{3}x^2 - z\right) - \frac{2}{3}z^{3/2} = 0.$$

Using $y = \dot{x}$, rearranging and integrating implies the indefinite integral expression, or "quadrature",

$$\int \frac{dx}{\sqrt{2z^{3/2} - x^3 + 3zx}} = \sqrt{\frac{2}{3}} \int dt\,.$$

After some work this integrates to

$$\frac{x(t) + \sqrt{z}}{3\sqrt{z}} = \operatorname{sech}^2\left(\frac{z^{1/4}t}{\sqrt{2}}\right).$$

From this equation, one may also compute the evolution of $\theta(t)$ on the homoclinic orbit by integrating the θ-equation,

$$\frac{d\theta}{dt} = -(x(t) + \sqrt{z}\,).$$

▲

12. Consider the solutions in the *complex* x-plane for $H < 0$. Also consider what happens in the complex x-plane when $z < 0$. To get started, take a look at [BeBrHo2008].

A.7 Lie derivatives and differential forms

Exterior calculus and actions of vector fields.

1 ***Operations on differential forms***

(i) Verify the following exterior derivative formulas,

(a) $df = \nabla f \cdot d\mathbf{x}\,,$

(b) $d(\mathbf{v} \cdot d\mathbf{x}) = \mathrm{curl}\,\mathbf{v} \cdot d\mathbf{S}\,,$

(c) $d(\boldsymbol{\omega} \cdot d\mathbf{S}) = \mathrm{div}\,\boldsymbol{\omega}\, d^{\,3}x\,.$

What does $d^2 = 0$ mean for each of (a), (b) and (c) above?

What does Stokes theorem $\int_\Omega d\alpha = \oint_{\partial\Omega} \alpha$ say about each of these relations?

(ii) Verify the following contraction formulas for $X = \mathbf{X} \cdot \nabla$,

(a) $X \lrcorner\, \mathbf{v} \cdot d\mathbf{x} = \mathbf{X} \cdot \mathbf{v}\,,$

(b) $X \lrcorner\, \boldsymbol{\omega} \cdot d\mathbf{S} = \boldsymbol{\omega} \times \mathbf{X} \cdot d\mathbf{x}\,,$

(c) $X \lrcorner\, f\, d^{\,3}x = f\mathbf{X} \cdot d\mathbf{S}.$

(iii) Verify the exterior derivatives of these contraction formulas for $X = \mathbf{X} \cdot \nabla$,

(a) $d(X \lrcorner\, \mathbf{v} \cdot d\mathbf{x}) = d(\mathbf{X} \cdot \mathbf{v}) = \nabla(\mathbf{X} \cdot \mathbf{v}) \cdot d\mathbf{x}\,,$

(b) $d(X \lrcorner\, \boldsymbol{\omega} \cdot d\mathbf{S}) = d(\boldsymbol{\omega} \times \mathbf{X} \cdot d\mathbf{x}) = \mathrm{curl}\,(\boldsymbol{\omega} \times \mathbf{X}) \cdot d\mathbf{S}\,,$

(c) $d(X \lrcorner\, f\, d^{\,3}x) = d(f\mathbf{X} \cdot d\mathbf{S}) = \mathrm{div}\,(f\mathbf{X})\, d^{\,3}x\,.$

(iv) Verify the following Lie derivative formulas,

(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) $\pounds_X(\boldsymbol{\omega} \cdot d\mathbf{S}) = \big(\mathrm{curl}\,(\boldsymbol{\omega} \times \mathbf{X}) + \mathbf{X}\,\mathrm{div}\,\boldsymbol{\omega}\big) \cdot d\mathbf{S}$
$= \big(-\boldsymbol{\omega} \cdot \nabla\mathbf{X} + \mathbf{X} \cdot \nabla\boldsymbol{\omega} + \boldsymbol{\omega}\,\mathrm{div}\,\mathbf{X}\big) \cdot d\mathbf{S}\,,$

(c′) $(*\pounds_X * \omega^\flat)^\sharp = [X, \omega] + (\mathrm{div}\,\mathbf{X})\omega,$
with $\omega^\flat := \boldsymbol{\omega} \cdot d\mathbf{x}$ and $*\omega^\flat := \boldsymbol{\omega} \cdot d\mathbf{S}$,

(d) $\pounds_X(f\, d^{\,3}x) = (\mathrm{div}\, f\mathbf{X})\, d^{\,3}x.$

(v) Verify the following Lie derivative identities by using Cartan's formula,

$$\pounds_X \alpha = X \lrcorner\, d\alpha + d(X \lrcorner\, \alpha)\,,$$

for a k-form α, $k = 0, 1, 2, 3$ in $\mathbb{R}^3$,

(a) $£_{fX}\alpha = f£_X\alpha + df \wedge (X \lrcorner\, \alpha)\,,$

(b) $£_X d\alpha = d\big(£_X\alpha\big)\,,$

(c) $£_X(X \lrcorner\, \alpha) = X \lrcorner\, £_X\alpha\,,$

(d) $£_X(\alpha \wedge \beta) = (£_X\alpha) \wedge \beta + \alpha \wedge £_X\beta\,.$

2 ***Operations among vector fields***

The Lie derivative of one vector field by another is called the ***Jacobi-Lie bracket***, defined as

$$£_X Y := [X\,,\,Y] := \nabla Y \cdot X - \nabla X \cdot Y = -\,£_Y X\,.$$

In components, the Jacobi-Lie bracket is

$$[X\,,\,Y] = \Big[X^k\frac{\partial}{\partial x^k}\,,\,Y^l\frac{\partial}{\partial x^l}\Big] = \left(X^k\frac{\partial Y^l}{\partial x^k} - Y^k\frac{\partial X^l}{\partial x^k}\right)\frac{\partial}{\partial x^l}\,.$$

The Jacobi-Lie bracket among vector fields satisfies the Jacobi identity,

$$[\,X\,,\,[Y\,,\,Z]\,] + [\,Y\,,\,[Z\,,\,X]\,] + [\,Z\,,\,[X\,,\,Y]\,] = 0\,.$$

Verify the following formulas:

(a) $X \lrcorner\,(Y \lrcorner\, \alpha) = -\,Y \lrcorner\,(X \lrcorner\, \alpha)\,,$

(b) $[X\,,\,Y] \lrcorner\, \alpha = £_X(Y \lrcorner\, \alpha) - Y \lrcorner\,(£_X\alpha)$, for zero-forms (functions) and one-forms.

(c) $£_{[X\,,\,Y]}\alpha = £_X£_Y\alpha - £_Y£_X\alpha$, as a result of **(b)**.

(d) Verify the Jacobi identity for the action of Lie derivatives with respect to vector fields X, Y, Z, on a differential form α.

★

Answer.
Problems **1(i)-1(iv)** are verified easily from Cartan's formula,

$$\pounds_X \alpha = X \,\lrcorner\, d\alpha + d(X \,\lrcorner\, \alpha)\,,$$

for $\alpha \in \Lambda^k$. Problems **1.v(a)-(d)** are answered as follows:

1 ***Lie derivative identities by using Cartan's formula***

1.v(a) By its linearity, contraction satisfies

$$d(fX \,\lrcorner\, \alpha) = fd(X \,\lrcorner\, \alpha) + df \wedge (X \,\lrcorner\, \alpha)\,.$$

Hence, the Lie derivative satisfies,

$$\begin{aligned} \pounds_{fX}\alpha &= fX \,\lrcorner\, d\alpha + d(fX \,\lrcorner\, \alpha) \\ &= fX \,\lrcorner\, d\alpha + fd(X \,\lrcorner\, \alpha) + df \wedge (X \,\lrcorner\, \alpha) \\ &= f\pounds_X\alpha + df \wedge (X \,\lrcorner\, \alpha)\,. \end{aligned}$$

1.v(b) Cartan's formula implies $\pounds_X d\alpha = d\big(\pounds_X\alpha\big)$ by comparing the definitions:

$$\begin{aligned} \pounds_X d\alpha &= X \,\lrcorner\, d^2\alpha + d(X \,\lrcorner\, d\alpha)\,, \\ d\big(\pounds_X\alpha\big) &= d(X \,\lrcorner\, d\alpha) + d^2(X \,\lrcorner\, \alpha)\,. \end{aligned}$$

By $d^2 = 0$, these both equal $d(X \,\lrcorner\, d\alpha)$ and the result follows.

1.v(c) One also proves $\pounds_X(X \,\lrcorner\, \alpha) = X \,\lrcorner\, \pounds_X\alpha$ by comparing the definitions:

$$\begin{aligned} \pounds_X(X \,\lrcorner\, \alpha) &= X \,\lrcorner\, d(X \,\lrcorner\, \alpha) + d(X \,\lrcorner\, (X \,\lrcorner\, \alpha))\,, \\ X \,\lrcorner\, \pounds_X\alpha &= X \,\lrcorner\, d(X \,\lrcorner\, \alpha) + X \,\lrcorner\, (X \,\lrcorner\, d\alpha)\,. \end{aligned}$$

By $X \,\lrcorner\, (X \,\lrcorner\, \alpha) = 0$, these both equal $X \,\lrcorner\, d(X \,\lrcorner\, \alpha)$ and the result follows.

1.v(d) $\pounds_X(\alpha \wedge \beta) = (\pounds_X\alpha) \wedge \beta + \alpha \wedge \pounds_X\beta$

This important identity follows immediately from the product rule for the dynamical definition of the Lie derivative in (3.4.1)

$$\pounds_X\alpha = \frac{d}{dt}\bigg|_{t=0}(\phi_t^*\alpha)\,,$$

and the formula (3.3.7) for the action of the pull-back on the wedge product of differential forms, rewritten here as

$$\phi_t^*(\alpha \wedge \beta) = \phi_t^*\alpha \wedge \phi_t^*\beta \,.$$

The product rule for the Lie derivative also follows from Cartan's formula, when the two defining properties for contraction

$$X \mathbin{\lrcorner} (\alpha \wedge \beta) = (X \mathbin{\lrcorner} \alpha) \wedge \beta + (-1)^k \alpha \wedge (X \mathbin{\lrcorner} \beta)\,,$$

and exterior derivative

$$d(\alpha \wedge \beta) = (d\alpha) \wedge \beta + (-1)^k \alpha \wedge d\beta\,,$$

are invoked. (The corresponding underlined terms cancel in the computation below.)

$$\begin{aligned}
\pounds_X(\alpha \wedge \beta) &= (X \mathbin{\lrcorner} d\alpha) \wedge \beta + \underline{[(-1)^{k+1} d\alpha \wedge (X \mathbin{\lrcorner} \beta)]} \\
&\quad + \underline{\underline{[(-1)^k (X \mathbin{\lrcorner} \alpha) \wedge d\beta\,]}} + (-1)^{2k} \alpha \wedge (X \mathbin{\lrcorner} d\beta) \\
&\quad + d(X \mathbin{\lrcorner} \alpha) \wedge \beta + \underline{\underline{[(-1)^{k-1} (X \mathbin{\lrcorner} \alpha) \wedge d\beta\,]}} \\
&\quad + \underline{[(-1)^k d\alpha \wedge (X \mathbin{\lrcorner} \beta)]} + (-1)^{2k} \alpha \wedge d(X \mathbin{\lrcorner} \beta) \\
&= (X \mathbin{\lrcorner} d\alpha + d(X \mathbin{\lrcorner} \alpha)) \wedge \beta \\
&\quad + (-1)^{2k} \alpha \wedge (X \mathbin{\lrcorner} d\beta + d(X \mathbin{\lrcorner} \beta)) \\
&= (\pounds_X \alpha) \wedge \beta + \alpha \wedge \pounds_X \beta\,.
\end{aligned}$$

2 ***Verifying the formulas for operations among vector fields***

2(a) By direct substitution

$$\begin{aligned}
X \mathbin{\lrcorner} (Y \mathbin{\lrcorner} \alpha) &= X^l Y^m \alpha_{mli_3\ldots i_k} dx^{i_3} \wedge \cdots \wedge dx^{i_k} \\
&= -\,X^l Y^m \alpha_{lmi_3\ldots i_k} dx^{i_3} \wedge \cdots \wedge dx^{i_k} \\
&= -\,Y \mathbin{\lrcorner} (X \mathbin{\lrcorner} \alpha)\,,
\end{aligned}$$

by antisymmetry of $\alpha_{mli_3\ldots i_k}$ in its first two indices.

2(b) For zero-forms (functions) all terms in the formula vanish identically. The formula

$$[X\,,\,Y]\,\lrcorner\,\alpha = \pounds_X(Y\,\lrcorner\,\alpha) - Y\,\lrcorner\,(\pounds_X\alpha)\,,$$

is seen to hold for a one-form $\alpha = \mathbf{v}\cdot d\mathbf{x}$ by comparing

$$[X\,,\,Y]\,\lrcorner\,\alpha \quad = \quad (X^kY^l_{,k} - Y^kX^l_{,k})v_l\,,$$

with

$$\begin{aligned}&\pounds_X(Y\,\lrcorner\,\alpha) - Y\,\lrcorner\,(\pounds_X\alpha)\\ &\qquad\qquad = X^k\partial_k(Y^lv_l) - Y^l(X^kv_{l,k} + v_jX^j_{,l})\,,\end{aligned}$$

to see that it holds in an explicit calculation.

(By a general theorem [AbMa1978], verification for zero-forms and one-forms is sufficient to imply the result for all k-forms. Notice that exercise 1.iv(c) is an example for 3-forms. Try writing the formula in vector notation for 2-forms!)

Remark A.7.1 *In fact, the general form of the relation required in part* **2(b)** *follows immediately from the product rule for the dynamical definition of the Lie derivative. By equation (3.3.10), insertion of a vector field into a k-form transforms under the flow ϕ_t of a smooth vector field Y as*

$$\phi_t^*(Y(m)\,\lrcorner\,\alpha) = Y(\phi_t(m))\,\lrcorner\,\phi_t^*\alpha\,.$$

A direct computation using the dynamical definition of the Lie derivative in (3.4.1)

$$\pounds_X\alpha = \frac{d}{dt}\Big|_{t=0}(\phi_t^*\alpha)\,,$$

then yields

$$\begin{aligned}\frac{d}{dt}\Big|_{t=0}\phi_t^*(Y\,\lrcorner\,\alpha) \quad &= \quad \Big(\frac{d}{dt}\Big|_{t=0}Y(\phi_t(m))\Big)\,\lrcorner\,\alpha\\ &\qquad + Y\,\lrcorner\,\Big(\frac{d}{dt}\Big|_{t=0}\phi_t^*\alpha\Big)\,.\end{aligned}$$

Hence, we recognise the desired formula in part **2(b)** *written as a* ***product rule****:*

$$\pounds_X(Y \,\lrcorner\, \alpha) = (\pounds_X Y) \,\lrcorner\, \alpha + Y \,\lrcorner\, (\pounds_X \alpha)\,,$$

which may also be written in an equivalent notation as

$$\pounds_X(Y(\alpha)) = (\pounds_X Y)(\alpha) + Y(\pounds_X \alpha)\,.$$

For Lie derivative $\pounds_u$ of a 2-form ω^2 this is

$$\pounds_u \omega(v_1, v_2) = (\pounds_u \omega)(v_1, v_2) + \omega(\pounds_u v_1, v_2) + \omega(v_1, \pounds_u v_2)\,,$$

and for a k-form,

$$\begin{aligned} \pounds_u \omega(v_1, v_2, \dots, v_k) &= (\pounds_u \omega)(v_1, v_2, \dots, v_k) \\ &\quad + \omega(\pounds_u v_1, v_2, \dots, v_k) + \dots \\ &\quad + \omega(v_1, v_2, \dots, \pounds_u v_k)\,. \end{aligned}$$

2(c) Given $[X\,,\,Y] \,\lrcorner\, \alpha = \pounds_X(Y \,\lrcorner\, \alpha) - Y \,\lrcorner\, (\pounds_X \alpha)$ as verified in part **2(b)** we use Cartan's formula to compute

$$\begin{aligned} \pounds_{[X\,,\,Y]}\alpha &= d([X\,,\,Y] \,\lrcorner\, \alpha) + [X\,,\,Y] \,\lrcorner\, d\alpha \\ &= d\big(\pounds_X(Y \,\lrcorner\, \alpha) - Y \,\lrcorner\, (\pounds_X \alpha)\big) \\ &\quad + \pounds_X(Y \,\lrcorner\, d\alpha) - Y \,\lrcorner\, (\pounds_X d\alpha) \\ &= \pounds_X d(Y \,\lrcorner\, \alpha) - d(Y \,\lrcorner\, (\pounds_X \alpha) \\ &\quad + \pounds_X(Y \,\lrcorner\, d\alpha) - Y \,\lrcorner\, d(\pounds_X \alpha) \\ &= \pounds_X(\pounds_Y \alpha) - \pounds_Y(\pounds_X \alpha)\,, \end{aligned}$$

as required. Thus, the product rule for Lie derivative of a contraction obtained in answering problem **2(b)** provides the key to solving **2(c)**. ■

Consequently,

$$\begin{aligned}\pounds_{[Z,[X,Y]]}\alpha &= \pounds_Z\pounds_X\pounds_Y\alpha - \pounds_Z\pounds_Y\pounds_X\alpha \\ &\quad - \pounds_X\pounds_Y\pounds_Z\alpha + \pounds_Y\pounds_X\pounds_Z\alpha\,,\end{aligned}$$

and summing over cyclic permutations immediately verifies that

$$\pounds_{[Z,[X,Y]]}\,\alpha + \pounds_{[X,[Y,Z]]}\,\alpha + \pounds_{[Y,[Z,X]]}\,\alpha = 0\,.$$

This is the ***Jacobi identity for the Lie derivative***. ■

▲

Appendix B

Exercises for review and further study

B.1 The reduced Kepler problem: Newton [1686]

Newton's equation [Ne1686] for the ***reduced Kepler problem*** of planetary motion is

$$\ddot{\mathbf{r}} + \frac{\mu \mathbf{r}}{r^3} = 0\,, \tag{B.1.1}$$

in which μ is a constant and $r = |\mathbf{r}|$ with $\mathbf{r} \in \mathbb{R}^3$.

Scale invariance of this equation under the changes $R \to s^2 R$ and $T \to s^3 T$ in the units of space R and time T for any constant (s) means that it admits families of solutions whose space and time scales are related by $T^2/R^3 = const$. This is ***Kepler's Third Law***.

1. Show that Newton's equation (B.1.1) conserves the quantities,

$$\begin{aligned} E &= \frac{1}{2}|\dot{\mathbf{r}}|^2 - \frac{\mu}{r} \quad \text{(energy)}\,, \\ \mathbf{L} &= \mathbf{r} \times \dot{\mathbf{r}} \quad \text{(specific angular momentum)}\,. \end{aligned}$$

Since, $\mathbf{r} \cdot \mathbf{L} = 0$, the planetary motion in $\mathbb{R}^3$ takes place in a plane to which vector $\mathbf{L}$ is perpendicular. This is the orbital plane.

2. The unit vectors for polar coordinates in the orbital plane are $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\theta}}$. Show that these vectors satisfy

$$\frac{d\hat{\mathbf{r}}}{dt} = \dot{\theta}\,\hat{\boldsymbol{\theta}} \quad \text{and} \quad \frac{d\hat{\boldsymbol{\theta}}}{dt} = -\,\dot{\theta}\,\hat{\mathbf{r}}\,, \quad \text{where} \quad \dot{\theta} = \frac{L}{r^2}\,.$$

Show that Newton's equation (B.1.1) also conserves the following two vector quantities,

$$\begin{aligned} \mathbf{K} &= \dot{\mathbf{r}} - \frac{\mu}{L}\hat{\boldsymbol{\theta}} \quad \text{(Hamilton's vector)},\\ \mathbf{J} &= \dot{\mathbf{r}} \times \mathbf{L} - \mu\mathbf{r}/r \quad \text{(Laplace-Runge-Lenz vector)}\,, \end{aligned}$$

which both lie in the orbital plane, since $\mathbf{J}\cdot\mathbf{L} = 0 = \mathbf{K}\cdot\mathbf{L}$. Hint: How are these two vectors related? Their constancy means that certain attributes of the orbit, particularly, its orientation, are fixed in the orbital plane.

3. From their definitions, show that these conserved quantities are related by

$$\begin{aligned} L^2 + \frac{J^2}{(-2E)} &= \frac{\mu^2}{(-2E)} \\ \text{and} \quad \mathbf{J}\cdot\mathbf{K}\times\mathbf{L} &= K^2L^2 = J^2\,, \end{aligned} \tag{B.1.2}$$

where $J^2 := |\mathbf{J}|^2$, etc. and $-2E > 0$ for bounded orbits.

4. Orient the conserved Laplace-Runge-Lenz vector vector $\mathbf{J}$ in the orbital plane to point along the reference line for the measurement of the polar angle θ, say from the center of the orbit (Sun) to the perihelion (point of nearest approach, at midsummer's day), so that

$$\mathbf{r}\cdot\mathbf{J} = rJ\cos\theta = \mathbf{r}\cdot(\dot{\mathbf{r}}\times\mathbf{L} - \mu\,\mathbf{r}/r)\,.$$

Use this relation to write the Kepler orbit $r(\theta)$ in plane polar coordinates, as

$$r(\theta) = \frac{L^2}{\mu + J\cos\theta} = \frac{l_\perp}{1 + e\cos\theta}\,, \tag{B.1.3}$$

with eccentricity $e = J/\mu$ and semi latus rectum $l_\perp = L^2/\mu$. The expression $r(\theta)$ for the Kepler orbit is the formula for a conic section. This is ***Kepler's First Law***. How is the value of the eccentricity associated to the types of orbits?

5. Use conservation of L to show that constancy of magnitude $L = |\mathbf{L}|$ means the orbit sweeps out equal areas in equal times. This is ***Kepler's Second Law***. For an elliptical orbit, write the period in terms of angular momentum and the area.

6. Use the result of Part 5 and the geometric properties of ellipses to show that the period of the orbit is given by

$$\left(\frac{T}{2\pi}\right)^2 = \frac{a^3}{\mu} = \frac{\mu^2}{(-2E)^3}$$

The relation $T^2/a^3 = \textit{constant}$ is Kepler's Third Law. The constant is Newton's constant.

7. Write the Kepler motion equation (B.1.1) in Hamiltonian form

$$\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}}\,, \qquad \dot{\mathbf{p}} = -\,\frac{\partial H}{\partial \mathbf{q}}\,.$$

Identify the position $\mathbf{q}$ and its canonical momentum $\mathbf{p}$ in terms of $\mathbf{r}$ and $\dot{\mathbf{r}}$. Write the Hamiltonian $H(\mathbf{q}, \mathbf{p})$ explicitly.

8. Fill in a table of canonical Poisson brackets

$$\big\{F,\, H\big\} := \frac{\partial F}{\partial \mathbf{q}} \cdot \frac{\partial H}{\partial \mathbf{p}} - \frac{\partial F}{\partial \mathbf{p}} \cdot \frac{\partial H}{\partial \mathbf{q}}\,, \tag{B.1.4}$$

among the variables in the set $\{q_k, p_k, |\mathbf{q}|^2, |\mathbf{p}|^2, \mathbf{q}\cdot\mathbf{p}\}$.

9. Write the Poisson brackets among the set of quadratic combinations

$$X_1 = |\mathbf{q}|^2 \geq 0\,, \quad X_2 = |\mathbf{p}|^2 \geq 0\,, \quad X_3 = \mathbf{p}\cdot\mathbf{q}\,, \tag{B.1.5}$$

as an $\mathbb{R}^3$-bracket and identify its Casimir function.

10. Three fundamental conserved phase-space variables in the Kepler problem are its Hamiltonian H, its angular momentum $\mathbf{L}$ and its Laplace-Runge-Lenz (LRL) vector $\mathbf{J}$. These are given by

$$\begin{aligned} H &= \tfrac{1}{2}|\mathbf{p}|^2 - \frac{\mu}{|\mathbf{q}|}\,, \\ \mathbf{L} &= \mathbf{q}\times\mathbf{p}\,, \\ \mathbf{J} &= \mathbf{p}\times(\mathbf{q}\times\mathbf{p}) - \mu\,\mathbf{q}/|\mathbf{q}| \\ &= -\,(\mathbf{q}\cdot\mathbf{p})\,\mathbf{p} - 2H\,\mathbf{q}\,. \end{aligned}$$

What does preservation of the LRL vector $\mathbf{J}$ and angular momentum vector $\mathbf{L}$ imply about the shape and orientation of a given planar orbit?

11. Check whether the Poisson brackets amongst the components of the vectors $\mathbf{L}$ and $\mathbf{J}$ satisfy the following relations:

$$\begin{aligned} \{L_i,\, L_j\} &= \epsilon_{ijk}L_k\,, \\ \{L_i,\, J_j\} &= \epsilon_{ijk}J_k\,, \\ \{J_i,\, J_j\} &= -\,2H\epsilon_{ijk}L_k\,. \end{aligned}$$

(Note the sign of last term: $-2H > 0$ for bounded orbits.) Importantly, this means that

$$\{J_i,\, J^2\} = -\,4H\epsilon_{ijk}J_jL_k = -\,4H(\mathbf{J}\times\mathbf{L})_i = -\,2H\{L^2,\, J_i\}\,.$$

The conservation laws $\{L^2, H\} = 0$ and $\{J_i, H\} = 0$ allow the use of formula (B.1.2) to check consistency of the previous Poisson bracket relations, since that formula implies

$$\Big\{J_i, \big(J^2 - 2H\,L^2\big)\Big\} = \Big\{J_i, \mu^2\Big\} = 0\,.$$

Upon referring to the relationships between the orbital parameters and the conservation laws derived in (B.1.3), explain how the canonical transformations generated by $\mathbf{J}$ affect the (i) energy, (ii) eccentricity and (iii) width of the orbit.

12. Suppose Newton had postulated the Yukawa potential with length scale $\alpha > 0$

$$V(r) = -\mu \frac{e^{-r/\alpha}}{r}$$

as the potential energy for planetary motion. Which of Kepler's Laws would have survived? Write the evolution equation for the LRL vector for Newtonian dynamics with a Yukawa potential.

B.2 Hamiltonian reduction by stages

1. For two $\mathbb{R}^3$ vectors $\mathbf{M}$ and $\mathbf{N}$, write Hamilton's equations using the Poisson brackets among the components,

$$\{M_i,\, M_j\} = \epsilon_{ijk} M_k\,, \quad \{\, N_i,\, N_j\,\} = \epsilon_{ijk} N_k\,, \quad \{M_i,\, N_j\} = 0\,,$$

2. Compute the equations of motion and identify the functionally independent conserved quantities for the following two Hamiltonians

$$H_1 = \hat{\mathbf{z}} \cdot (\mathbf{M} \times \mathbf{N}) \quad \text{and} \quad H_2 = \mathbf{M} \cdot \mathbf{N}\,. \tag{B.2.1}$$

3. Determine whether these Hamiltonians have sufficiently many symmetries and associated conservation laws to be completely integrable (i.e., reducible to Hamilton's canonical equations for a single degree of freedom) and explain why.

4. Transform the Hamiltonians in (B.2.1) from Cartesian components of the vectors $(\mathbf{M}, \mathbf{N}) \in \mathbb{R}^3 \times \mathbb{R}^3$ into spherical coordinates $(\theta, \phi) \in S^2$ and $(\bar{\theta}, \bar{\phi}) \in S^2$, respectively.

5. Use the S^1 symmetries of the Hamiltonians H_1, H_2 and their associated conservation laws to reduce the dynamics of $(\mathbf{M},\, \mathbf{N})$ in $\mathbb{R}^3 \times \mathbb{R}^3$ to canonical Hamiltonian equations. First reduce to $S^2 \times S^2$ and then to S^2 by a two-stage sequence of canonical transformations.

B.3 $\mathbb{R}^3$ bracket for the spherical pendulum

1. Identify the degrees of freedom for the spherical pendulum in $T\mathbb{R}^3$ and show that the constraint of constant length reduces the number of degrees of fredom from three to two by sending $T\mathbb{R}^3 \to T\mathbb{S}^2$.

2. Formulate the constrained Lagrangian in $T\mathbb{R}^3$ and show that its symmetry under rotations about the vertical axis implies a conservation law for angular momentum via Noether's theorem.

3. Derive the motion equations on $T\mathbb{R}^3$ from Hamilton's principle. Note that these equations preserve the defining conditions for
$$T\mathbb{S}^2 : \{(\mathbf{x}, \dot{\mathbf{x}}) \in T\mathbb{R}^3 \,\big|\, \|\mathbf{x}\|^2 = 1 \text{ and } \mathbf{x}\cdot\dot{\mathbf{x}} = 0\}\,.$$
That is, $T\mathbb{S}^2$ is an invariant manifold of the equations in $\mathbb{R}^3$. Conclude that the constraints for remaining on $T\mathbb{S}^2$ may be regarded as dynamically preserved initial conditions for the spherical pendulum equations in $T\mathbb{R}^3$.

4. Legendre transform the Lagrangian defined on $T\mathbb{R}^3$ to find a constrained Hamiltonian (Routhian) with variables $(\mathbf{x}, \mathbf{y}) \in T^*\mathbb{R}^3$ whose dynamics preserves $T\mathbb{S}^2$.

5. There are six independent linear and quadratic variables in $T^*\mathbb{R}^3/S^1$
$$\begin{aligned} \sigma_1 &= x_3 & \sigma_3 &= y_1^2 + y_2^2 + y_3^2 & \sigma_5 &= x_1y_1 + x_2y_2 \\ \sigma_2 &= y_3 & \sigma_4 &= x_1^2 + x_2^2 & \sigma_6 &= x_1y_2 - x_2y_1 \end{aligned}$$
These are not independent. They satisfy a cubic algebraic relation. Find this relation and write the TS^2 constraints in terms of the S^1 invariants.

6. Write closed Poisson brackets among the six independent linear and quadratic S^1-invariant variables
$$\sigma_k \in T^*\mathbb{R}^3/S^1, \quad k = 1, 2, \ldots, 6.$$

7. Show that the two quantities

$$\sigma_3(1-\sigma_1^2)-\sigma_2^2-\sigma_6^2=0 \quad \text{and} \quad \sigma_6$$

are Casimirs for the Poisson brackets on $T^*\mathbb{R}^3/S^1$.

8. Use the orbit map $T\mathbb{R}^3 \to \mathbb{R}^6$

$$\pi:\ (\mathbf{x},\mathbf{y}) \to \{\sigma_j(\mathbf{x},\mathbf{y}),\ j=1,\dots,6\} \tag{B.3.1}$$

to transform the energy Hamiltonian to S^1-invariant variables.

9. Find the reduction $T^*\mathbb{R}^3/S^1 \cap T\mathbb{S}^2 \to \mathbb{R}^3$. Show that the motion follows the intersections of level surfaces of angular momentum and energy in $\mathbb{R}^3$. Compute the associated Nambu bracket in $\mathbb{R}^3$ and use it to characterise the types of motion available in the motion of this system.

10. Write the Hamiltonian, Poisson bracket and equations of motion in terms of the variables $\sigma_k \in T^*\mathbb{R}^3/S^1$, $k=1,2,\dots,3$.

11. Interpret the solutions geometrically as intersections of Hamiltonian level sets (planes in $\mathbb{R}^3$) with a family of cup-shaped surfaces (Casimirs of the $\mathbb{R}^3$ bracket depending on angular momentum) whose limiting surface at zero angular momentum is a *pinched cup*.

12. Show that this geometrical interpretation implies that the two equilibria along the vertical axis at the North and South poles have the expected opposite stability. Explain why this was to be expected.

13. Show that the unstable vertical equilibrium at the North pole is connected to itself by homoclinic orbits.

14. Show that all other orbits are periodic.

15. Reduce the dynamics on a family of planes representing level sets of the Hamiltonian to single particle motion in a phase plane and compute the behaviour of its solutions. Identify its critical points and their stability.

16. See [CuBa1997] for further analysis of this problem, including a study of its monodromy by using the energy-momentum map.

B.4 Maxwell-Bloch equations

The real-valued Maxwell-Bloch system for $\mathbf{x} = (x_1, x_2, x_3)^T \in \mathbb{R}^3$ is given by

$$\dot{x}_1 = x_2\,, \quad \dot{x}_2 = x_1 x_3\,, \quad \dot{x}_3 = -x_1 x_2\,.$$

1. Write this system in three-dimensional vector $\mathbb{R}^3$-bracket notation as
$$\dot{\mathbf{x}} = \nabla H_1 \times \nabla H_2\,,$$
where H_1 and H_2 are two conserved functions, one of whose level sets (let it be H_1) may be taken as ***circular cylinders*** oriented along the x_1-direction and the other (let it be H_2) whose level sets may be taken as ***parabolic cylinders*** oriented along the x_2-direction.

2. Restrict the equations and their $\mathbb{R}^3$ Poisson bracket to a level set of H_2. Show that the Poisson bracket on the circular cylinder $H_2 = const$ is symplectic.

3. Derive the equation of motion on a level set of H_2 and express them in the form of Newton's Law. Do they reduce to something familiar?

4. Identify steady solutions and determine which are unstable (saddle points) and which are stable (centers).

5. Determine the geometric and dynamic phases of a closed orbit on a level set of H_2.

B.5 Modulation equations

The real 3-wave modulation equations on $\mathbb{R}^3$ are

$$\dot{X}_1 = X_2 X_3\,, \quad \dot{X}_2 = X_3 X_1\,, \quad \dot{X}_3 = -X_1 X_2\,.$$

1. Write these equations using an $\mathbb{R}^3$ bracket of the form,

$$\dot{\mathbf{X}} = \nabla C \times \nabla H \,,$$

   where level sets of C and H are each circular cylinders.

2. Characterise the equilibrium points geometrically in terms of the gradients of C and H. How many are there? Which are stable?

3. Choose cylindrical polar coordinates along the axis of the circular cylinder that represents the level set of C and restrict the $\mathbb{R}^3$ Poisson bracket to that level set. Show that the Poisson bracket on the parabolic cylinder C is symplectic.

4. Write the equations of motion on that level set. Do they reduce to something familiar?

5. Determine the geometric and dynamic phases of a closed orbit on a level set of C.

B.6 The Hopf map

In coordinates $(a_1, a_2) \in \mathbb{C}^2$, the Hopf map $\mathbb{C}^2/S^1 \to S^3 \to S^2$ is obtained by transforming to the four quadratic S^1-invariant quantities

$$(a_1, a_2) \to Q_{jk} = a_j a_k^* \,, \quad \text{with} \quad j, k = 1, 2 \,.$$

Let the $\mathbb{C}^2$ coordinates be expressed as

$$a_j = q_j + i p_j$$

in terms of canonically conjugate variables satisfying the fundamental Poisson brackets

$$\{q_k, p_m\} = \delta_{km} \quad \text{with} \quad k, m = 1, 2.$$

1. Compute the Poisson brackets $\{a_j, a_k^*\}$ for $j, k = 1, 2$.

2. Is the transformation $(q, p) \to (a, a^*)$ canonical? Explain why, or why not.

3. Compute the Poisson brackets among the Q_{jk}, with $j, k = 1, 2$.

4. Make the linear change of variables,

$$X_0 = Q_{11} + Q_{22}\,, \quad X_1 + iX_2 = Q_{12}\,, \quad X_3 = Q_{11} - Q_{22}\,.$$

   Compute the Poisson brackets among the (X_0, X_1, X_2, X_3).

5. Express the Poisson bracket $\{F(\mathbf{X}), H(\mathbf{X})\}$ in vector form among functions F and H of $\mathbf{X} = (X_1, X_2, X_3)$

6. Show that the quadratic invariants (X_0, X_1, X_2, X_3) themselves satisfy a quadratic relation. How is this relevant to the Hopf map?

B.7 2:1 resonant oscillators

The Hamiltonian $\mathbb{C}^2 \to \mathbb{R}$ for a certain $2:1$ resonance is given by

$$H = \tfrac{1}{2}|a_1|^2 - |a_2|^2 + \tfrac{1}{2}\mathrm{Im}(a_1^{*2} a_2)\,,$$

in terms of canonical variables $(a_1, a_1^*, a_2, a_2^*) \in \mathbb{C}^2$ whose Poisson bracket relation is

$$\{a_j, a_k^*\} = -2i\delta_{jk}, \quad \text{for} \quad j, k = 1, 2.$$

is invariant under the $2:1$ resonance S^1 transformation

$$a_1 \to e^{i\phi} \quad \text{and} \quad a_2 \to e^{2i\phi}.$$

1. Write the motion equations in terms of the canonical variables $(a_1, a_1^*, a_2, a_2^*) \in \mathbb{C}^2$

2. Introduce the orbit map $\mathbb{C}^2 \to \mathbb{R}^4$

$$\pi:\ (a_1, a_1^*, a_2, a_2^*) \to \{X, Y, Z, R)\} \qquad \text{(B.7.1)}$$

and transform the Hamiltonian H on $\mathbb{C}^2$ to new variables $X, Y, Z, R \in \mathbb{R}^4$ given by

$$\begin{aligned} R &= \tfrac{1}{2}|a_1|^2 + |a_2|^2 \,, \\ Z &= \tfrac{1}{2}|a_1|^2 - |a_2|^2 \,, \\ X + iY &= 2a_1^{*2} a_2 \,, \end{aligned}$$

that are invariant under the $2:1$ resonance S^1 transformation.

3. Show that these variables are functionally dependent, because they satisfy a cubic algebraic relation $C(X, Y, Z, R) = 0$.

4. Use the orbit map $\pi : \mathbb{C}^2 \to \mathbb{R}^4$ to make a table of Poisson brackets among the four quadratic $2 : 1$ resonance S^1-invariant variables $X, Y, Z, R \in \mathbb{R}^4$.

5. Show that both R and the cubic algebraic relation $C(X, Y, Z, R) = 0$ are Casimirs for these Poisson brackets.

6. Write the Hamiltonian, Poisson bracket and equations of motion in terms of the remaining variables $\mathbf{X} = (X, Y, Z)^T \in \mathbb{R}^3$.

7. Describe this motion in terms of level sets of the Hamiltonian H and the orbit manifold for the 2:1 resonance, given by $C(X, Y, Z, R) = 0$.

8. Restrict the dynamics to a level set of the Hamiltonian and show that it reduces there to the equation of motion for a point particle in a cubic potential. Explain its geometrical meaning.

9. Compute the geometric and dynamic phases for any closed orbit on a level set of H.

B.8 A steady Euler fluid flow

A steady Euler fluid flow in a rotating frame satisfies

$$\pounds_u (\mathbf{v} \cdot d\mathbf{x}) = -\, d(p + \tfrac{1}{2}|\mathbf{u}|^2 - \mathbf{u} \cdot \mathbf{v}) \,,$$

where $\pounds_u$ is Lie derivative with respect to the divergenceless vector field $u = \mathbf{u} \cdot \nabla$, with $\nabla \cdot \mathbf{u} = 0$, and $\mathbf{v} = \mathbf{u} + \mathbf{R}$, with Coriolis parameter curl$\mathbf{R} = 2\boldsymbol{\Omega}$.

1. Write out this Lie-derivative relation in Cartesian coordinates.

2. By taking the exterior derivative, show that this relation implies that the exact two-form

$$\mathrm{curl}v \,\lrcorner\, d^3x = \mathrm{curl}\mathbf{v}\cdot\nabla \,\lrcorner\, d^3x = \mathrm{curl}\mathbf{v}\cdot d\mathbf{S} = d(\mathbf{v}\cdot d\mathbf{x}) =: d\Xi \wedge d\Pi$$

   is left invariant under the flow of the divergenceless vector field u, where we have used Darboux's theorem to introduce smooth potentials Ξ and Π to parameterise the exact 2-form.

3. Show that Cartan's formula for the Lie derivative in the steady Euler flow condition implies that

$$u \,\lrcorner\, \Big(\mathrm{curl}v \,\lrcorner\, d^3x\Big) = dH(\Xi, \Pi)$$

   and identify the function H.

4. Use the result of (3) to write $\pounds_u \Xi = \mathbf{u}\cdot\nabla\Xi$ and $\pounds_u \Pi = \mathbf{u}\cdot\nabla\Xi$ in terms of the partial derivatives of H.

5. What do the results of (4) mean geometrically? Hint: Is a symplectic form involved?

B.9 Dynamics of vorticity gradient

1. Write Euler's fluid equations in $\mathbb{R}^3$

$$\partial_t \mathbf{u} + \mathbf{u}\cdot\nabla\mathbf{u} = -\nabla p\,, \quad \mathrm{div}\,\mathbf{u} = 0\,, \tag{B.9.1}$$

   in geometric form using the Lie derivative of the circulation 1-form.

2. State and prove Kelvin's circulation theorem for Euler's fluid equations in geometric form.

3. The vorticity of the Euler fluid velocity $\mathbf{u} = (u, v, w) \in \mathbb{R}^3$ is given by $\boldsymbol{\omega} := \operatorname{curl} \mathbf{u}$. Write the Euler fluid equation for vorticity by taking the exterior derivative of the geometric form in part 1 above.

4. For Euler fluid motion restricted to the (x, y) plane with normal unit vector $\hat{\mathbf{z}}$, the vorticity equation for $\boldsymbol{\omega} = \omega\hat{\mathbf{z}}$ with scalar vorticity $\omega(x, y, t)$ simplifies to

$$\partial_t \omega + \mathbf{u} \cdot \nabla \omega = 0 \quad \text{with} \quad \mathbf{u} = (u, v, 0).$$

   Using this scalar vorticity dynamics, compute the equation for

$$\left(\frac{\partial}{\partial t} + \pounds_u \right) (dz \wedge d\omega)$$

   and express it as a 2D vector equation for the quantity

$$\mathcal{B} := \hat{\mathbf{z}} \times \nabla \omega \in \mathbb{R}^2.$$

5. How is $\mathcal{B}$ related to $\mathbf{u}$? Compare the result of part 4 for the dynamics of $\mathcal{B}$ with the dynamical equation for the vorticity vector in part 3 and the defining equation for the flow velocity in terms of the stream function.

B.10 The C. Neumann problem [1859]

For the origin of this problem see [Ne1859] and for some recent progress on it, see [De1978, Ra1981].

1. Derive the equations of motion

$$\ddot{\mathbf{x}} = -A\mathbf{x} + (A\mathbf{x} \cdot \mathbf{x} - \|\dot{\mathbf{x}}\|^2)\mathbf{x}$$

   of a particle of unit mass moving on the sphere S^{n-1} under the influence of a quadratic potential

$$V(\mathbf{x}) = \tfrac{1}{2} A\mathbf{x} \cdot \mathbf{x} = \tfrac{1}{2} a_1 x_1^2 + \tfrac{1}{2} a_2 x_2^2 + \cdots + \tfrac{1}{2} a_n x_n^2 \,,$$

for $\mathbf{x} \in \mathbb{R}^n$, where $A = \operatorname{diag}(a_1, a_2, \dots, a_n)$ is a fixed $n \times n$ diagonal matrix. Here $V(\mathbf{x})$ is a harmonic oscillator with spring constants that are taken to be fully anisotropic, with $a_1 < a_2 < \cdots < a_n$.

Hint: These are the Euler-Lagrange equations obtained when a Lagrange multiplier μ is used to restrict the motion to a sphere by adding a term,

$$\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}) = \tfrac{1}{2}\|\dot{\mathbf{x}}\|^2 - \tfrac{1}{2}A\mathbf{x} \cdot \mathbf{x} - \mu(1 - \|\mathbf{x}\|^2)\,, \tag{B.10.1}$$

on the tangent bundle

$$TS^{n-1} = \{(\mathbf{x}, \dot{\mathbf{x}}) \in \mathbb{R}^n \times \mathbb{R}^n |\ \|\mathbf{x}\|^2 = 1,\ \mathbf{x} \cdot \dot{\mathbf{x}} = 0\}.$$

2. Form the matrices

$$Q = (x^i x^j) \quad \text{and} \quad L = (x^i \dot{x}^j - x^j \dot{x}^i)\,,$$

and show that the Euler-Lagrange equations for the Lagrangian in (B.10.1) are equivalent to

$$\dot{Q} = [L, Q] \quad \text{and} \quad \dot{L} = [Q, A]\,.$$

Show further that for a constant parameter λ these Euler-Lagrange equations imply

$$\frac{d}{dt}(-Q + L\lambda + A\lambda^2) = [-Q + L\lambda + A\lambda^2, -L - A\lambda]\,.$$

Explain why this formula is important from the viewpoint of conservation laws.

3. Verify that the energy

$$E(Q, L) = -\frac{1}{4}\operatorname{trace}(L^2) + \frac{1}{2}\operatorname{trace}(AQ)$$

is conserved for this system.

4. Prove that the following $(n-1)$ quantities for $j = 1, 2, \ldots, n-1$ are also conserved

$$\Phi_j = \dot{x}_j^2 + \frac{1}{2} \sum_{i \neq j} \frac{(x^i \dot{x}^j - x^j \dot{x}^i)^2}{a_j - a_i} \,,$$

where $(\mathbf{x}, \dot{\mathbf{x}}) = (x_1, x_2, \ldots, x_n, \dot{x}_1, \dot{x}_2, \ldots, \dot{x}_n) \in TS^{n-1}$ and the a_j are the eigenvalues of the diagonal matrix A.

Bibliography

[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. Roy. Soc. (London) A* **318**, 349-409.

[AbMa1978] Abraham, R. and Marsden, J. E. [1978], *Foundations of Mechanics*. Addison-Wesley, second edition.

[AcHoKoTi1997] Aceves, A., Holm, D. D., Kovacic, G. and Timofeyev, I. [1997] Homoclinic orbits and chaos in a second-harmonic generating optical cavity. *Phys. Lett. A* **233**, 203-208.

[AbMaRa1988] Abraham, R., Marsden, J. E. and Ratiu, T. S. [1988], *Manifolds, Tensor Analysis, and Applications*. Volume 75 of *Applied Mathematical Sciences*. Springer-Verlag, second edition.

[AlLuMaRo1998] Alber, M. S., Luther, G. G., Marsden, J. E. and Robbins, J. M. [1998] Geometric phases, reduction and Lie-Poisson structure for the resonant three-wave interaction. *Phys. D*, **123**, pp. 271-290.

[AlLuMaRo1999] Alber, M. S., Luther, G. G., Marsden, J. E. and Robbins, J. M. [1999] Geometry and control of three-wave interactions, in *The Arnoldfest* (Toronto, ON, 1997), *Fields Inst. Commun.* **24**, AMS, Providence, RI, pp. 55-80.

[AlBaNe1990] Allen, J. S., Barth, J. A. and Newberger, P. A. [1990] On intermediate models for barotropic continental shelf and slope flow fields. Part I: Formulation and comparison of exact solutions. *J. of Phys. Oceanog.* **20**, 1017Ð1042

[AlHo1996] Allen, J. S., and Holm, D. D. [1996] Extended-geostrophic Hamiltonian models for rotating shallow water motion. *Physica D*, **98** 229-248.

[AlHoNe2002] Allen, J. S., Holm, D. D., and Newberger, P. A. [2002] Toward an extended-geostrophic Euler–Poincaré model for mesoscale oceanographic flow. In *Large-Scale Atmosphere-Ocean Dynamics 1: Analytical Methods and Numerical Models*. Edited by J. Norbury & I. Roulstone, Cambridge University Press: Cambridge, pp. 101-125.

[AlEb1975] Allen, L., and Eberly, J. H. [1975] *Optical Resonance and Two-Level Atoms*, Wiley: NewYork. (Reprinted by Dover, Mineola, NY, 1987)

[AmCZ96] Ambrosetti, A. and Coti Zelati, V. [1996], *Periodic Solutions of Singular Lagrangian Systems*. Birkhäuser

[AnMc1978] Andrews, D. G. and McIntyre, M. E. [1978] An exact theory of nonlinear waves on a Lagrangian-mean flow. *J. Fluid Mech.* **89**, 609–646.

[Ar350BC] Aristotle [350BC] *Physics*.

[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-Verlag.

[Ar1994] Arnold, V. I. [1994] *Topological Invariants of Plane Curves and Caustics*. Amer. Math. Soc. University Lecture Series **5**.

[ArKh1992] Arnold, V. I. and Khesin, B. A. [1992] Topological methods in hydrodynamics. *Ann. Rev. of Fluid Mech.* **24**, 145-166.

[ArKh1998] Arnold, V. I and Khesin, B. A. [1998] *Topological Methods in Hydrodynamics.* Springer: New York.

[Be1986] Benci, V. [1986], Periodic solutions of Lagrangian systems on a compact manifold, *Journal of Differential Equations* **63**, 135–161.

[BeBrHo2008] Bender, C. M., Brody, D. C. and Hook, D. W. [2008] Quantum effects in classical systems having complex energy, arXiv:0804.4169v1.

[BeBuHo1994] Berman, G. P., Bulgakov, E. N. and Holm, D. D. [1994] *Crossover-Time in Quantum Boson and Spin Systems*, Springer Lecture Notes in Physics: Monographs **m21**, Springer: New York.

[BlBr92] Blanchard, P. and Bruning, E. [1992], *Variational Methods in Mathematical Physics*. Springer Verlag.

[Bl2004] Bloch, A. M. (with the collaboration of: Baillieul, J., Crouch, P. E. and Marsden, J. E.) [2004] *Nonholonomic Mechanics and Control*, Springer-Verlag NY.

[BlBrCr1997] Bloch, A. M., Brockett, R. W. and Crouch, P. E. [1997] Double bracket equations and geodesic flows on symmetric spaces. *Comm. Math Phys* **187**, 357-373.

[BlCr1997] Bloch, A. M. and Crouch, P. E. [1996] Optimal control and geodesic flows. *Systems and Control Letters* **28**, 65-72.

[BlCrHoMa2001] Bloch, A.M., Crouch, P. E., Holm, D. D. and Marsden, J. E. [2001] Proc. of the 39th IEEEE Conference on Decision and Control, *Proc. CDC* **39**, 1273-1279.

[BlCrMaRa1998] Bloch, A. M., Crouch, P. E.,Marsden, J. E. and Ratiu, T. S. [1998] Discrete rigid body dynamics and optimal control. *Proc. CDC IEEE*, **37**, 2249-2254.

[Bl1965] Bloembergen, N. [1965] *Nonlinear optics*, Benjamin. New York.

[BoWo1965] Born, M. and Wolf, E. [1965] *Principles of Optics* (Oxford: Pergamon Press) p. 31.

[Bo71] Bourbaki, N. [1971] *Variétés différentielles et analytiques. Fascicule de résultats*. Hermann.

[Bo1989] Bourbaki, N. [1989] *Lie Groups and Lie Algebras. Chapters 1–3.* Springer Verlag.

[Br1964] Bretherton, F. P. [1964] Resonant interactions between waves: The case of discrete oscil lations, *J. Fluid Mech.*, **20**, 457-479.

[Br1895] Bruns, H. [1895] *Leipzig. Math. Phys. Ber.* **21**, 325.

[Bu 1993] Buchdahl, H. A. [1993] *An Introduction to Hamiltonian Optics*. New York: Dover.

[Ca1995] Calogero, F. [1995] An integrable hamiltonian system. *Phys. Lett. A* **201**, 306-310.

[CaFr1996] Calogero, F. and Francoise, J.-P. [1996] An integrable hamiltonian system. *J. Math. Phys.* **37**, 2863-2871.

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

[CeMaPeRa2003] Cendra, H., Marsden, J. E., Pekarsky, S., and Ratiu, T. S. [2003] Variational principles for Lie-Poisson and Hamilton-Poincaré equations, *Moskow Math. Journ.* **3**(3), 833–867.

[ChChLa1999] Chern, S. S., Chen, W. H. and Lam, K. S. [1999] *Lectures on Differential Geometry*. Series on University Mathematics Vol. 1, World Scientific.

[ChMa1974] Chernoff, P. R. and Marsden, J. E. [1974] *Properties of Infinite Dimensional Hamiltonian Systems*, Lecture Notes in Mathematics, **425**, Springer-Verlag.

[CuBa1997] Cushman, R. H. and Bates, L. M. [1997] *Global Aspects of Integrable Systems*. Birkhäuser.

[Cu-etal-2004] Cushman, R. H., Dullin, H. R., Giacobbe, A. Holm, D. D., Joyeux, M., Lynch, P., Sadovskií, D. A. and Zhilinskií, B. I. [2004] CO_2 Molecule as a quantum realization of the 1:1:2 resonant swing-spring with monodromy. *Phys. Rev. Lett.* **93**, 024302.

[DaBaBi1992] Dandoloff, R., Balakrishnan, R. and Bishop, A. R. [1992] Two-level systems: space curve formalism, Berry's phase and Gauss-Bonnet theorem. *J. Phys. A: Math. Gen.* **25**, L1105-L1110.

[DaHo1992] David, D. and Holm, D. D. [1992] Multiple Lie-Poisson structures, reductions, and geometric phases for the Maxwell-Bloch traveling-wave equations. *J. Nonlin. Sci.* **2**, 241–262.

[DaHoTr1989] David, D., Holm, D. D. and Tratnik, M. V. [1989] Horseshoe chaos in a periodically perturbed polarized optical beam. *Phys. Lett. A* **138**, 29-36.

[DaHoTr1990] David, D., Holm, D. D. and Tratnik, M. V. [1990] Hamiltonian chaos in nonlinear polarization dynamics. *Phys. Rep.* **187**, 281-367.

[DeMe1993] Dellnitz, M. and Melbourne, I. [1993] The equivariant Darboux theorem. *Lectures in Appl. Math.* **29**, 163-169.

[De1978] Devaney, R. L. [1978] Transversal homoclinic orbits in an integrable system. *Am. J. Math.* **100**, 631-642.

[Dirac1933] Dirac, P. A. M. [1933] The Lagrangian in quantum mechanics. *Physikalische Zeitschrift der Sowjetunion* **3**, 64-72.

[Dirac1981] Dirac, P. A. M. [1981] *The Principles of Quantum Mechanics*. 4th Revised edition, Oxford University Press, UK.

[Dr2007] Dragt, A. J. [2007] *Lie Methods for Nonlinear Dynamics with Applications to Accelerator Physics*. In preparation.

[DrJo1989] Drazin, P. G. and Johnson, R. S. [1989] *Solitons: An Introduction*. Cambridge University Press.

[DuNoFo1995] Dubrovin, B., Novikov, S. P. and Fomenko, A. T. [1995] *Modern Geometry I, II, III*, Volumes 93, 104, 124 of *Graduate Texts in Mathematics*. Springer-Verlag.

[DuGiCu2004] Dullin, H., Giacobbe, A. and Cushman, R. H. [2004] Monodromy in the resonant swing spring. *Physica D* **190**, 15.

[DuGoHo2001] Dullin, H., Gottwald, G. and Holm, D. D. [2001] An integrable shallow water equation with linear and nonlinear dispersion. *Phys. Rev. Lett.* **87**, 194501-04.

[DuGoHo2003] Dullin, H., Gottwald, G. and Holm, D. D. [2003] Camassa-Holm, Korteweg-de Vries-5 and other asymptotically equivalent equations for shallow water waves. *Fluid Dyn. Res.* **33**, 73-95.

[DuGoHo2004] Dullin, H., Gottwald, G. and Holm, D. D. [2004] On asymptotically equivalent shallow water wave equations. *Physica D* **190**, 1-14.

[EbMa1970] Ebin, D. G. and Marsden, J. E. [1970] Groups of diffeomorphisms and the motion of an incompressible fluid, *Ann. of Math.* **92**, 102-163.

[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 University Press.

[El2006] Ellis, D. C. P. [2006] The Hopf fibration and the $1-1$ resonant optical travelling wave pulse. MSci Project, Imperial College London.

[Er1942] Ertel, H. [1942] Ein neuer hydrodynamischer Wirbelsatz. *Met. Z., Braunschweig*, pp. 277-281.

[Fe1948] Feynman, R. P. [1948] Space-time approach to non-relativistic quantum mechanics. *Rev. Mod. Phys.* **20**, 367.

[FeHi1965] Feynman, R. P. and Hibbs, A. R. [1965] *Quantum Mechanics and Path Integrals*, McGraw-Hill.

[FeVeHe1957] Feynman, R., Vernon, F. and Hellwarth, R. [1957] Geometrical representation of the Schrödinger equation for solving maser problems. *Appl. Phys.* **28**, 49-52.

[Fl1963] Flanders, H. [1963] *Differential Forms with Applications to the Physical Sciences*, Academic Press, New York.

[FoHo1991] Fordy, A. P. and Holm, D. D. [1991] A tri-Hamiltonian formulation of the self-induced transparency equations. *Phys. Lett A* **160**, 143-148.

[FrHo2001] Fringer, O. B. and Holm, D. D. [2001] Integrable vs non-integrable geodesic soliton behavior. *Physica D* **150**, 237-263.

[Ga1632] Galilei, G. [1632] *Dialogue Concerning the Two Chief World Systems: The Ptolemaic and the Copernican*. `http://en.wikipedia.org/wiki/Dialogue_Concerning_the_Two_Chief_World_Systems`

[Ge1961] Gelfand, I. M. [1961] *Lectures on Linear Algebra*. Interscience, New York.

[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.

[GjHo1996] Gjaja, I. and Holm, D. D. [1996] Self-consistent wave-mean flow interaction dynamics and its Hamiltonian formulation for a rotating stratified incompressible fluid. *Physica D* **98**, 343-378.

[GuHo1983] Guckenheimer, J. and Holmes, P. [1983] *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer, Berlin.

[GuSt1984] Guillemin, V. and Sternberg, S. [1984] *Symplectic Techniques in Physics*. Cambridge University Press.

[HaMe1998] Haller, G. and Mezic, I. [1998] Reduction of three-dimensional, volume-preserving flows with symmetry. *Nonlinearity* **11**, 319-339.

[Ha1830] Hamilton, W. R. [1830] Supplement to an essay on the theory of systems of rays. *Transactions of the Royal Irish Academy*, **16**, part 1 pp. 1-61.

[Ha1837] Hamilton, W. R. [1837] Third supplement to an essay on the theory of systems of rays. *Transactions of the Royal Irish Academy*, **17**, part 1, pp. 1-144.

[Ha2006] Hanson, A. J. [2006] *Visualizing Quaternions*. MK-Elsevier.

[HaMi1977] Hasegawa, A. and Mima, K. [1997] Pseudo-three-dimensional turbulence in magnetized nonuniform plasmas. *Phys. Fluids* **21**, 87-92.

[He1858] Helmholtz, H., [1858] On integrals of the hydrodynamical equations which express vortex motion. Transl. Tait, P. G. [1867] *Phil. Mag.* **33**, no. 4, 485-512.

[Ho2002a] 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. Springer, pp. 113-167.

[Ho2002b] Holm, D. D. [2002] Variational principles for Lagrangian averaged fluid dynamics. *J. Phys. A: Math. Gen.* **35**, 1-10.

[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 University Press (2005).

[HoKo1991] Holm, D. D. and Kovacic, G. [1991] Homoclinic chaos for ray optics in a fiber. *Physica D* **51**, 177-188.

[HoKo1992] Holm, D. D. and Kovacic, G. [1992] Homoclinic chaos in a laser-matter system. *Physica D* **56**, 270-300.

[HoKoSu1991] Holm, D. D., Kovacic, G. and Sundaram, B. [1991] Chaotic laser-matter interaction. *Phys. Lett. A* **154**, 346–352.

[HoKoWe1995] Holm, D. D., Kovacic, G. and Wettergren, T. A. [1995] Near integrability and chaos in a resonant-cavity laser model, *Phys. Lett. A* **200**, 299-307.
Ibid [1996] Homoclinic orbits in the Maxwell-Bloch equations with a probe, *Phys. Rev. E* **54**, 243-256.

[HoLy2002] Holm, D. D. and Lynch, P. [2002], Stepwise precession of the resonant swinging spring. *SIAM J. Applied Dyn. Syst.* **1**, (1) 44-64.

[HoMa1991] Holm, D. D. and Marsden, J. E. [1991] The rotor and the pendulum. in *Symplectic Geometry and Mathematical Physics*, P. Donato, C. Duval, J. Elhadad, G. M. Tuynman, ed., *Prog. in Math.* Vol. **99**, Birkhauser: Boston, pp. 189-203.

[HoMa2004] Holm, D. D. and Marsden, J. E. [2004] Momentum maps and measure valued solutions (peakons, filaments, and sheets) of the Euler-Poincarńe equations for the diffeomorphism group. *In The Breadth of Symplectic and Poisson Geometry, A Festshrift for Alan Weinstein*, pp. 203-235, *Progr. Math.*, **232**, J.E. Marsden and T.S. Ratiu, Editors, Birkhäuser Boston, Boston, MA. 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.

[HoMaRa1998b] Holm, D. D. Marsden, J. E. and Ratiu, T. S. [1998b] Euler-Poincaré models of ideal fluids with nonlinear dispersion. *Phys. Rev. Lett.* **80**, 4173-4177.

[HoMaRa2002] Holm, D. D. Marsden, J. E. and Ratiu, T. S. [2002] The Euler-Poincaré equations in geophysical fluid dynamics. In *Large-Scale Atmosphere-Ocean Dynamics 2: Geometric Methods and Models*. Edited by J. Norbury and I. Roulstone, Cambridge University Press: Cambridge, pp. 251-299.

[HoRaTrYo2004] Holm, D. D., Rananather, J. T., Trouvé, A. and Younes, L. [2004] Soliton dynamics in computational anatomy. *NeuroImage* **23**, S170-178 (2004). http://arxiv.org/abs/nlin.SI/0411014

[HoMaRaWe1985] Holm, D. D., Marsden, J. E., Ratiu, T. S. and Weinstein, A. [1985] Nonlinear stability of fluid and plasma equilibria. *Physics Reports* **123** 1-116.

[HoSt2002] Holm, D. D. and Staley, M. F. [2002] Wave Structures and Nonlinear Balances in a Family of 1+1 Evolutionary PDEs. http://arxiv.org/abs/nlin.CD/0202059.

[HoWo1991] Holm, D. D. and Wolf, K. B. [1991] Lie-Poisson description of Hamiltonian ray optics. *Physica D* **51**, 189-199.

[Ho1931] Hopf, H. [1931] Über die Abbildungen der dreidimensionalen Sphäre auf die Kugelfläche. *Mathematische Annalen* **104**, 637-665.

[HoHa1994] Horton, W. and Hasegawa, A. [1994] Quasi-two-dimensional dynamics of plasmas and fluids. *Chaos* **4**, 227-251.

[HuZh1994] Hunter, J. K. and Zheng, Y. [1994] On a completely integrable nonlinear hyperbolic variational equation. *Physica D* **79**, 361-386.

[Is1999] Isham, C. J. [1999] *Differential Geometry for Physicists*. World Scientific Lecture Notes in Physics **61**.

[JoSa98] José, J. V. and Saletan, E. J. [1998] *Classical Dynamics : A Contemporary Approach*. Cambridge University Press.

[Jo1998] Jost, J. [1998] *Riemannian Geometry and Geometric Analysis. University Text*. Springer-Verlag, second edition.

[KaKoSt1978] Kazhdan, D., Kostant, B., and Sternberg, S. [1978] Hamiltonian group actions and dynamical systems of Calogero type, *Comm. Pure Appl. Math.* **31**, 481–508.

[Ke1962] Keller, J. B. [1962] Geometrical theory of diffraction, *J. Optical Soc. America* **52**, 116–130.

[Ke1869] Kelvin, L., [1869] On vortex motion. *Trans. Roy. Soc. Edin.* **25**, 217-260; Maximum and minimum energy in vortex motion. *Mathematical and Physical Papers*, vol. 4, Cambridge Univ. Press, 172-183.

[KhMis03] Khesin, B. and Misiolek, G. [2003] Euler equations on homogeneous spaces and Virasoro orbits, *Adv. in Math.* **176**, 116-144.

[Ki1979] Kibble, T. [1979] Geometrization of quantum mechanics, *Comm. Math. Phys.* **65**, 189-201.

[KnHj2001] Knudsen, J. M. and Hjorth, P. G. [2001] *Elements of Newtonian Mechanics: Including Nonlinear Dynamics*. Springer-Verlag, New York, third edition, second printing.

[Ko1984] Koiller, J. [1984] A mechanical system with a "wild" horseshoe. *J. Math. Phys.* **25**, 1599-1604.

[Ko1966] Kostant, B. [1966] Orbits, symplectic structures and representation theory. *Proc. US-Japan Seminar on Diff. Geom., Kyoto. Nippon Hyronsha, Tokyo* **77**.

[KrSc2001] Kruse, H. P., Schreule, J. and Du, W. [2001] A two-dimensional version of the CH equation. In *Symmetry and Perturbation Theory: SPT 2001* Edited by D. Bambusi, G. Gaeta and M. Cadoni. World Scientific: New York, pp 120-127.

[Ku1999] Kuipers, J. B. [1999] *Quaternions and Rotation Sequences: a Primer with Applications to Orbits, Aerospace and Virtual Reality*. Princeton University Press.

[Ku1976] Kummer, M. [1976] On resonant nonlinearly coupled oscillators with two equal frequencies. *Commun. Math. Phys.* **48**, 53-79 (1976).

[Ku1978] Kummer, M. [1978] On resonant classical Hamiltonians with two equal frequencies. *Commun. Math. Phys.* **58**, 85-112.

[Ku1981] Kummer, M. [1981] On the construction of the reduced phase space of a Hamiltonian system with symmetry. *Indiana Univ. Math. J.* **30**, 281-291.

[Ku1986] Kummer, M. [1986] in *Local and Global Methods in Nonlinear Dynamics*, Lecture Notes in Physics Vol. **252**, edited by A. V. Sáenz, Springer-Verlag, New York, pp. 19-31.

[La1999] Lang, S. [1999] *Fundamentals of Differential Geometry*. Volume 191 of *Graduate Texts in Mathematics*. Springer-Verlag, New York.

[Le2003] Lee, J. [2003] *Introduction to Smooth Manifolds*, Springer-Verlag.

[LiMa1987] Libermann, P. and Marle, C.-M. [1987] *Symplectic Geometry and Analytical Mechanics*. Reidel.

[Lie1890] Lie, S. [1890] *Theorie der Transformationsgruppen. Zweiter Abschnitt*. Teubner.

[LoGi1967] Longuet-Higgins, M. S. and Gill, A. E. [1967] Resonant interactions between planetary waves. *Proc. Roy. Soc. Edinburgh Sect. A.* **299**, 120-140.

[Lo1963] Lorenz, E. N. [1963] Deterministic non-periodic flow. *J. Atmospheric Sci.* **20**, 130-141.

[LuAlMaRo2000] Luther, G. G., Alber, M. S., Marsden, J. E. and Robbins, J. M. [2000] Geometric analysis of optical frequency conversion and its control in quadratic nonlinear media, *J. Opt. Soc. Amer. B Opt. Phys.* **17**, 932-941.

[Ly2002a] Lynch, P. [2002] Resonant motions of the three-dimensional elastic pendulum, *Internat. J. Non-Linear Mech.* **37**, 345-367.

[Ly2002b] Lynch, P. [2002] The swinging spring: A simple model for atmospheric balance, in *Large-Scale Atmosphere-Ocean Dynamics: Vol II: Geometric Methods and Models*, Cambridge University Press, Cambridge, UK, pp. 64-108.

[Mac1970] Mac Lane, S. [1970] Hamiltonian mechanics and geometry, *American Mathematical Monthly*, **77**, 570-586.

[Man1976] Manakov, S. V. [1976] Note on the integration of Euler's equations of the dynamics of and n-dimensional rigid body. *Funct. Anal. and its Appl.* **10**, 328–329.

[Mar1976] Marle, C.-M. [1976] Symplectic manifolds, dynamical groups, and Hamiltonian mechanics, in *Differential Geometry and Relativity*, Cahen, M. and Flato, M. eds., D. Reidel, Boston, 249–269.

[Ma1981] Marsden, J. E. [1981] *Lectures on Geometric Methods in Mathematical Physics*. Volume 37, SIAM, Philadelphia.

[Ma1992] Marsden, J. E. [1992] *Lectures on Mechanics*. Volume 174 of *London Mathematical Society Lecture Note Series*. Cambridge University Press.

[MaHu1983] Marsden, J. E. and Hughes, T. J. R. [1983] *Mathematical Foundations of Elasticity*. Prentice Hall. Reprinted by Dover Publications, NY, 1994.

[MaMiOrPeRa2007] Marsden, J. E., Misiolek, G., Ortega, J.-P., Perlmutter, M., and Ratiu, T. S. [2007] *Hamiltonian Reduction by Stages*. Lecture Notes in Mathematics, **1913**. Springer-Verlag.

[MaMoRa1991] Marsden, J. E., Montgomery, R., and Ratiu, T. S. [1991] Reduction, symmetry, and phases in mechanics. *Memoirs Amer. Math. Soc.* **88**, (436) 1–110.

[MaRa1994] Marsden, J. E. and Ratiu, T. S. [1994] *Introduction to Mechanics and Symmetry*. Volume 75 of *Texts in Applied Mathematics*, second printing of second edition 2003. Springer-Verlag.

[MaRa95] Marsden, J. E. and Ratiu, T. S. [2003] *Geometric Fluid Dynamics*. Unpublished notes.

[MaRa03] Marsden, J. E. and Ratiu, T. S. [2003] *Mechanics and Symmetry. Reduction Theory*. In preparation.

[MaRaWe84a] Marsden, J. E., Ratiu, T. S., and Weinstein, A. [1984a] Semidirect products and reduction in mechanics. *Trans. Amer. Math. Soc.* **281**, (1) 147–177.

[MaRaWe84b] Marsden, J. E., Ratiu, T. S., and Weinstein, A. [1984b] Reduction and Hamiltonian structures on duals of semidirect product Lie algebras. *Contemporary Math.*, **28**, 55–100.

[MaWe74] Marsden, J. E. and Weinstein, A. [1974] Reduction of symplectic manifolds with symmetry. *Rep. Math. Phys.* **5**, 121–130.

[MaWe83] Marsden, J. E. and Weinstein, A. [1983] Coadjoint orbits, vortices and Clebsch variables for incompressible fluids, *Physica D*, **7**, 305–323.

[MaWil989] Mawhin, J. and Willem, M. [1989] *Critical Point Theory and Hamiltonin Systems*. Volume 74 of *Applied Mathematical Sciences*. Springer-Verlag, second edition.

[McSa1995] McDuff, D. and Salamon, D. [1995] *Introduction to Symplectic Topology*. Clarendon Press.

[MeDe1993] Melbourne, I. and Dellnitz, M. [1993] Normal forms for linear Hamiltonian vector fields commuting with the action of a compact Lie group. *Proc. Camb. Phil. Soc.* **114**, 235–268.

[Mi1963] Milnor, J. [1963] *Morse Theory*. Princeton University Press.

[MiFo1978] Mishchenko, A. S. and Fomenko, A. T. [1978] Euler equations on finite dimensional Lie groups. *Izv. Acad. Nauk SSSR*, Ser. Matem. **42**, (2) 396-415 (Russian); English translation: *Math. USSR-Izv.* **12**, (2) 371-389.

[Mi2002] Misiolek, G. [2002] Classical solutions of the periodic Camassa-Holm equation. *Geom. Funct. Anal.*, **12**, 1080–1104.

[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.

[Mo1985] Moffatt, H. K. [1985] Magnetostatic equilibria and analogous Euler flows of arbitrarily complex topology, Part 1. *J. Fluid Mech.* **159**, 359-378; Part 2 [1986] *J. Fluid Mech.* **166**, 359-378.

[MoTs1992] Moffatt, H. K. and Tsinober, A. [1992] Helicity in laminar and turbulent flow. *Ann. Rev. Fluid Mech.* **24**, 281-312.

[Mo1991] Montgomery, R. [1991] How much does the rigid body rotate? A Berry's phase from the 18th Century, *Am. J. Phys.* **59**, 394-398.

[Na1973] Nambu, Y. [1973] Generalized Hamiltonian mechanics, *Phys. Rev. D* **7**, 2405-2412.

[Ne1859] Neumann, C. [1859] De problemate quodam mechanica, quod ad primam integralium ultra-ellipticorum classem revocatur, *J. Reine Angew. Math.* bf 56, 54-66.

[Ne1686] Newton, I. [1686] *Principia Mathematica Philosophiae Naturalis*. Reprinted by University of California Press, Berkeley, California, 1934.

[No1918] Noether, E. [1918] *Nachrichten Gesell. Wissenschaft. Göttingen* **2**, 235. See also C. H. Kimberling [1972] *Am. Math. Monthly* **79**, 136.

[OcoRo1998] O'Connor, J. J. and Robertson, E. F., [1998] *Sir William Rowan Hamilton*, `http://www-groups.dcs.st-and.ac.uk/~history/Mathematicians/Hamilton.html`

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

[OrRa2004] Ortega, J.-P. and Ratiu, T. S. [2004] *Momentum Maps and Hamiltonian Reduction*. Volume 222 of *Progress in Mathematics*. Birkhäuser.

[Ot1993] Ott, E. [1993] *Chaos in Dynamical Systems*, Cambridge University Press, Cambridge, UK, p. 385.

[OKh87] Ovsienko, V. Y. and Khesin, B. A. [1987] Korteweg-de Vries superequations as an Euler equation. *Funct. Anal. Appl.*, **21**, 329-331.

[Pa1968] Palais, R. [1968] *Foundations of Global Non-Linear Analysis*, W. A. Benjamin, Inc., New York-Amsterdam.

[Po1892] Poincaré H. [1892] *Theorie Mathematique de la Lumiere*, (ed.) George Carre, Paris, p. 275.

[Po1901] Poincaré, H. [1901] Sur une forme nouvelle des équations de la méchanique *C.R. Acad. Sci.* **132**, 369-371.

[Ra1980] Ratiu, T. [1980] The C. Neumann problem as a completely integrable system on an adjoint orbit. *Trans. Amer. Math. Soc.* **264**, 321-329.

[Ra1981] Ratiu, T. [1981] The motion of the free n-dimensional rigid body. *Indiana U. Math. J.* **29**, 609-627.

[RaTuSbSoTe2005] Ratiu, T. S., Tudoran, R., Sbano, L., Sousa Dias, E., 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 University Press.

[Ru1959] Rund, H. [1959] *The Differential Geometry of Finsler Spaces*, Springer-Verlag.

[SaCa1981] Sarlet, W. and Cantrijn, F. [1981] *Generalizations of Noether's theorem in classical mechanics*. *SIAM Review* **23**, 467-494.

[Sc1987] Schmid, R. [1987] *Infinite Dimensional Hamiltonian Systems*. Bibliopolis.

[Se1992] Serre, J.-P. [1992] *Lie Algebras and Lie Groups*, Volume 1500 of *Lecture Notes in Mathematics*. Springer-Verlag.

[ShWi1989] Shapere, A. and Wilcek, F. [1989] *Geometric Phases in Physics*, World Scientific: Singapore.

[Sh1984] Shen, Y. R. [1984] *The Principles of Nonlinear Optics*, Wiley-lnterscience, New York.

[Sh1998] Shkoller, S. [1998] Geometry and curvature of diffeomorphism groups with H^1 metric and mean hydrodynamics *J. Funct. Anal.* **160**, 337-365.

[Sh2000] Shkoller, S. [2000] Analysis on groups of diffeomorphisms of manifolds with boundary and the averaged motion of a fluid *J. Diff. Geom.* **55**, 145-191.

[Si1913] Silberstein, L. [1913] *Vectorial Mechanics*, MacMillan and Co., Limited. London. 2nd ed. (1926).

[Sm1970] Smale, S. [1970] Topology and mechanics, *Inv. Math.* **10**, 305-331; **11**, 45-64.

[So1970] Souriau, J. M. [1970] *Structure des Systèmes Dynamiques*, Dunod, Paris.

[Sp1982] Sparrow, C. [1982] *The Lorenz Equations: Bifurcations, Chaos and Strange Attractors*, Springer-Verlag, New York, p. 269.

[Sp1979] Spivak, M. [1979] *Differential Geometry*, Volume I. New printing with corrections. Publish or Perish, Inc. Houston, Texas.

[St1852] Stokes, G. G. [1852] *Trans. Camb. Phil. Soc.* **9**, 399.

[Sy1937] Synge, J. L. [1937] *Geometrical Optics: An Introduction to Hamilton's Method*, Cambridge U. Press, London.
Ibid [1937] Hamilton's method in geometrical optics, *J. Opt. Soc. Am.* **27**, 75-82. `http://www.opticsinfobase.org/abstract.cfm?URI=josa-27-2-75`
Ibid [1960] Classical Dynamics, in *Handbuch der Physik*, edited by S. Flügge Springer-Verlag, Berlin, Vol. III/1, p. 11.

[Ta1994] Takhtajan, L. [1994] On foundation of the generalized Nambu mechanics. *Commun. Math. Phys.* **160**, 295-315.

[Ur2003] Urbantke, H. K. [2003] The Hopf fibration - seven times in physics. *J. Geom. Phys.* **46** 125-150.

[Va1996] Vaisman, I. [1996] *Lectures on the Geometry of Poisson Manifolds*. Volume 118 of *Progress in Mathematics*, Birkhäuser.

[ViGo1933] Vitt, A. and Gorelik, G. [1933] Kolebaniya uprugogo mayatnika kak primer kolebaniy dvukh parametricheski svyazannykh linejnykh sistem, *Zh. Tekh. Fiz.* (*J. Tech. Phys.*), **3**, pp. 294Ð307. Available in English translation: Oscillations of an Elastic Pendulum as an Example of the Oscillations of Two Parametrically Coupled Linear Systems, Translated by Lisa Shields with an Introduction by Peter Lynch, Historical Note 3, Met Eireann, Dublin, 1999.

[Wa1983] Warner, F. W. [1983] *Foundation of Differentiable Manifolds and Lie Groups*. Volume 94 of *Graduate Texts in Mathematics*, Springer-Verlag.

[We1983] Weinstein, A. [1983a] Sophus Lie and symplectic geometry, *Exposition Math.* **1**, 95-96.

[We1983b] Weinstein, A. [1983b] The local structure of Poisson manifolds. *Journ. Diff. Geom* **18**, 523–557.

[We2002] Weinstein, A. [2002] Geometry of momentum (preprint); ArXiv:math/SG0208108 v1.

[WeFiOt1980] Wersinger, J.-M., Finn, J. M. and Ott, E. [1980] Bifurcations and strange behavior in instability saturation by nonlinear mode coupling. *Phys. Rev. Lett.* **44**, 453-456.

[Wh1974] Whitham, G. B. [1974] *Linear and Nonlinear Waves*, John Wiley and Sons, New York.

[Wi1988] Wiggins, S. [1988] *Global Bifurcations and Chaos – Analytical Methods*, Springer. Berlin.

[Wo2004] Wolf, K. B. [2004] *Geometric Optics on Phase Space*, Springer. ISBN 3540220399.

[Wo1958] Woltjer, L. [1958] A theorem on force-free magnetic fields. *Proc. Nat. Acad. Sci. USA* **44**, 489-491.

Index