Thesis for the degree of Doctor of Philosophy

Geometric Numerical Methods: From Random Fields to Shape Matching

Erik Jansson

CHALMERS

Department of Mathematical Sciences
Division of Applied Mathematics and Statistics
Chalmers University of Technology
Göteborg, Sweden 2025

Geometric Numerical Methods: From Random Fields to Shape Matching
Erik Jansson

ISBN 978-91-8103-208-6
Acknowledgements, dedications, and similar personal statements in this thesis, reflect the author's own views.

Doktorsavhandlingar vid Chalmers tekniska högskola Ny serie nr 5666 ISSN 0346-718X

Department of Mathematical Sciences
Division of Applied Mathematics and Statistics
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone: +46 (0)31-772 1000

Cover: The evolution of the spherical Zeitlin–Euler equations (see Section 4.5) with initial data generated by a random field sampled using the Chebyshev–Galerkin method described in Section 3.4.

Typeset with LaTeX
Printed by Chalmers Digitaltryck
Göteborg, Sweden 2025

Geometric Numerical Methods: From Random Fields to Shape Matching

Erik Jansson
Department of Mathematical Sciences
Division of Applied Mathematics and Statistics
Chalmers University of Technology

Abstract

Geometry is central to many applied problems, though its influence varies. Some problems are inherently geometric, requiring numerical methods that preserve the underlying structure to remain accurate. Others are well understood in Euclidean space but demand different techniques when extended to curved settings. This thesis addresses such geometric challenges through studying numerical methods for two main types of problems: matching problems and stochastic (partial) differential equations. It is based on seven papers—the first three focus on SPDEs and SDEs, while the remaining consider matching problems and related differential equations. The first develops a numerical method for fractional SPDEs on the sphere, combining a recursive splitting scheme with surface finite elements. The second studies a Chebyshev–Galerkin approach for simulating non-stationary Gaussian random fields on hypersurfaces. The third introduces a geometric integrator for stochastic Lie–Poisson systems, derived via a reduction of the implicit midpoint method for canonical Hamiltonian systems. The fourth explores sub-Riemannian shape matching, where shapes are matched using constrained motions, and shows how this problem can be interpreted as a neural network. The fifth studies the convergence of a gradient flow for the Gaussian Monge problem. The sixth adapts geometric shape matching to recover protein conformations from single-particle Cryo-EM data by using rigid deformations of chains of particles. The seventh investigates the numerical signature of blow-up in hydrodynamic equations, showing that numerical solutions can be used to detect the onset in a class of hydrodynamic equations.

Keywords: Stochastic partial differential equations, shape analysis, optimal transport, Gaussian random fields, Lie–Poisson systems, hydrodynamics, surface finite element methods, geometric numerical integration.

Contents

List of Publications

The following papers and preprints are included in this thesis:

Paper I: Erik Jansson, Annika Lang, Mihály Kovács, Surface Finite Element Approximation of Spherical Whittle–Matérn Gaussian Random Fields, *SIAM Journal on Scientific Computing* 44 (2), 2021, p. A825–A842.

Paper II: Erik Jansson, Annika Lang, Mike Pereira, Non-stationary Gaussian random fields on hypersurfaces: Sampling and strong error analysis, submitted, preprint available at *arXiv:2406.08185*, 2024.

Paper III: Sagy Ephrati, Erik Jansson, Annika Lang, Erwin Luesink, An exponential map free implicit midpoint method for stochastic Lie–Poisson systems, submitted, preprint available at *arXiv:2408.16701*, 2024.

Paper IV: Erik Jansson, Klas Modin, Sub-Riemannian Landmark Matching and its interpretation as residual neural networks, accepted for publication in *Journal of Computational Dynamics*.

Paper V: Erik Jansson, Klas Modin, Convergence of the vertical gradient flow for the Gaussian Monge problem, *Journal of Computational Dynamics*, 11 (1), 2024, p. 1-9.

Paper VI: Erik Jansson, Jonathan Krook, Klas Modin, Ozan Öktem, Geometric shape matching for recovering protein conformations from single-particle Cryo-EM data, submitted, preprint available at *arXiv:2410.00833*, 2024

Paper VII: Erik Jansson, Klas Modin, On the numerical signature of blow-up in hydrodynamic equations, submitted, preprint available at *arXiv:2210.02328*, 2024

Author contribution:

Paper I: E.J. derived the splitting scheme, performed the error analysis, implemented the numerical experiments, and wrote the first version of the paper, which was completed together with the coauthors.

Paper II: E.J. contributed significantly to early approaches to the problem, contributed to the final theoretical approach and wrote, in collaboration with the coauthors, the submitted version of the manuscript.

Paper III: E.J. derived the method in collaboration with the coauthors, performed the error analysis, devised, in collaboration with the coauthors, the numerical experiments, and contributed to the submitted version of the manuscript in collaboration with the coauthors.

Paper IV: E.J. developed most of the theoretical framework and drafted the manuscript together with the coauthor. E.J. also implemented the numerical experiments presented in the paper.

Paper V: E.J. carried out the convergence analysis of the gradient flow and implemented the numerical experiment. E.J. wrote the first version of the paper, which was completed together with the coauthor.

Paper VI: E.J. developed the shape-based framework for the problem, together with the coauthors, implemented the final version of the numerical experiments and completed the manuscript together with the coauthors.

Paper VII: E.J. derived the equation and its discretization, designed and performed the numerical experiments, and wrote the first version of the paper, which was completed together with the coauthor.

Acknowledgements

First, I want to express my sincere gratitude to my supervisors, Klas Modin and Annika Lang, for their support and guidance throughout these past years. I'm also thankful to my examiner, Ann-Brith Strömberg, for her valuable feedback on the thesis and for helping me prepare for the defense. I've also been fortunate to have a great head of unit in Irina Petterson——thank you.

A big thank you to my coauthors and collaborators: Mihály Kovács, Mike Pereira, Erwin Luesink, Sagy Ephrati, Ozan Öktem, Jonathan Krook, Ruben Seyer, Moritz Schauer, Akash Sharma, Björn Müller, Andrea Papini, and Magnus Röding. It's been a real pleasure working with all of you, and I've learned a great deal through our collaborations.

To my colleagues at the department—thank you for fostering such a stimulating and friendly environment. Philipp, thanks for being a good sport about sharing an office with me, and for all the valuable discussions. And Ruben, a special thanks for making the thesis better by getting the example boxes to work.

I'm grateful to Kjell-Ove Widman and Stiftelsen för Vetenskaplig Forskning och Utbildning i Matematik, the Royal Society of Arts and Sciences in Gothenburg, Chalmersska forskningsfonden, and the Knut och Alice Wallenberg Foundation for supporting several of my conference travels. The funders of this project, the Wallenberg AI, Autonomous Systems and Software Program (WASP) are gratefully acknowledged for financial support.

Finally, I'd like to thank my family and friends for their support and encouragement. Esmée—thank you for everything.

Erik Jansson
Göteborg, 2025

Chapter 1: Introduction

In the early stages of most mathematics curricula, emphasis is placed on developing analytical problem-solving skills. For example, students spend significant time manually computing roots of polynomials, solving differential equations, and evaluating integrals. At this stage, it's easy to assume that every problem has a closed-form solution, and that solving it simply requires learning the right techniques. However, as the curriculum progresses, the problems grow increasingly complex, and the techniques more intricate, revealing that there are many problems for which we do not have a closed-form solution. In fact, as Stibitz and Larrivee (1957) observed, many of the problems encountered in applications are not amenable to the analytical techniques the typical student has learned. This is where numerical methods come into play.

Numerical methods have played a central role in the application of mathematics throughout history. Indeed, many famous mathematicians of the past—such as Newton, Euler, Gauss, and Legendre—used numerical approximation in their work (Goldstine, 1977). Yet there's a key distinction between merely using numerical methods and systematically studying them. The field that studies numerical methods from a mathematical point of view, as a branch of analysis, is called numerical analysis. Numerical analysts not only develop algorithms but also rigorously analyze their behavior—examining when they work, when they fail, and how they can be improved. The field's roots extend back to the pre-digital era, when mechanical devices and numerical tables were used to approximate solutions (Stibitz and Larrivee, 1957; Goldstine, 1977).

However, according to Grcar (2011), the inception of modern numerical analysis is closely linked to the development of the first digital computers. Indeed, first modern paper in numerical analysis is often cited as (von Neumann and Goldstine, 1947), the authors of which also played a significant role in the development of the first digital computers. In this paper, John von Neumann and Herman Goldstine rigorously study the sources of error in a numerical matrix inversion algorithm. The paper is important not only because it connects numerical analysis and its error analysis with digital computers, but also, as Grcar (2011) notes, because it served as von Neumann's

way of establishing numerical analysis as a serious branch of mathematics. Since then, computers and numerical analysis have remained closely linked. As computers have grown more powerful and widespread across scientific disciplines, numerical analysis has become an increasingly significant field.

Depending on whom you ask, one of the perks or drawbacks of numerical analysis is that it can be applied to virtually any problem, and as such, a numerical analyst can find themselves working on problems from a wide range of fields. This thesis is a contribution to the field of numerical analysis, and it is concerned with the development and analysis of numerical methods for problems in mechanics, hydrodynamics, protein imaging, and stochastic differential equations. These applications all have in common that they are problems where numerical analysis is interacting with one of the oldest branches of mathematics: geometry. Geometry is the study of space and its properties, such as shape, distance and area, and it is a field that has been studied for thousands of years. To combine it with numerical analysis is not new: a famous example is one of the first numerical methods for solving ordinary differential equations. It was invented by Newton in his proof of Kepler's second law (Wanner, 2010). This method is an early example of a method we call *geometric*, since it preserves the underlying mechanical properties of the true solution of the differential equation, properties that arise from the geometry of the problem. However, it should be noted that Newton used the numerical methods not to solve a known continuous problem, but to infer the properties of the continuous problem from the numerical method. In fact, as noted by Wanner (2010), this inverse use of numerical methods has not been uncommon, and several of the mathematicians mentioned above used numerical methods in this way.

In this thesis, geometry enters the picture in a few different ways. Sometimes, we are, like Newton, interested in methods that preserve the underlying geometry of a mechanical problem. In these cases, the geometric aspects may not be immediately obvious, and the motivation for the method becomes clear only when we understand the underlying geometric mechanics of the problem. In other cases, geometry plays a more concrete role: we are interested in problems posed on curved spaces, rather than flat domains. This type of problem could show up, for instance, if one is interested in modelling some sort of hydrodynamical system on a global scale, i.e., a system not confined to a flat subsection of the earth but rather one that evolves on the entire surface of the Earth. A suitable domain in this case is the sphere, even if, as Arnold remarks in (Arnold, 1989): a convenient simplifying assumption in hydrodynamics is that the Earth is modeled as a flat torus. If we disregard this, and accept that the Earth is not flat, we are faced with the problem of solving differential equations on a curved space. Thus, geometry immediately enters the picture, and the methods used to solve the problem must take this

into account.

In more detail, this thesis consists of some introductory chapters, and more importantly, of seven papers which can be divided into two categories of applications.

1. Matching problems and hydrodynamics.

2. Stochastic (partial) differential equations.

The first category encompasses a variety of problems, but they are unified by their connection to Lie–Poisson equations, a class of mechanical systems. An important goal of the introductory chapters is to provide an overview of geometric mechanics and differential geometry, which form the foundation for these problems. This background will also be helpful for the second category, where we study stochastic differential equations, both partial and ordinary. More specifically, we focus on the numerical solution of elliptic stochastic partial differential equations (SPDEs) on surfaces, as well as the numerical solution of stochastic Lie–Poisson equations. The latter equations are a type of stochastic (ordinary) differential equation that arises in mechanics, providing a bridge to the first category. Elliptic SPDEs are used to simulate *random fields* on surfaces, that is, mappings that assign a random value to each point on the surface.

The introductory chapters are structured as follows: We begin by introducing some geometric and mechanical basics in Chapter 2. The final goal of that chapter is the formulation of the Euler–Arnold equations, which are central to many of the systems studied in this thesis. In Chapter 3, we discuss the stochastic differential equations relevant to the second category of problems. Chapter 4 covers equations arising in matching problems and hydrodynamics. Finally, in Chapter 5, we summarize the papers included in the thesis. The introductory chapters are based on those of my licentiate thesis (Jansson, 2022).

Chapter 2: Mechanics and Differential Geometry

Classical mechanics and differential geometry are intimately connected. In this section, we introduce geometric concepts as they arise in mechanical examples. The focus here is not to provide a self-contained introduction, but rather to give a brief overview of some concepts and tools that appear in several of the papers included in the thesis. For a complete introduction to differential geometry, we refer the reader to (Lee, 2012), and for mechanics, to (Arnold, 1989; Marsden and Ratiu, 1999).

2.1 The Basics: Positions and Velocities

We start with a question.

What is mechanics?

The answer: Mechanics is the study of motions, or displacements, of objects in space. The displacements occur due to forces acting on the object.

Let us warm up with a few basic mechanical examples where displacements take place in certain subspaces of the Euclidean space.

Example 2.1: Pendulums

Consider a pendulum. It consists of a mass m attached to a rigid rod of length l. The mass is free to rotate in the vertical plane, i.e., it is constrained to $\mathbb{R}^2$, and we let the angle between the rod and the vertical axis be θ. The displacements are thus the possible positions of the mass as determined by the rod, described by the angle θ. See Figure 2.1a for an illustration. Mathematically, the possible positions are modelled by some space $Q \subset \mathbb{R}^2$. Assuming for simplicity that $l = 1$ and that the mass is free to rotate one full circle, Q is the circle $\mathbb{S}^1 = \{x \in \mathbb{R}^2 : \|x\|_{\mathbb{R}^2} = 1\}$.
A more complicated model is the double pendulum, with two masses

m_1 and m_2, attached to two connected rigid rods of length l_1 and l_2; see Figure 2.1b. In this case, we need two angles to describe the displacements, so the possible positions of the two masses are $Q = \mathbb{S}^1 \times \mathbb{S}^1$, i.e., a flat torus.
Returning to the single pendulum, and removing the *constraint* that the mass only rotates in the vertical plane, we get a spherical pendulum; see Figure 2.1c. Now, the possible positions are $Q = \mathbb{S}^2$, i.e., the sphere.

We have seen three examples of mechanical systems and how the constraints determine the possible positions, i.e., the space of possible configurations Q. These spaces are all examples of *smooth manifolds*. While the examples we have seen, $\mathbb{S}^1$, $\mathbb{S}^1 \times \mathbb{S}^1$ and $\mathbb{S}^2$, are easy to visualize and understand without much background, in general, mechanical systems can have much more complicated *configuration manifolds*. Intuitively, a manifold is a space that is locally Euclidean, i.e., it looks like $\mathbb{R}^n$ close to each point, but globally, it can be more complicated.

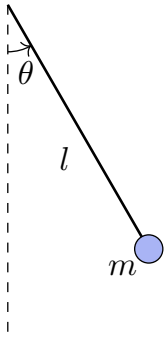
Formally, an n-dimensional *topological manifold* M is a completely separable Hausdorff topological space such that each point $x \in M$ has a neighborhood U homeomorphic to an open subset of $\mathbb{R}^n$. Denoting the corresponding homeomorphism by ϕ, a pair (U, ϕ) is known as a *chart*.

A point x is in (U, ϕ) if $x \in U$. A family of charts $(U_i, \phi_i)_{i=1}^m$ such that every $x \in M$ is in at least one charts is called an *atlas*. An atlas is smooth if, for every pair of charts (U_i, ϕ_i) and (U_j, ϕ_j), it holds that $\phi_i(U_i \cap U_j)$ and $\phi_j(U_i \cap U_j)$ are open sets in $\mathbb{R}^n$ and the *transition functions* $\phi_j \circ \phi_i^{-1} : \phi_i(U_i \cap U_j) \to \phi_j(U_i \cap U_j)$ are smooth. A smooth manifold is a topological manifold equipped with a smooth atlas that is maximal, meaning it cannot be contained in a larger smooth atlas. If there is an atlas for which all transition functions have positive Jacobian determinants, then the manifold is said to be *orientable*.

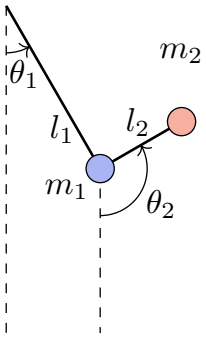
Charts are used to define *local coordinates*. The set of local coordinates for a point x in (U, ϕ) is $(x_1, x_2, \ldots, x_n) = \phi(x)$. Working with local coordinates is sometimes easier than working on the manifold itself.

Charts allow us to define what is meant by smooth mappings of manifolds. A mapping $F \colon M \to N$ between two smooth manifolds M and N is smooth if, for every $x \in M$, there is a chart (U, ϕ) around x and a chart (V, ψ) around $F(x)$ such that $F(U) \subset V$ and $\psi \circ F \circ \phi^{-1} \colon \phi(U) \to \psi(V)$ is smooth. A smooth map that is bijective and whose inverse is also smooth is called a *diffeomorphism*. Two manifolds are said to be *diffeomorphic* if there exists a diffeomorphism between them.

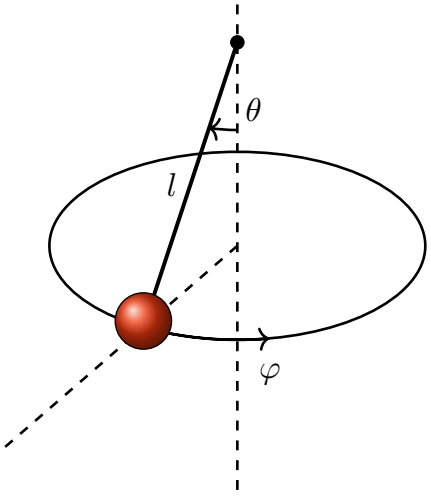
The concept of charts is illustrated in Figure 2.2. For a concrete example, let's again consider the sphere.



(a) Single pendulum. The position of the mass is determined by the angle θ, so the configuration space is $Q = \mathbb{S}^1$.



(b) Double pendulum. The position of the masses are determined by the angles θ_1 and θ_2, so the configuration space is $Q = \mathbb{S}^1 \times \mathbb{S}^1$, i.e., a torus.



(c) Spherical pendulum. The position of the mass is determined by the angles θ, between 0 and π radians, and φ between 0 and 2π radians, so the configuration space is $Q = \mathbb{S}^2$, i.e., the sphere.

Figure 2.1: Examples of pendulums.

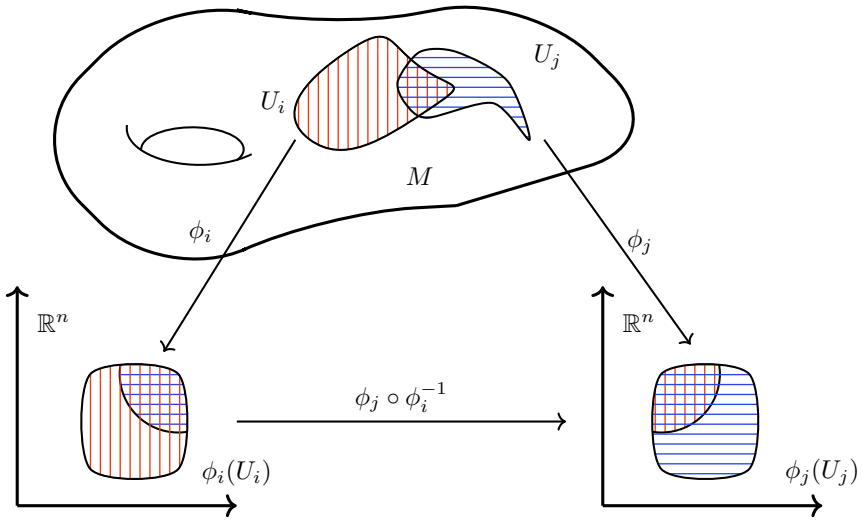


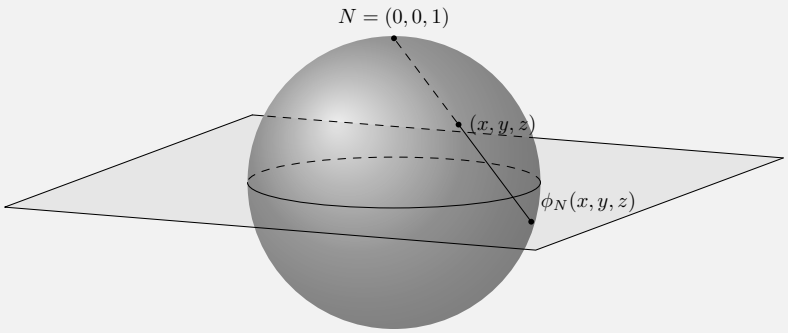
Figure 2.2: Illustration of two overlapping charts on a smooth manifold M. Inspired by (Lee, 2012, Figure 1.6). Image from (Jansson, 2022).

Example 2.2: The sphere

The sphere $\mathbb{S}^2 = \{x \in \mathbb{R}^3 : \|x\|_{\mathbb{R}^3} = 1\}$ is a smooth manifold of dimension 2. To construct an atlas for $\mathbb{S}^2$, we can use the stereographic projection. Let $N = (0, 0, 1)$ be the North Pole and $S = (0, 0, -1)$ be the South Pole. Let $U_N = \mathbb{S}^2 \setminus \{N\}$ and $U_S = \mathbb{S}^2 \setminus \{S\}$. The stereographic projection from the North Pole is the mapping $\phi_N : U_N \to \mathbb{R}^2$ given by

$$\phi_N(x, y, z) = \left(\frac{x}{1 - z}, \frac{y}{1 - z} \right).$$

Similarly, we can construct the stereographic projection from the South Pole, $\phi_S : U_S \to \mathbb{R}^2$ given by $\phi_S(x, y, z) = -\phi_N(-x, -y, -z)$.
The stereographic projection is illustrated below. The figure is inspired by (Lee, 2012, Figure 1.13).



Both (U_N, ϕ_N) and (U_S, ϕ_S) are charts, and the union of these two charts is a smooth atlas for $\mathbb{S}^2$ (Arnold, 1989, Section 4.18B).

The manifold of possible positions Q of a mechanical system does not alone determine the dynamics of the system, i.e., how the system evolves in time. In Lagrangian mechanics, we use positions and velocities to describe the motion of systems. For instance, to describe the motion of the pendulum, we need to know how the angle θ changes over time, i.e., the angular velocity $\dot{\theta} = \frac{\mathrm{d}\theta}{\mathrm{d}t}$. In general, the velocities at $x \in Q$ are described in terms of *tangent vectors* on Q, i.e., the set of vectors tangent to Q at x. These vectors form a vector space known as the *tangent space at* x and is denoted T_xQ. Equivalently, T_xQ is the set of all *derivations* of smooth functions $C^\infty(Q)$ at x, that is, the linear maps $v : C^\infty(Q) \to \mathbb{R}$ satisfying the product rule

$$v(fg) = f(x)v(g) + g(x)v(f)$$

for $f, g \in C^\infty(Q)$. By taking the disjoint union of the tangent spaces at each $x \in Q$, we get the *tangent bundle* TQ of Q. The tangent bundle is a manifold of dimension $2n$, where n is the dimension of Q (Lee, 2012, Proposition 3.18). In particular, the tangent bundle can be equipped with local coordinates. By (Lee, 2012, Corollary 3.3), given local coordinates $(x_1, \ldots, x_n)$ on Q, the tangent space T_xQ at x has a basis $\frac{\partial}{\partial x_1}\Big|_x, \ldots, \frac{\partial}{\partial x_n}\Big|_x$ where

$$\frac{\partial}{\partial x_i}\bigg|_x f = \frac{\partial f}{\partial x_i}(x),$$

meaning that a tangent vector $v \in T_xQ$ in local coordinates has the expression

$$v = \sum_{i=1}^{n} v_i \frac{\partial}{\partial x_i},$$

so by (Lee, 2012, Chapter 3), the tangent bundle TQ has local coordinates $(x_1, \ldots, x_n, v_1, \ldots, v_n)$, where v_i are the components of the tangent vector v.

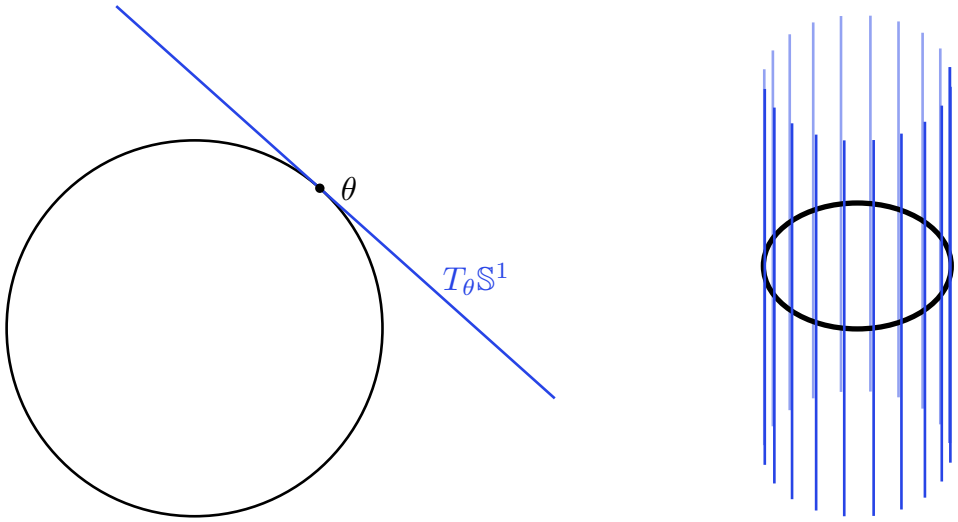
The single pendulum, or rather, its configuration manifold, provides a good example of the tangent bundle.

Example 2.3: Pendulum velocities

For the single pendulum, $Q = \mathbb{S}^1$. The tangent space at θ is $T_\theta\mathbb{S}^1 = \mathbb{R}$. Therefore, the tangent bundle is $T\mathbb{S}^1 = \mathbb{S}^1 \times \mathbb{R}$, i.e., a cylinder. This is illustrated in Figure 2.3.

Tangent vectors allow us to define the *differential* of a smooth mapping between manifolds. Given a smooth map $F\colon M \to N$ between smooth manifolds M and N, the differential of F at $x \in M$ is a linear map $dF_x\colon T_xM \to T_{F(x)}N$. If tangent vectors are viewed as derivations, then the differential of F at x is given by

$$dF_x(v)(f) = v(f \circ F),$$



(a) $T_x\mathbb{S}^1$ (b) $T\,\mathbb{S}^1$

Figure 2.3: Illustration of $T\mathbb{S}^1$. The tangent space $T_\theta\mathbb{S}^1$ at $\theta \in \mathbb{S}^1$ is the real line. By attaching a real line to each $\theta \in \mathbb{S}^1$, we obtain a cylinder of infinite height. Image from (Jansson, 2022).

for all $f \in C^\infty(N)$.

Given a configuration manifold Q of a mechanical system, we can assign to every point a tangent vector, i.e., a field of velocities. Mathematically, we define a mapping X from Q into the tangent bundle $T\,Q$. If this mapping is a section of TQ, meaning that it is a continuous map $X : Q \to TQ$ with the property that for all $x \in Q$,

$$p(X(x)) = x, \tag{2.1}$$

where p is the natural projection sending $v \in T_xQ$ to x, then X is a *vector field*. Equation (2.1) means that we attach a tangent vector to each point of Q. If the map from Q to $T\,Q$ is smooth, then X is a smooth vector field. Equivalently, vector fields are linear maps on $C^\infty(Q)$ that satisfy

$$X(fg) = gX(f) + fX(g),$$

for all $f, g \in C^\infty(Q)$, i.e., they are derivations. The space of all smooth vector fields on Q is denoted $\mathfrak{X}(Q)$. For any $f \in C^\infty(Q)$, the *commutator* $X, Y \in \mathfrak{X}(Q)$ is the vector field given by

$$[X, Y](f) = X(Y(f)) - Y(X(f)). \tag{2.2}$$

If we know the vector field X, we can describe the trajectory of a system by the curve $\gamma(t)$ on M obtained by solving the differential equation

$$\dot{\gamma} \coloneqq \frac{\mathrm{d}}{\mathrm{d}t}\gamma = X \circ \gamma$$

for all $t \in \mathrm{Dom}(\gamma) \subset \mathbb{R}$. The curve γ is called an *integral curve* of X.

2.2 Equations of Motion: Lagrangian and Hamiltonian Mechanics

Let us return to the pendulum. The state of the pendulum is described by the angle θ and the angular velocity $\dot{\theta}$. The movement of the pendulum is determined by how gravity acts on the mass, i.e., by pulling the pendulum downwards. One of the basic tenets of mechanics is Newton's second law, which states that the force acting on an object is equal to the mass of the object times its acceleration. By inserting the gravitational force $F = mg\sin\theta$, where m is the mass, we recover the equations of motion of the simple pendulum,

$$\ddot{\theta} + \frac{g}{l}\sin\theta = 0, \tag{2.3}$$

where g is the gravitational constant and, as before, l is the length of the rod.

An alternative characterization of motion is that a system behaves as to extremize a certain quantity, the *action functional*. In other words, mechanical systems obey the *principle of least action*. Formally, a system is described by a smooth function $\mathcal{L} : TQ \to \mathbb{R}$, called the *Lagrangian*. Given a path, i.e., a smooth curve $\gamma\colon [t_0, t_1] \to Q$, the action functional is

$$A[\gamma] = \int_{t_0}^{t_1} \mathcal{L}(\gamma(t), \dot{\gamma}(t))\,\mathrm{d}t.$$

If γ extremizes the action functional, it is known as a *motion* of the mechanical system. The principle of least action is an example of a *variational principle*. The equations of motion of the system are given by

$$\frac{\partial \mathcal{L}}{\partial \gamma} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\gamma}} = 0. \tag{2.4}$$

These equations are known as the *Euler–Lagrange equations*, and are derived by extremizing the action functional, see e.g., (Marsden and Ratiu, 1999, Section 7.1).

As an example, let us again consider the single pendulum.

Example 2.4: Pendulum as a Lagrangian system

The potential energy of the mass is $V = mgl(1 - \cos\theta)$, and its kinetic energy is $T = \frac{1}{2}ml^2\dot{\theta}^2$. The Lagrangian is given by the difference of kinetic energy and potential energy, i.e.,

$$\mathcal{L}(\theta, \dot{\theta}) = T - V = \frac{1}{2}ml^2\dot{\theta}^2 - mgl(1 - \cos\theta).$$

Therefore, by Equation (2.4), the equation of motion for the pendulum

is

$$\ddot{\theta} + \frac{g}{l}\sin\theta = 0,$$

and we recover Equation (2.3).

Another basic example is that of a particle moving in $\mathbb{R}^3$ under the influence of a potential V.

Example 2.5: Particle in potential part I

Let x be the position of a particle of mass m in $\mathbb{R}^3$. Its velocity $\dot{x}$ is denoted by v, so the kinetic energy is given by $T = \frac{m}{2}\|v\|^2_{\mathbb{R}^3}$. The potential energy is denoted by V. In this case, the Lagrangian is

$$\mathcal{L}(x, v) = \frac{m}{2}\|v\|^2_{\mathbb{R}^3} - V(x).$$

By the Euler–Lagrange equations (2.4), the equation of motion is $m\dot{v} = m\ddot{x} = -\nabla V(x)$, and we recover Newton's second law.
Consider now a first order perturbation, or variation, of the velocity v, i.e., $v + \varepsilon\delta w$. The perturbation of the velocity induces a perturbation of the kinetic energy, i.e.,

$$\delta T = \langle mv, \delta w\rangle_{\mathbb{R}^3}.$$

Note that $P(\cdot) = \langle mv, \cdot\rangle_{\mathbb{R}^3} \colon T_x\mathbb{R}^3 \to \mathbb{R}$ is a linear functional that assigns to each perturbation of the velocity the associated perturbation of the kinetic energy. We can identify P with $p = mv$, and we recognize the *momentum* of the particle. Having introduced the momentum, we can rewrite the equations of motion as

$$\begin{aligned} \dot{x} &= \frac{1}{m}p, \\ \dot{p} &= -\nabla V(x), \end{aligned} \tag{2.5}$$

which are *Hamilton's equations* for the particle. This is a first example of a *Hamiltonian system*.

For a general mechanical system, a similar procedure as in the above example can be applied to arrive at a Hamiltonian system from a Lagrangian. First, consider the dual of the tangent space, i.e., the *cotangent space* T^*_qQ consisting of the linear functionals of T_qQ. The *cotangent bundle* T^*Q is the disjoint union of the cotangent spaces at each $q \in Q$. The cotangent bundle is a manifold of dimension $2n$, where n is the dimension of Q (Lee, 2012,

Proposition 3.18).

Let us, as in (Marsden and Ratiu, 1999, Chapter 7), consider a mechanical system on Q with Lagrangian $\mathcal{L} : TQ \to \mathbb{R}$. If q is fixed, then $\mathcal{L}(q, \cdot) \coloneqq \mathcal{L}_q$ is a function from T_qQ to $\mathbb{R}$. Take then the directional derivative of $\mathcal{L}_q$ at $v \in T_qQ$, i.e.,

$$F\mathcal{L}(v)w = \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} \mathcal{L}_q(v + \varepsilon w), \tag{2.6}$$

where $w \in T_qQ$ is another tangent vector. In other words, we have a mapping $F\mathcal{L} : TQ \to T^*Q$ that assigns to each $v \in T_qQ$ an element of the cotangent space T_q^*Q, in analogy with Example 2.5. This mapping is known as the *fiber derivative* of $\mathcal{L}$. As noted in (Marsden and Ratiu, 1999, Chapter 7), if we work in local coordinates $(q_1, \ldots, q_n, v_1, \ldots, v_n)$, then the Equation (2.6) is given by

$$(q_1, \ldots, q_n, v_1, \ldots, v_n) \mapsto \left(q_1, \ldots, q_n, \frac{\partial \mathcal{L}}{\partial v_1}, \ldots, \frac{\partial \mathcal{L}}{\partial v_n}\right),$$

and we set $p_i = \frac{\partial \mathcal{L}}{\partial v_i}$. We refer to $F\mathcal{L}$ as the Legendre transform and define the associated energy function $E : TQ \to \mathbb{R}$ by

$$E(v) = F\mathcal{L}(v)v - \mathcal{L}_q(v).$$

If $F\mathcal{L}$ is a diffeomorphism (the Lagrangian is *hyperregular*), we can define the *Hamiltonian* $\mathcal{H} : T^*Q \to \mathbb{R}$ by

$$\mathcal{H} = E \circ (F\mathcal{L})^{-1} \colon T^*Q \to \mathbb{R}.$$

In this case, we can translate the Euler–Lagrange equations to Hamilton's equations (Marsden and Ratiu, 1999, Section 7.4) and obtain in local coordinates

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i}. \tag{2.7}$$

Thus, in the case that the Lagrangian is hyperregular, we can move from a description of the motion on the tangent bundle to a description on the cotangent bundle and the Hamiltonian and Lagrangian viewpoints are two sides of the same coin.

To illustrate the Legendre transform, we consider a Lagrangian system that is hyperregular and can be translated to a Hamiltonian system, namely the particle in a potential.

Example 2.6: Particle in potential part II

Let us revisit Example 2.5, i.e., the particle moving in a potential V. The configuration space is $Q = \mathbb{R}^3$. The cotangent space at x is $T_x^*\mathbb{R}^3 = \mathbb{R}^3$, and the cotangent bundle is $T^*\mathbb{R}^3 = \mathbb{R}^3 \times \mathbb{R}^3$. As noted in Example 2.5, the Lagrangian is $\mathcal{L}(x, v) = \frac{m}{2}\|v\|_{\mathbb{R}^3}^2 - V(x)$. The fiber derivative is just

$$F\mathcal{L}(v)w = \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} \mathcal{L}(x, v + \varepsilon w) = \langle mv, w\rangle_{\mathbb{R}^3},$$

which we recall from Example 2.5 as the induced perturbation of the kinetic energy. The associated energy function is

$$E(v) = \langle mv, v\rangle_{\mathbb{R}^3} - \frac{m}{2}\|v\|_{\mathbb{R}^3}^2 + V(x) = \frac{m}{2}\|v\|_{\mathbb{R}^3}^2 + V(x)$$

Note that in this case, $F\mathcal{L}$ is an isomorphism (we are merely working with vector spaces). As we can identify $p = mv \in \mathbb{R}^3$ so that $v = \frac{1}{m}p$, the inverse of $F\mathcal{L}$ is the mapping $(x, \langle p, \cdot\rangle_{\mathbb{R}^3}) \mapsto (x, p/m)$. Thus, the Hamiltonian $\mathcal{H} : T^*\mathbb{R}^3 \to \mathbb{R}$ is defined by

$$\mathcal{H}(x, p) = \frac{1}{2m}\|p\|_{\mathbb{R}^3}^2 + V(x), \tag{2.8}$$

and we can recover Hamilton's equation for the particle (2.5) by inserting the Hamiltonian (2.8) into (2.7).

The cotangent bundle is an example of a *phase space* of a mechanical system. Its generalization is the concept of *symplectic manifolds*, which is a manifold equipped with an additional structure known as the *symplectic form*. To understand symplectic manifolds, we need to introduce the concept of *differential forms*. Informally, a differential form is a function that assigns to each point of a manifold a multilinear mapping that inputs tangent vectors and outputs the oriented volume spanned by the vectors.

More formally, a k-form is a field of alternating k-linear maps on the tangent spaces of a manifold, i.e., for each $x \in M$, a k-form ω is a map $\omega_x : T_xM \times \ldots \times T_xM \to \mathbb{R}$ that is linear in each argument and antisymmetric, i.e., the sign changes if we swap two arguments. We denote the space of k-forms on M by $\Omega^k(M)$. The *wedge product* of two forms $\omega \in \Omega^k(M)$ and $\eta \in \Omega^l(M)$ is a $(k+l)$-form, denoted $\omega \wedge \eta$. The wedge product is bilinear and associative. If kl is even, the wedge product is commutative, and otherwise it is anti-commutative.

The *interior product* of a k-form ω and a vector field X is a $(k-1)$-form,

denoted $\iota_X\omega$ and is given by

$$\iota_X\omega(Y_1,\ldots,Y_{k-1}) = \omega(X,Y_1,\ldots,Y_{k-1}),$$

where $Y_1,\ldots,Y_{k-1}$ are vector fields.

The *exterior derivative* $d\colon \Omega^k(M) \to \Omega^{k+1}(M)$ is the $\mathbb{R}$-linear mapping that satisfies $d^2 = 0$ and the product rule

$$d(\alpha\wedge\beta) = d\alpha\wedge\beta + (-1)^k\alpha\wedge d\beta,$$

where $\alpha\in\Omega^k(M)$ and $\beta\in\Omega^l(M)$. A k-form α is exact if there is a $(k-1)$-form β such that $\alpha = d\beta$. Finally, we say that α is closed if $d\alpha = 0$. Note in particular that all exact forms are closed, but that not all closed forms are exact.

Example 2.7: Integration of differential forms

An important feature of differential forms is that they can be integrated. On a Euclidean space $\mathbb{R}^n$, an n-form ω is given by $\omega = \phi(x)dx_1\wedge\ldots\wedge dx_n$, for some smooth function ϕ. The integration of ω over a bounded and convex domain D is the same as in the case of multivariate calculus, i.e.,

$$\int_D \omega \coloneqq \int_D \phi(x)\ \mathrm{d}x_1\cdots\mathrm{d}x_k,$$

To integrate a k-form over a manifold M, we must carry over the notion of integration from Euclidean space to the manifold. One way to do this is to triangulate the manifold with bounded convex polyhedra, and then to integrate over the k-chains defined by the triangulation. A k-chain is a formal sum of k-cells, where a k-cell is a triple (D,F,O), where D is a bounded convex polyhedron in $\mathbb{R}^k$, $F : D\to M$ is a differentiable map, and O is an orientation of $\mathbb{R}^k$. The integral of a k-form ω over a k-cell $\sigma = (D,F,O)$ is

$$\int_\sigma \omega = \int_D F^*\omega.$$

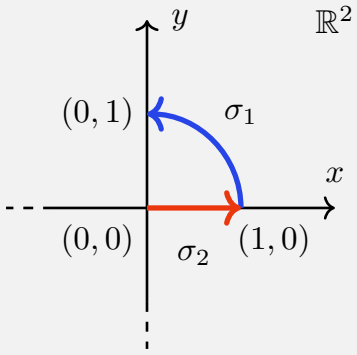
Here, $F^*\omega(X_1,\ldots,X_k) \coloneqq \omega(dF_x(X_1),\ldots,dF_x(X_k))$ is the *pullback* of ω, where $X_1,\ldots,X_k$ are tangent vectors on D. The orientation of the k-cell is used to determine the sign of the integral over D, as usual in multivariate calculus.

Given $r\in\mathbb{N}$ k-cells $(D_1,F_1,O_1),\ldots,(D_r,F_r,O_r)$ and integers $m_1,\ldots,m_r$, a k-chain c_k on M is the formal sum $c_k = \sum_{i=1}^r m_i\sigma_i$.

The integral of ω over c_k is

$$\int_{c_k} \omega = \sum_{i=1}^{r} m_i \int_{\sigma_i} \omega. \tag{2.9}$$

As an example, consider the integral of the 1-form $\omega = xdx - ydy$ on $\mathbb{R}^2$ over the submanifold K shown below.



The figure is from (Jansson, 2022). Since K consists of two parts, a quarter circle σ_1 and a line segment σ_2, we just need two 1-cells. To cover the circular arch, we take $D_1 = [0, \pi/2], \phi_1^{(t)} = (\cos(t), \sin(t))$. The line segment is covered by $D_2 = [0, 1], \phi_2^{(t)} = (t, 0)$. Equation (2.9) then gives

$$\int_K \omega = \int_{\sigma_1+\sigma_2} (xdx - ydy) = \int_0^{\pi/2} -\sin(2t)\ \mathrm{d}t + \int_0^1 t\ \mathrm{d}t = -\frac{1}{2}.$$

Now we can note something interesting about the integral of ω over K. Note that $\omega = dF$, where $F = x^2/2 - y^2/2$, so ω is exact. The boundary of K is $\partial K = [(0, 1)] - [(0, 0)]$. If we integrate F over ∂K, we get

$$\int_{\partial K} F = F(0, 1) - F(0, 0) = -1/2,$$

i.e., the integral of F over the boundary of K equals the integral of dF over K.

This is no coincidence, but a consequence of *Stokes' theorem* (Fortney, 2018, Theorem 11.1). Stokes' theorem states that for a manifold M of dimension k with boundary ∂M, it holds that

$$\int_{\partial M} \omega = \int_M d\omega,$$

where $\omega \in \Omega^{k-1}(M)$.

A symplectic form is just a non-degenerate closed 2-form. Non-degeneracy means that for every $x \in M$, if $\omega_x(u, v) = 0$ for all $v \in T_xM$, then $u = 0$. Not all manifolds can be equipped with a symplectic form. Indeed, all symplectic manifolds are even-dimensional (Lee, 2012, Proposition 22.7), thereby excluding many manifolds, such as the circle, the ball and odd-dimensional vector spaces. However, the fact that symplectic manifolds are even-dimensional is indicative of how they are used in mechanics. In a phase space, one doubles the number of variables by considering both position and momentum variables, meaning that we need two sets of coordinates, i.e., we necessarily work in an even-dimensional space.

We have already encountered quite a few examples of symplectic manifolds.

Example 2.8: Symplectic manifolds

1. The sphere, $\mathbb{S}^2$, is a symplectic manifold with symplectic form.

$$\omega_x(u, v) = \langle x, (u \times v) \rangle_{\mathbb{R}^3},$$

   where $x \in \mathbb{S}^2$, $u, v \in T_x\mathbb{S}^2$, $\times$ is the cross product and $\langle \cdot, \cdot \rangle_{\mathbb{R}^3}$ is the Euclidean inner product. Here, we think of $\mathbb{S}^2$ as a submanifold of $\mathbb{R}^3$.

2. If V is a real vector space of dimension n, then

$$T^*V = V \times V^*,$$

   where V^* is the dual of V, is a symplectic manifold when equipped with the *canonical symplectic form* given by

$$\omega((v_1, \alpha_1), (v_2, \alpha_2)) = \alpha_2(v_1) - \alpha_1(v_2),$$

   where $v_1, v_2 \in V$ and $\alpha_1, \alpha_2 \in V^*$. Cotangent bundles can in general equipped with a canonical symplectic form (Lee, 2012, Proposition 22.11).

The symplectic form is used to describe the dynamics of a system evolving on a symplectic manifold. Given a smooth function $\mathcal{H}$ on a symplectic manifold (M, ω), its *Hamiltonian vector field* $X_\mathcal{H}$ is defined by

$$d\mathcal{H} = \omega(\cdot, X_\mathcal{H}). \tag{2.10}$$

The function $\mathcal{H}$ is called the *Hamiltonian*, and the dynamics of the system it defines are given by the general form of Hamilton's equations

$$\dot{z} = X_\mathcal{H}(z),$$

where $z(t) \in M$. Thus, all we need to do Hamiltonian mechanics is a Hamiltonian function on a symplectic manifold.

2.3 Symmetries: Lie Groups and Lie Algebras

An important part of mechanics is the study of symmetries and conserved quantities of a system. Let us start with an example.

Example 2.9: Particle in potential part III

Recall the setting from Examples 2.5 and 2.6, i.e., the particle moving in $\mathbb{R}^3$ under the influence of a potential V. Assume that the potential only depends on the distance to the origin, $V(x) = \tilde{V}(\|x\|_{\mathbb{R}^3})$ for some $\tilde{V}\colon \mathbb{R} \to \mathbb{R}$. If this is the case, then the Lagrangian is *invariant* under rotations, meaning that if R is a rotation matrix (the set of which is denoted SO(3)), then

$$\mathcal{L}(Rx, R\dot{x}) = \mathcal{L}(x, \dot{x}).$$

In other words, the system is rotationally *symmetric*.
This has an interesting consequence. Consider the *angular momentum* given by $L = x \times p$, where $p = m\dot{x}$ is the momentum of the particle. By using that $\dot{p} = -\nabla V(x)$ and $\dot{x} = \frac{1}{m}p$ we see that the time derivative of L along a solution is

$$\dot{L} = \dot{x} \times p + x \times \dot{p} = x \times \dot{p} = \nabla V(\|x\|_{\mathbb{R}^3}) \times x = \frac{V'(\|x\|_{\mathbb{R}^3})}{\|x\|_{\mathbb{R}^3}} x \times x = 0,$$

i.e., L is a *conserved quantity* of the system. The conservation of angular momentum is a direct consequence of the rotational symmetry of the system.
Consider now the case when the potential is invariant under arbitrary translations in direction of $x_1 = (1, 0, 0)$, the first coordinate direction of $\mathbb{R}^3$. In formulas, this means that $V(x) = V(x + ax_1)$ for all $a \in \mathbb{R}$. Then, the Lagrangian is invariant under translations in the x_1-direction,

$$\mathcal{L}(x + ax_1, \dot{x}) = \mathcal{L}(x, \dot{x}).$$

Similarly to the rotational invariant case, the system has a conserved quantity, the *first component of the linear momentum* $p_1 = m(\dot{x})_1$ since

$$\dot{p}_1 = m(\ddot{x})_1 = -(\nabla V(x))_1 = 0.$$

Thus, the conservation of the first component of the linear momentum is a direct consequence of the translational symmetry in the x_1-direction.

We have now seen two examples of how a symmetry of a system leads to a conserved quantity. This is, as we shall see, not a coincidence.

To study symmetry, we need to introduce a way to consider continuous transformations acting on systems, for instance, a way to mathematically describe how a system can be rotated or translated. The properties that remain unchanged under continuous transformations are known as continuous symmetries, and these are mathematically described by *Lie groups*. A Lie group H is a smooth manifold that is also a group such that the map $(h_1, h_2) \mapsto h_1 h_2^{-1}$ is smooth for all $h_1, h_2 \in H$. A Lie subgroup of H is a submanifold that is closed under the group operations.

The *Lie algebra* $\mathfrak{h}$ of a Lie group H is defined as the tangent space of H at the identity $e \in H$. In general, a Lie algebra is just a vector space equipped with a *Lie bracket*, which is a bilinear map $[\cdot, \cdot] \colon \mathfrak{h} \times \mathfrak{h} \to \mathfrak{h}$ satisfying

- *Anticommutativity*: $[X, Y] = -[Y, X]$,
- *Jacobi identity*: $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$,

for all $X, Y, Z \in \mathfrak{h}$. A subspace of a Lie algebra that is closed under the Lie bracket is called a *Lie subalgebra*. The space of vector fields $\mathfrak{X}(Q)$ on a manifold Q is a Lie algebra with bracket given by Equation (2.2).

The Lie bracket on a Lie algebra $\mathfrak{h}$ of a Lie group H arises from the commutator of vector fields on H. To see this, we denote left translation by

$$L_{h_1} h_2 = h_1 h_2,$$

and right translation by

$$R_{h_1} h_2 = h_2 h_1,$$

where $h_1, h_2 \in H$. Left and right translations are by definition smooth mappings. Because inversion is also smooth on Lie groups, the inverse of left and right translation are also smooth, that is to say, $L_h^{-1} = L_{h^{-1}}$ and $R_h^{-1} = R_{h^{-1}}$ are smooth maps. The differential of left translation at a point $h' \in H$ acting on a vector $v \in T_{h'}H$ is denoted by $d(L_h)_{h'}(v)$ and produces a vector in $T_{hh'}H$.

Now, let ξ be a fixed element of $\mathfrak{h} = T_eH$. We have that $d(L_h)_e(\xi) \in T_hH$, and so we have a mapping from H to TH given by $h \mapsto (h, v^\xi(h))$ where $v^\xi(h) = d(L_h)_e(\xi)$, meaning that this is a vector field. Moreover, $d(L_{h'})v^\xi(h) = v^\xi(h'h)$ meaning that v^ξ is a *left-invariant* vector field. We notice directly that a left-invariant vector field is completely determined by its value at the identity, so we can identify the set of left-invariant vector fields with $\mathfrak{h}$ as vector spaces (Marsden and Ratiu, 1999, Section 9.1). In fact,

a bit more can be said. It turns out that the Lie bracket of two left-invariant vector fields is also left-invariant, meaning that the set of left-invariant vector fields is a Lie subalgebra of $\mathfrak{X}(H)$. The Lie bracket on $\mathfrak{h}$ is given by the Lie bracket of left-invariant vector fields, i.e.,

$$[\xi_1, \xi_2] = [v^{\xi_1}, v^{\xi_2}](e),$$

where v^{ξ_1} and v^{ξ_2} are left-invariant vector fields corresponding to ξ_1 and ξ_2.

Lie groups come equipped with an *exponential map* $\exp\colon \mathfrak{h} \to H$ that maps elements of the Lie algebra to the Lie group. The exponential map of an element X is defined by considering the flow, that is, the solution to (2.1) of the left-invariant vector field v_X at time 1;

$$\exp(X) = \gamma_{v_X}(1).$$

An important example of Lie groups can be found in the various matrix groups.

Example 2.10: Matrix Lie groups

A classical example of a Lie group is the *general linear group* $\mathrm{GL}(n)$, which consists of all real invertible $n \times n$ matrices. Its Lie algebra, denoted by $\mathfrak{gl}(n)$, is the space of all $n \times n$ matrices. The Lie bracket on $\mathfrak{gl}(n)$ is $[A, B] = AB - BA$, i.e., the matrix commutator.
There are several important subgroups of $\mathrm{GL}(n)$ that are also Lie groups. One example is $\mathrm{O}(n)$, which contains all orthogonal matrices of dimension n. Another example is $\mathrm{SO}(n)$, which consists of elements of $\mathrm{O}(n)$ with determinant 1. Note that while $\mathrm{SO}(n) \subset \mathrm{O}(n)$, their Lie algebras are the same: $\mathfrak{so}(n) = \mathfrak{o}(n) = \{A \in \mathfrak{gl}(n) : A^\top = -A\}$.

A key property of Lie groups is that they can act on other manifolds. If H is a Lie group and X is a set, a *left action* of H on X is a map $\Phi\colon H \times X \to X$ that satisfies

$$\Phi(e, x) = x, \quad \Phi(g, \Phi(h, x)) = \Phi(gh, x),$$

for all $g, h \in H$ and $x \in X$. A *right action* is defined similarly, as a map $\Phi\colon X \times H \to X$, but with the composition rule now being from the right,

$$\Phi(x, gh) = \Phi(\Phi(x, g), h).$$

If the group is clear from the context, we often write $\Phi(h, x)$ as $h \cdot x$. If Φ is a smooth map and X is a smooth manifold, then we say that the group action is smooth. Now, if H acts on X, the *orbit* of $x \in X$ is the set

$$\mathcal{O}_x = \{h \cdot x : h \in H\},$$

in other words, the set of all elements that x can be mapped to by the group action. The *stabilizer* of x is the set of all elements that leave x fixed,

$$H_x = \{h \in H : h \cdot x = x\}.$$

The stabilizer is a subgroup of H. A group action is *transitive* if there is only one orbit, i.e., for all $x, y \in X$, there is an $h \in H$ such that $h \cdot x = y$. If all stabilizers are trivial, meaning that $H_x = \{e\}$ for all $x \in X$, then the action is *free*.

We have already seen how Lie groups can act on themselves by right or left translation. Another important example is the action of a Lie group H on itself by conjugation,

$$C_{h_1} h_2 = L_{h_1} R_{h_1^{-1}} h_2 = h_1 h_2 h_1^{-1},$$

where $h_1, h_2 \in H$. The conjugation action gives rise to the *adjoint representation*, denoted by $\mathrm{Ad}_h \in \mathrm{End}(\mathfrak{h})$, which is given by the differential of C_h at the identity. Here, $\mathrm{End}(\mathfrak{h})$ denotes the space of endomorphisms of $\mathfrak{h}$, that is, linear maps from $\mathfrak{h}$ to $\mathfrak{h}$. By varying h we get the map $\mathrm{Ad} : H \to \mathrm{End}(\mathfrak{h})$. Differentiating Ad we obtain the map $\mathrm{ad} : \mathfrak{h} \to \mathrm{End}(\mathfrak{h})$. Explicitly, $\mathfrak{h} \ni Y \mapsto [Y, \cdot] = \mathrm{ad}_Y$.

The coadjoint representation is defined by

$$\langle \mathrm{Ad}_h^* \, \xi, Y \rangle = \langle \xi, \mathrm{Ad}_{h^{-1}} \, Y \rangle,$$

for all $\xi \in \mathfrak{h}^*$ and $Y \in \mathfrak{h}$. Here $\langle \cdot, \cdot \rangle$ is the dual pairing between $\mathfrak{h}$ and $\mathfrak{h}^*$.

Likewise, the dual of ad is denoted by ad^*, and is a map from $\mathfrak{h}$ to $\mathrm{End}(\mathfrak{h}^*)$. Explicitly, it is given by

$$\langle \mathrm{ad}_X^* \, \eta, Y \rangle = \langle \eta, -\mathrm{ad}_X \, Y \rangle = \langle \eta, [Y, X] \rangle,$$

for all $\eta \in \mathfrak{h}^*$ and $X, Y \in \mathfrak{h}$.

The orbits of Ad^* are known as *coadjoint orbits*. The coadjoint orbit of $\xi \in \mathfrak{h}$ is the set

$$\mathcal{O}_\xi = \{\mathrm{Ad}_h^* \, \xi, h \in H\}.$$

The coadjoint orbits can be endowed with a symplectic structure known as the *Kirillov–Kostant–Souriau form*. To write down this form, we first remark that by Marsden and Ratiu (1999, Section 14.2), coadjoint orbits are manifolds, and that the tangent space of the coadjoint orbit $\mathcal{O}_\xi$ at $\eta \in \mathcal{O}_\xi$ is given by

$$T_\eta \mathcal{O}_\xi = \{\mathrm{ad}_\xi^* \, \eta : \xi \in \mathfrak{h}\}.$$

The symplectic form is then given by

$$\omega_\eta(\mathrm{ad}_{\xi_1}^* \, \eta, \mathrm{ad}_{\xi_2}^* \, \eta) = \langle \eta, [\xi_1, \xi_2] \rangle,$$

where $\xi_1, \xi_2 \in \mathfrak{h}$ and $\eta \in \mathcal{O}_\xi$. Moreover, the coadjoint orbits "slice up" the dual algebra. This means that the coadjoint orbits provide a special kind of partition of $\mathfrak{h}^*$ known as a *foliation* (Kirillov, 2004).

To give an example of the above concepts, we consider the special orthogonal group $\mathrm{SO}(n)$.

Example 2.11: Special orthogonal group

Let $G, H \in \mathrm{SO}(n)$. The conjugation action of G on H is given by

$$C_G H = GHG^{-1} = GHG^\top.$$

The Lie algebra of $\mathrm{SO}(n)$ is $\mathfrak{so}(n)$, which consists of all $n \times n$ skew-symmetric matrices. To compute the adjoint representation, we differentiate C_G at the identity, i.e., we insert the curve $H(t) = I + tX + \mathsf{O}(t^2)$ for some $X \in \mathfrak{so}(n)$, where I is the identity matrix, into C_G and differentiate at $t = 0$. Thus, we obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} C_G H(t)\bigg|_{t=0} &= \frac{\mathrm{d}}{\mathrm{d}t} GIG^\top + tGXG^\top + \mathsf{O}(t^2)\bigg|_{t=0} \\ &= GXG^\top = \mathrm{Ad}_G X.\end{aligned}$$

To find the coadjoint representation, we first identify $\mathfrak{so}(n)$ with $\mathfrak{so}(n)^*$ by the Frobenius inner product $\langle A, B\rangle = -\operatorname{Tr}(AB)$. Therefore, the coadjoint representation is given by

$$\begin{aligned}\langle \mathrm{Ad}^*_G \xi, Y\rangle &= \langle \xi, \mathrm{Ad}_{G^\top} \ Y\rangle = -\operatorname{Tr}(\xi G^\top Y G) \\ &= -\operatorname{Tr}(G\xi G^\top Y) = \langle G\xi G^\top, Y\rangle,\end{aligned}$$

where $\xi \in \mathfrak{so}(n)^*$ and $Y \in \mathfrak{so}(n)$. We conclude that $\mathrm{Ad}^*_G \xi = G\xi G^\top$. Likewise, ad^* is computed by

$$\begin{aligned}&\langle \mathrm{ad}^*_X \eta, Y\rangle = \langle \eta, -\mathrm{ad}_X Y\rangle = -\operatorname{Tr}(\eta[Y, X]) = -\operatorname{Tr}(\eta YX - \eta XY) \\ &\quad = -\operatorname{Tr}((X\eta Y - \eta X)Y) = -\operatorname{Tr}([X, \eta]Y),\end{aligned}$$

meaning that $\mathrm{ad}^*_X \eta = [X, \eta]$. Finally, the coadjoint orbit of $\xi \in \mathfrak{so}(n)^*$ is the set of all $G\xi G^\top$ for $G \in \mathrm{SO}(n)$.
If $n = 3$ we can be even more explicit. We can identify $\mathfrak{so}(3)$ with $\mathbb{R}^3$ by the isomorphism

$$\xi = \begin{pmatrix} 0 & -\xi_3 & \xi_2 \\ \xi_3 & 0 & -\xi_1 \\ -\xi_2 & \xi_1 & 0 \end{pmatrix} \mapsto \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \hat{\xi}. \tag{2.11}$$

Thus, the coadjoint representation is $G\hat{\xi}$, and so the coadjoint orbit is

$$\mathcal{O}_{\hat{\xi}} = \{G\hat{\xi} : G \in \mathrm{SO}(3)\}.$$

Since G is a rotation matrix, the coadjoint orbit is the set of all vectors that can be obtained by rotating $\hat{\xi}$, i.e., the sphere of radius $\|\hat{\xi}\|_{\mathbb{R}^3}$. If $\mathfrak{so}(3)$ is identified with $\mathbb{R}^3$, then $\mathrm{ad}_{\hat{\xi}_1}\hat{\xi}_2 = \hat{\xi}_1 \times \hat{\xi}_2$. This means the tangent space to a coadjoint orbit $\mathcal{O}_{\hat{\xi}}$ at $G\hat{\xi}$ is

$$T_{G\hat{\xi}}\mathcal{O}_{\hat{\xi}} = \{v \times G\hat{\xi} : v \in \mathbb{R}^3\},$$

that is, the set of all vectors that are orthogonal to $G\hat{\xi}$, which we already knew from Example 2.2. The Kostant–Kirillov–Souriau form is given by

$$\omega_{G\hat{\xi}}(v_1 \times G\hat{\xi}, v_2 \times G\hat{\xi}) = \langle G\hat{\xi}, v_1 \times v_2\rangle$$

This is recognized as the standard symplectic form on the sphere from Example 2.8. The foliation of $\mathbb{R}^3$ by coadjoint orbits is just the foliation of $\mathbb{R}^3$ by concentric spheres, thus obtaining something resembling an onion, where the layers are the coadjoint orbits.

We are now ready to discuss symmetries of mechanical systems. Let us consider a system with configuration manifold Q and Lagrangian $\mathcal{L}\colon TQ \to \mathbb{R}$. Let H be a Lie group that acts on Q. This action induces the *tangent lifted action* of H on TQ. If the action of H on Q is given by $\Phi\colon H \times Q \to Q$, then $\Phi_h = \Phi(h, \cdot)$ defines a diffeomorphism of Q for each $h \in H$. The tangent lifted action $T\Phi\colon H \times TQ \to TQ$ is given by

$$T\Phi_h(q, v) = (\Phi_h(q), d(\Phi_h)_q(v)).$$

Example 2.12: Tangent lifted action of $\mathrm{SO}(3)$

Let us return to Example 2.9. The group $\mathrm{SO}(3)$ acts on $\mathbb{R}^3$ by rotation, i.e., $\Phi_R(x) = Rx$ where $R \in \mathrm{SO}(3)$ and $x \in \mathbb{R}^3$. One verifies that $d(\Phi_R)(v) = Rv$, so the tangent lifted action is given by $T\Phi_R(x, v) = (Rx, Rv)$.

Now, H is a symmetry of the system if the Lagrangian is invariant under the tangent lifted action of H, meaning that

$$\mathcal{L}(T\Phi_g(q, v)) = \mathcal{L}(q, v).$$

This implies that if $\gamma(t)$ follows the equations of motion, that is, solves the Euler–Lagrange equations (2.4), then so does $\Phi_g(\gamma(t))$.

The observation from Example 2.9—that if the system has a symmetry, then there is a conserved quantity—is formalized in Noether's theorem, which in its Lagrangian formulation states that

Theorem 1 (Noether's theorem, Lagrangian version)**.** *Let the Lagrangian $\mathcal{L}\colon TQ \to \mathbb{R}$ be invariant under the tangent lifted action of a Lie group H. Then there is a conserved quantity $I\colon TQ \to \mathfrak{h}^*$ given by*

$$I(q,v) = \frac{\partial \mathcal{L}(q,v)}{\partial v} \left.\frac{\partial \Phi_h(q)}{\partial h}\right|_{h=e}.$$

2.4 Observables: Poisson Systems

Hamiltonian mechanics gives rise to a natural understanding of *observables*, namely, simply as functions of the phase space. Given a symplectic manifold (M,ω), an observable is a smooth function $C^\infty(M) \ni f : M \to \mathbb{R}$. Two functions $f, g \in C^\infty(M)$ can be combined to form a new observable by taking their *Poisson bracket*, defined by

$$\{f,g\} = \omega(X_f, X_g),$$

where X_f and X_g are the Hamiltonian vector fields of f and g, respectively.

The Poisson bracket is a bilinear map $\{\cdot,\cdot\} : C^\infty(M) \times C^\infty(M) \to C^\infty(M)$ that satisfies

- *anticommutativity*: $\{f,g\} = -\{g,f\}$,
- *the Jacobi identity*: $\{f,\{g,h\}\} + \{g,\{h,f\}\} + \{h,\{f,g\}\} = 0$ and
- *the Leibniz rule*: $\{f,gh\} = g\{f,h\} + \{f,g\}h$.

In other words, the Poisson bracket is a Lie bracket on $C^\infty(M)$ which in addition satisfies the Leibniz rule. Further, if $\mathcal{H}$ is a Hamiltonian, that is, a function that describes the energy of a system, then the time evolution of an observable F is given by

$$\dot{F} = \{F, \mathcal{H}\}. \tag{2.12}$$

The concept of Poisson brackets allows for a generalization of symplectic manifolds to *Poisson manifolds*, where the bracket is not directly related to a symplectic form. Indeed, a Poisson manifold $(M, \{\cdot,\cdot\})$ is a smooth manifold M equipped with a bilinear map $\{\cdot,\cdot\} : C^\infty(M) \times C^\infty(M) \to C^\infty(M)$ that satisfies the same properties as the Poisson bracket on a symplectic manifold

above, i.e., anticommutativity, Jacobi identity, and the Leibniz rule. This mapping is also called a Poisson bracket, and, to reiterate, it does not need to arise from a symplectic form. It is in this case still possible to define Hamiltonian vector fields, by the mapping $f \mapsto X_f = \{f, \cdot\}$. In contrast to symplectic manifolds, odd-dimensional manifolds can be Poisson manifolds.

An important example of a Poisson bracket is that of the Lie–Poisson bracket on the dual $\mathfrak{h}^*$ of a Lie algebra $\mathfrak{h}$. Systems governed by this bracket are known as *Lie–Poisson systems*.

Example 2.13: Lie–Poisson systems

Let $\mathfrak{h}$ be a Lie algebra with Lie bracket $[\cdot, \cdot]$. To define the Lie–Poisson bracket on $C^\infty(\mathfrak{h}^*)$, we first note that since $T_\xi\mathfrak{h}^* \cong \mathfrak{h}^*$ for all $\xi \in \mathfrak{h}^*$, and since $df_\xi \colon T_\xi\mathfrak{h}^* \to \mathbb{R}$ is linear, the differential of f at ξ is an element of $\mathfrak{h}^{**}$ that we in turn identify with $\mathfrak{h}$.
The Lie–Poisson bracket on $C^\infty(\mathfrak{h}^*)$ is given by

$$\{F, G\}(\xi) = \langle \xi, [dF_\xi, dG_\xi] \rangle, \tag{2.13}$$

where $F, G \in C^\infty(\mathfrak{h}^*)$.
A Lie–Poisson system is a system whose dynamics are governed by the Lie–Poisson bracket. Let $\mathcal{H} \in C^\infty(\mathfrak{h}^*)$ denote the Hamiltonian. Then, the time evolution of an observable $F \in C^\infty(\mathfrak{h}^*)$ is given by Equation (2.12) with the Lie–Poisson bracket. Lie–Poisson systems have a number of conserved quantities, namely the *Casimir functions*, which are functions $C \in C^\infty(\mathfrak{h}^*)$ that satisfy $\{C, h\} = 0$ for all $h \in C^\infty(\mathfrak{h}^*)$. Another conserved quantity is the Hamiltonian, as the time evolution of the Hamiltonian is zero, i.e., $\dot{\mathcal{H}} = \{\mathcal{H}, \mathcal{H}\} = 0$.
There is an alternative formulation of Lie–Poisson systems. Let $\xi \in \mathfrak{h}^*$ be arbitrary. Then, the time evolution of the observable $F \in C^\infty(M)$ is given by, according to the chain rule,

$$\frac{\mathrm{d}}{\mathrm{d}t} F(\xi) = \langle \dot{\xi}, dF_\xi \rangle.$$

Further, by Equation (2.13), we have that

$$\langle \dot{\xi}, dF_\xi \rangle = \langle \xi, [dF_\xi, d\mathcal{H}_\xi] \rangle = \langle \xi, \mathrm{ad}_{d\mathcal{H}_\xi} dF_\xi \rangle = \langle \mathrm{ad}^*_{d\mathcal{H}_\xi} \xi, dF_\xi \rangle,$$

Since F is arbitrary, we conclude that

$$\dot{\xi} = \mathrm{ad}^*_{d\mathcal{H}_\xi} \xi. \tag{2.14}$$

Lie–Poisson systems are an important example of Poisson systems. They

arise in many contexts, such as in fluid dynamics, rigid body dynamics (Marsden and Ratiu, 1999, Chapter 1) and magneto-hydrodynamics (Morrison and Greene, 1980). A very concrete example of a Lie–Poisson system is the rigid body.

Example 2.14: The rigid body part I

Consider a rigid body of uniform density, i.e., a body that does not deform, meaning that the distance between any two points on the body is constant. The rigid body moves by rotating around its center of mass, which is assumed to be at the origin.
The equation of motion of the rigid body is

$$I\dot{\omega} = I\omega \times \omega, \tag{2.15}$$

where $\omega \in \mathfrak{so}(3) \cong \mathbb{R}^3$ is the angular velocity of the body, and I is the inertia tensor of the body, a real, positive definite, and symmetric 3×3 matrix. We assume that I is diagonal, i.e., $I = \operatorname{diag}(I_1, I_2, I_3)$ with $I_1, I_2, I_3 > 0$.
These equations do not have an immediate interpretation as a Hamiltonian system, as ω lives in a three-dimensional space. However, it can be interpreted as a Lie–Poisson system. To see this, we first note that the angular velocity ω is an element of $\mathfrak{so}(3)$, the Lie algebra of SO(3). The angular momentum $L = (L_1, L_2, L_3)$ is given by

$$L_1 = I_1\omega_1, \quad L_2 = I_2\omega_2, \quad L_3 = I_3\omega_3,$$

so that Equation (2.15) can be written as

$$\dot{L} = L \times I^{-1}L. \tag{2.16}$$

Now, we view L as an element of $\mathfrak{so}(3)^*$, which we identify with $\mathfrak{so}(3)$ through the standard inner product. Then, the Lie–Poisson bracket on $C^\infty(\mathfrak{so}(3)^*)$ is, by Equation (2.13),

$$\{F, \mathcal{H}\}(L) = \langle L, [dF_L, dH_L]\rangle.$$

As we identify $\mathfrak{so}(3)$ with $\mathbb{R}^3$, the Lie bracket can be verified to be the standard cross product, $[L_1, L_2] = L_1 \times L_2$. Furthermore, the differentials dF_L and dH_L coincide with the gradients of F and $\mathcal{H}$ with respect to L, and the Lie–Poisson bracket becomes

$$\{F, \mathcal{H}\}(L) = \langle L, \nabla F \times \nabla \mathcal{H}\rangle$$

Thus, if an appropriate Hamiltonian $\mathcal{H} \in C^\infty(\mathfrak{so}(3))$ can be identified, the rigid body equation (2.16) can be written as a Lie–Poisson equation.

Recall from example Example 2.11 that the coadjoint action of $\mathfrak{so}(3)$ on its dual is $\operatorname{ad}^*_\omega L = \omega \times L$, so

$$\{F, \mathcal{H}\}(L) = -\langle L \times \nabla\mathcal{H}, \nabla F\rangle = \langle \operatorname{ad}^*_{\nabla\mathcal{H}} L, \nabla F\rangle.$$

By Equation (2.14), the rigid body equations can be written as

$$\dot{L} = -L \times \nabla\mathcal{H},$$

where

$$\mathcal{H}(L) = -\frac{1}{2}(L_1/I_1, L_2/I_2, L_3/I_3).$$

This means in particular that $\mathcal{H}(L)$ is a preserved quantity. Further, any Casimir function of the Lie–Poisson bracket is a conserved quantity of the rigid body equations. In this case, a Casimir function is $C(L) = \|L\|^2_{\mathbb{R}^3}$. Therefore, $C(L) = C$ for some constant $C > 0$, and $\mathcal{H}(L) = \mathsf{H}$ for some constant $\mathsf{H} > 0$. Note that $C(L) = C$ describes a sphere, and $\mathcal{H}(L) = \mathsf{H}$ describes an ellipsoid, so the rigid body evolves on the intersection of a sphere and an ellipsoid.

Finally, we remark that there is a Hamiltonian formulation of symmetries and Noether's theorem. This leads to the concept of *momentum maps* as well as a Hamiltonian version of Noether's theorem, which for us will be important in the context of numerical integration of Poisson systems. To this end, let H be a Lie group that acts on a Poisson manifold $(M, \{\cdot,\cdot\})$. The action of H on M is said to be *canonical* if

$$\{F \circ \Phi_h, G \circ \Phi_h\} = \{F, G\} \ \circ \Phi_h \tag{2.17}$$

for all $F, G \in C^\infty(M)$ and all $h \in H$. A $\xi \in \mathfrak{h}$ generates a path in H by $\gamma(t) = \exp(t\xi)$. The *infinitesimal generator* of the action of H on M is the vector field given by

$$v^\xi(q) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi_{\exp(t\xi)}(q),$$

where $q \in M$. For $\xi, \eta \in \mathfrak{h}$, the Lie bracket of the infinitesimal generators is given by

$$[v^\xi, v^\eta] = v^{[\eta,\xi]},$$

meaning that $\xi \mapsto v^\xi$ is a Lie algebra anti-homomorphism from $\mathfrak{h}$ to the Lie algebra of vector fields on M. The infinitesimal version of Equation (2.17) is

$$v^\xi\{F, G\} = \{v^\xi F, G\} + \{F, v^\xi G\},$$

for all $F, G \in C^\infty(M)$.

Assuming that v^ξ is a Hamiltonian vector field, i.e., $v^\xi = X_{J(\xi)}$ for some $J\colon \mathfrak{h} \to C^\infty(M)$, then the map $\mu\colon M \to \mathfrak{h}^*$ given by

$$\langle \mu(q), \xi \rangle = J(\xi)(q)$$

is called the *momentum map*. These maps allow us to formulate a Hamiltonian version of Noether's theorem (Marsden and Ratiu, 1999, Theorem 11.4.1).

Theorem 2 (Noether's theorem, Hamilton version)**.** *Let H be a Lie group acting canonically on a Poisson manifold $(M, \{\cdot,\cdot\})$ with momentum map $\mu\colon M \to \mathfrak{h}^*$. If the Hamiltonian $\mathcal{H} \in C^\infty(M)$ is invariant under the action of H, that is, $\mathcal{H} \circ \Phi_h = \mathcal{H}$ for all $h \in H$, then the momentum map μ is a conserved quantity. This means that μ is preserved under the flow of the Hamiltonian vector field of $\mathcal{H}$.*

It is illustrative to see how this works for a concrete example.

Example 2.15: Particle in potential part IV

Recall the Hamiltonian formulation of the particle in a potential V from Example 2.6. The Hamiltonian was given by $\mathcal{H}(q,p) = \frac{1}{2}p^2 + V(q)$, where $(q,p) \in T^*\mathbb{R}^3$. Consider an arbitrary matrix Lie group H acting on $\mathbb{R}^3$ by $\Phi_A x = Ax$. Just as in Example 2.9, we shall see that if V is invariant under the action of H, then there is a conserved quantity. As an example, if $H = \mathrm{SO}(3)$, then the angular momentum is preserved. We must first see how $\mathrm{SO}(3)$ acts on $T^*\mathbb{R}^3$, i.e., the cotangent version of the tangent lifted action. As in Example 2.12 one sees that

$$T\Phi_A(x,v) = (Ax, Av),$$

where $(x,v) \in T\mathbb{R}^3$. We denote the Euclidean inner product between two vectors $x, y \in \mathbb{R}^3$ by $x \cdot y$. To compute the *cotangent lifted action* $T^*\Phi_A\colon T^*\mathbb{R}^3 \mapsto T^*\mathbb{R}^3$, we take, for an arbitrary $A \in \mathrm{SO}(3)$, a co-vector $p \in T_x^*\mathbb{R}^3$ and a vector $v \in T_{Ax}\mathbb{R}^3$, and simply compute (keeping track of the base points for clarity),

$$\langle T^*\Phi_A(x,p), (Ax,v)\rangle = \langle (x,p), T\Phi_{A^{-1}}(Ax,v)\rangle = p \cdot A^{-1}v = \langle A^{-\top}p, v\rangle,$$

where the pairing between $A^{-1}p$ and v makes sense only if $A^{-\top}p \in T_{Ax}^*\mathbb{R}^3$, so the cotangent lifted action is just $T^*\Phi_A(x,p) = (Ax, A^{-\top}p)$. Note that since $H = \mathrm{SO}(3)$, then $A^{-\top} = A$, so the cotangent lifted action is just $T^*\Phi_A(x,p) = (Ax, Ap)$. The infinitesimal generator of the action is given by taking the curve $A(t) = I + t\xi + \mathcal{O}(t^2)$ and

differentiating at $t = 0$, so that

$$v^\xi(x,p) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (A(t)x, A^{-\top}(t)p) = (\xi x, -\xi^\top p).$$

Now, $v^\xi(x,p)$ should be a Hamiltonian vector field for some function $J(\xi)$, meaning that it should be of the form

$$v^\xi = X_{J(\xi)} = \begin{pmatrix} \nabla_p J(\xi) \\ -\nabla_x J(\xi) \end{pmatrix}.$$

Thus, we have that $\nabla_p J(\xi) = \xi x$, so $J(\xi) = \xi x \cdot p + C(x)$ for some function $C(x)$, and $-\nabla_x J(\xi) = -\xi^\top p$, so $\nabla_x J(\xi) = \xi^\top p \cdot x + D(p)$ for some function $D(p)$. However, as $\xi^\top p \cdot x = \xi x \cdot p$, we see that $D(p) = C(x) = 0$, so $J(\xi) = \xi x \cdot p$. Finally, using the Frobenius inner product $\langle A, B\rangle = \mathrm{Tr}(A^\top B)$ to identify $\mathfrak{g}$ with its dual, the momentum map is given by

$$\begin{aligned}\langle \mu(x,p), \xi\rangle &= \mathrm{Tr}(\mu(x,p)\xi^\top) = J(\xi)(x,p) \\ &= \xi x \cdot p = \mathrm{Tr}(x^\top \xi^\top p) = \mathrm{Tr}(p x^\top \xi),\end{aligned}$$

where in the last step we used the transpose invariance of the trace. This means that the momentum map is given by the projection of $px^\top$ onto $\mathfrak{h}$ induced by the Frobenius inner product.
If V is invariant under the action of H, then by Noether's theorem, the projection of $px^\top$ is a conserved quantity. In the case of $H = \mathrm{SO}(3)$, the projection is simply $(px^\top - xp^\top)/2$, and applying the mapping (2.11), we identify this with $p \times x$, the angular momentum.

2.5 The Arnold Approach: Geodesics on Lie groups

Let us now consider a large class of mechanical systems, namely, those whose motion can be understood as *geodesics* on Lie groups. A geodesic is a generalization of straight lines (the shortest path in flat spaces) to manifolds. To define geodesics, we first need to introduce a way to measure length on a manifold. To this end, we introduce *Riemannian metrics*. A Riemannian metric on a manifold M is a smoothly varying inner product on the tangent space at each point. That is, for each $x \in M$, we have a positive definite symmetric bilinear form $g_x : T_xM \times T_xM \to \mathbb{R}$. All smooth manifolds admit a Riemannian metric, meaning that all smooth manifolds

can be made into *Riemannian manifolds*. The Riemannian structure allows for the definition of concepts such as length, angles and curvature. In local coordinates $(x_1, \dots, x_n)$, the Riemannian metric is given by a symmetric positive definite matrix $[g] = (g_{ij})$ where $g_{ij} = g(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j})$.

Given a vector $v \in T_x M$, its *length* is $|v|_g = \sqrt{g_x(v, v)}$. Further, the length of a smooth curve $\gamma\colon [a, b] \to M$ is defined as

$$\operatorname{len}(\gamma) = \int_a^b |\dot{\gamma}(t)|_g \, \mathrm{d}t.$$

The distance between two points $x, y \in M$ is defined as the infimum of the lengths of all curves that start at x and ends at y. A geodesic is a curve that minimizes the *energy functional* (Lee, 2018, Chapter 6)

$$E(\gamma) = \int_a^b |\dot{\gamma}(t)|_g^2 \, \mathrm{d}t.$$

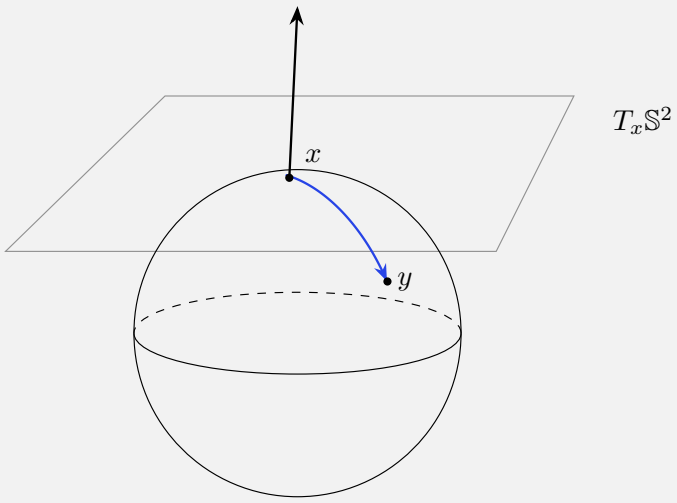
Let us illustrate these concepts by revisiting the sphere.

Example 2.16: The sphere as a Riemannian manifold

Consider the sphere. To describe its tangent space, we use the local coordinates $(\theta, \phi) \in [0, 2\pi) \times [0, \pi]$. Seeing tangent vectors as derivations, $T_x \mathbb{S}^2$ is the span of $\frac{\partial}{\partial \theta}$ and $\frac{\partial}{\partial \phi}$. Thus, to specify a Riemannian metric on $\mathbb{S}^2$, we need to give the values of the inner products of these basis vectors. On the sphere, the usual metric is the *round metric*

$$[g] = \begin{bmatrix} 1 & 0 \\ 0 & \sin^2 \theta \end{bmatrix}.$$

The round metric on $\mathbb{S}^2$ is the metric *induced* by the immersion of $\mathbb{S}^2$ into $\mathbb{R}^3$. Below, we see this illustrated, with a geodesic between two points on the sphere. Image from (Jansson, 2022).



Riemannian metrics also allow us to define the *gradient* of a function. If $f : M \to \mathbb{R}$ is a smooth function, the gradient $\nabla f \in \mathfrak{X}(M)$ of f at $x \in M$ is the vector $\nabla f(x) \in T_x M$ that satisfies

$$g_x(\nabla f(x), v) = \mathrm{d}f_x(v)$$

for all $v \in T_x M$. Note the similarity between this definition of the gradient and the definition of Hamiltonian vector fields in Equation (2.10).

To understand *acceleration* on a manifold, we must be able to generalize the concept of directional derivatives of vector fields to manifolds. To do this, we have to compare vector fields in different tangent spaces. This is done by *connections*. A connection is formally is a map $\nabla_{(\cdot)}\cdot \colon \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ that satisfies

- $C^\infty(M)$*-linearity in the first argument*: $\nabla_{fX+hY} Z = f\nabla_X Z + h\nabla_Y Z$,
- *linearity in the second argument*: $\nabla_X(Y + Z) = \nabla_X Y + \nabla_X Z$ and
- *the Leibniz rule in the second argument*: $\nabla_X(fY) = X(f)Y + f\nabla_X Y$,

for all $X, Y, Z \in \mathfrak{X}(M)$ and $f, h \in C^\infty(M)$. In other words, a covariant derivative is $C^\infty(M)$-linear in the first argument and is a derivation in the second argument. We call $\nabla_X Y$ the *covariant derivative* of Y in the direction of X. A connection is *symmetric* if $\nabla_X Y - \nabla_Y X = [X, Y]$ for all $X, Y \in \mathfrak{X}(M)$.

Further, connections act on tensor fields such as Riemannian metrics, namely by

$$\nabla_X \left(g(Y, Z)\right) = X(g(Y, Z)) - g(\nabla_X Y, Z) - g(Y, \nabla_X Z),$$

where $X, Y, Z \in \mathfrak{X}(M)$ and $g(X, Y)$ is the smooth function that sends $x \in M$ to $g_x(X(x), Y(x))$. If, for all $X \in \mathfrak{X}(M)$, $\nabla_X g = 0$, then ∇ is said to be *compatible* with g.

Connections may be used to define geodesics without the need for a Riemannian metric. As all manifolds admit a connection we can always have a notion of geodesics (Lee, 2018, Proposition 4.12), but there may be many connections on a manifold. If, however, we are given a Riemannian metric, there is a symmetric connection that is compatible with the metric known as the *Levi–Civita connection*. In the following, whenever a connection is mentioned, it is assumed to be the Levi–Civita connection. Moreover, we shall simplify notation somewhat and write $g(X, Y)$ as $X \cdot Y$.

Equipped with Riemannian geometry, we return to mechanics. Let H be a Lie group. Consider a mechanical system with configuration manifold H and a left-invariant Lagrangian $\mathcal{L} \colon TH \to \mathbb{R}$, i.e., for any $h' \in H$ and $(h, \dot{h})$ in TH, we have that

$$\mathcal{L}(L_{h'^{-1}} h, L_{h'^{-1}} \dot{h}) = \mathcal{L}(h, \dot{h}).$$

Note that we slightly abuse notation and denote left translation and its tangent lift with the same symbol. A left-invariant Lagrangian is determined by its behavior at only one point, say the identity $e \in H$. Therefore, it is sufficient to provide only a mapping $\ell\colon \mathfrak{h} \to \mathbb{R}$ to specify a left-invariant Lagrangian on H. Indeed, it holds that

$$\mathcal{L}(h, \dot{h}) = \mathcal{L}(L_{h^{-1}} h, L_{h^{-1}} \dot{h}) = \mathcal{L}(e, L_{h^{-1}} \dot{h}) \coloneqq \ell(L_{h^{-1}} \dot{h}). \tag{2.18}$$

The following theorem, due to Poincaré (1901), gives the equations of motion of a mechanical system described by a left-invariant Lagrangian on a Lie group.

Theorem 3. *Let the Lagrangian $\mathcal{L}$ be as in Equation* (2.18) *on a Lie group H Consider the extremizer $\gamma\colon [0,1] \to H$ of*

$$\int_0^1 \mathcal{L}(\gamma, \dot{\gamma})\, \mathrm{d}t.$$

The curve γ determines a curve on $\mathfrak{h}$ by $v(t) = L_{\gamma(t)^{-1}} \dot{\gamma}(t)$. It satisfies the Euler–Arnold equation

$$\dot{m} - \operatorname{ad}_v^* m = 0, \quad m = \frac{\delta \ell}{\delta v} \in \mathfrak{h}^*. \tag{2.19}$$

See (Modin, 2019) for a proof and further details. An important class of left-invariant Lagrangians are the quadratic Lagrangians, which are determined by an inner product on the Lie algebra $\mathfrak{h}$. To this end, let A be a positive-definite, symmetric linear operator $A : \mathfrak{h} \to \mathfrak{h}^*$, so that the inner product is given by $\langle v, w \rangle = Av(w)$. The inner product gives rise to a left-invariant Lagrangian $\mathcal{L}\colon TH \to \mathbb{R}$ by

$$\mathcal{L}(h, \dot{h}) = \frac{1}{2} \langle L_{h^{-1}} \dot{h}, L_{h^{-1}} \dot{h} \rangle = \ell(L_{h^{-1}} \dot{h}), \tag{2.20}$$

where $h \in H$, $\dot{h} \in T_h H$ and $\ell\colon \mathfrak{h} \to \mathbb{R}$ is the function $\ell(v) = \langle v, v \rangle$. The Lagrangian (2.20) determines a Riemannian metric, so that the Euler–Arnold equations (2.19) can be interpreted as geodesics on the Lie group H.

As an example of a mechanical system of this geodesic type, we consider again the rigid body in $\mathbb{R}^3$, as in (Marsden and Ratiu, 1999, Chapter 15).

Example 2.17: The rigid body part II

Recall the rigid body from Example 2.14. Just as it has a Hamiltonian formulation, it also has a Lagrangian formulation. Consider a rigid body in $\mathbb{R}^3$ with a fixed point at the origin, with a reference configuration given by a compact domain $B \subset \mathbb{R}^3$. Any configuration $\tilde{B}$ of

a rigid body is of the form $\tilde{B} = AB$, where $A \in \mathrm{SO}(3)$ is a rotation matrix. Thus, the configuration manifold is the Lie group $\mathrm{SO}(3)$ and the motion of a rigid body is a curve $A\colon [0,1] \to \mathrm{SO}(3)$. The velocity of the curve is $\dot{A}(t) \in T_{A(t)}\,\mathrm{SO}(3)$. By left translation, we obtain the angular velocity $\omega(t) = A(t)^{-1}\dot{A}(t) \in \mathfrak{so}(3) \cong \mathbb{R}^3$.
The velocity at $x \in B$ is given by $\frac{\mathrm{d}}{\mathrm{d}t}Ax = \dot{A}x$. Under the assumption that there are no external forces acting on the body, the Lagrangian is

$$\mathcal{L}(A, \dot{A}) = \frac{1}{2}\int_B \|\dot{A}x\|^2_{\mathbb{R}^3}\,\mathrm{d}x,$$

that is, the kinetic energy of the body. We note immediately that $\mathcal{L}$ is quadratic and positive definite, so it is a Riemannian metric on $\mathrm{SO}(3)$. Furthermore, it is left-invariant, as $\mathcal{L}(XA, X\dot{A}) = \mathcal{L}(A, \dot{A})$ for all $X \in \mathrm{SO}(3)$. Thus, we can reduce the problem to the Lie algebra $\mathfrak{so}(3) \cong \mathbb{R}^3$. However, in this case, we can nicely reason explicitly about how the reduced Lagrangian arises.
First, since $A^{-1}\dot{A} \in \mathfrak{so}(3)$, action of $A^{-1}\dot{A}$ is given by $\omega \times x$ where ω is the natural identification of $A^{-1}\dot{A} = A^\top \dot{A}$ with a vector in $\mathbb{R}^3$, as described by Equation (2.11). By the left-invariance of the metric, the Lagrangian can be written as

$$\mathcal{L}(A, \dot{A}) = \frac{1}{2}\int_B (A^\top \dot{A}x)^\top (A^\top \dot{A}x)\,\mathrm{d}x = \frac{1}{2}\int_B \|\omega \times x\|^2_{\mathbb{R}^3}\,\mathrm{d}x = \frac{1}{2}\langle \omega, \omega\rangle_B,$$

where $\langle \cdot, \cdot \rangle_B$ is the inner product on $\mathfrak{so}(3) \cong \mathbb{R}^3$ given by

$$\langle v, w\rangle_B = \int_B (v \times \hat{x})^\top (w \times \hat{x})\,\mathrm{d}\hat{x}.$$

Thus, as we are on $\mathbb{R}^3$, we can define a mapping $I\colon \mathfrak{so}(3) \cong \mathbb{R}^3 \to \mathfrak{so}(3)^* \cong \mathbb{R}^3$ by $w^\top I v = \langle v, w\rangle_B$. I is a positive definite and symmetric matrix, which in fact coincides with the moment of inertia tensor of the rigid body (Marsden and Ratiu, 1999, Section 15.3). We see that the reduced Lagrangian is just

$$\ell(\omega) = \frac{1}{2}\omega^\top I \omega.$$

In the context of Theorem 3, the operator A is I, and the Euler–Arnold equation (2.19) becomes,

$$\dot{L} = -\operatorname{ad}^*_\omega L = L \times \omega,$$

since by Example 2.11, the coadjoint action of $\mathfrak{so}(3)$ on its dual is just $\operatorname{ad}^*_\omega L = \omega \times L$. We have recovered the rigid body equations of motion from Example 2.14. And, moreover, we see that the motion of the rigid body follows a geodesic on $\mathrm{SO}(3)$.

Systems that follow the Euler–Arnold equation are known as *Euler–Arnold systems*, and examples include not only the rigid body, but also for instance hydrodynamical systems or equations in shape analysis, topics we return to in Chapter 4.

Chapter 3: Stochastic (Partial) Differential Equations

In this chapter, we briefly introduce the theory of SDEs and elliptic SPDEs. The goal is to describe the two remaining applications, namely *stochastic Lie–Poisson systems* on a class of Lie algebras, as well as *elliptic stochastic partial differential equations on surfaces*. While stochastic Lie–Poisson systems are decidedly mechanical in nature, elliptic SPDEs arise in a different context. For the purpose of this thesis, their purpose is to produce structured spatial noise on surfaces, also called random fields, that can be used as input to other algorithms. For instance, in (Modin and Viviani, 2019a), random fields on the sphere are used to generate random initial conditions for the simulation of fluid systems. The noise generated by elliptic stochastic partial differential equations is different from the noise used to drive the stochastic Lie–Poisson systems, as the former is a noise on the spatial domain whereas the latter systems are driven by time-dependent scalar noise.

3.1 Stochastic Differential Equations

We start by introducing the notation and basic concepts of stochastic differential equations. For a more detailed introduction, see for instance (Karatzas and Shreve, 1998) and (Øksendal, 2003).

SDEs are used to describe the evolution of a finite-dimensional system that is influenced by random noise, and are examples of stochastic processes, i.e., collections of random variables defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, taking values in some measurable space $(E, \mathcal{E})$. A stochastic process is indexed by a parameter in some set T, which is often referred to as time, thus motivating the notation $X_t(\omega)$, where $t \in T$ and $\omega \in \Omega$. The dependence on ω is often suppressed, and we write X_t instead.

In the remainder of this section, we assume that $T = \mathbb{R}^+$ and that the process takes values in $\mathbb{R}$. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in T}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions, i.e., that $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets and the filtration is right-continuous. Then X is *adapted* to the filtration $(\mathcal{F}_t)_{t\in T}$ if

X_t is $\mathcal{F}_t$-measurable for all $t \in T$.

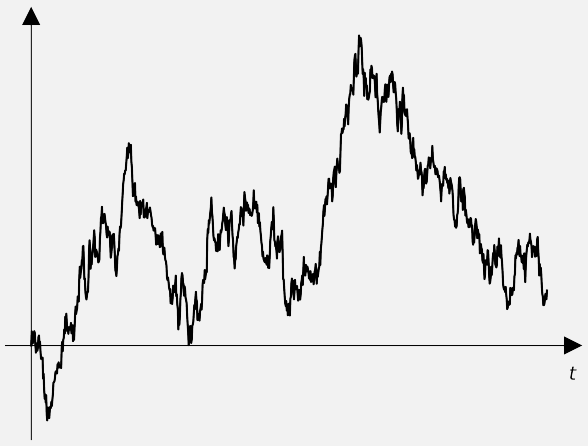
A famous, and for the remaining presentation very useful, example is Brownian motion.

Example 3.1: Brownian motion

Brownian motion $B = \{B_t, t \in \mathbb{R}^+\}$ is a stochastic process with the following properties:

1. $B_0 = 0$.
2. B_t has independent increments, that is to say, for $0 \leq t_1 < t_2 \leq t_3 < t_4$, the increments $B_{t_2} - B_{t_1}$ and $B_{t_4} - B_{t_3}$ are independent.
3. $B_t - B_s$ is normally distributed with mean zero and variance $t - s$.
4. B_t has almost surely continuous paths, that $t \mapsto B_t$ is continuous except on a set of measure zero.

A sample path of Brownian motions, also called a realization of the process, is depicted below.



In the remainder of this section, we assume that B is adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$.

Brownian motion can be used to build many other stochastic processes through the concept of stochastic calculus. First, consider the Itô integral

$$\int_0^t X_s \, \mathrm{d}B_s,$$

where X is a process adapted to $(\mathcal{F}_t)_{t\geq 0}$ and satisfies that $\mathbb{E}[\int_0^t X_s^2 \, \mathrm{d}s] < \infty$. We omit the formal construction of the Itô integral, which can be found in (Øksendal, 2003, Section 3.1). Morally, however, the Itô integral is a limit of

sums of the form

$$\sum_{i=0}^{n-1} X_{t_i}(B_{t_{i+1}} - B_{t_i}),$$

where $t_0 = 0 < t_1 < \ldots < t_n = t$ is a partition of $[0, t]$. This sum does not converge in the usual sense, but in mean-square sense. The proof of this is based on the Itô isometry, which allows us to convert the mean of the squared stochastic integral to an integral of the square of the integrand. Thus, the properties of the Lebesgue integral can be used. The Itô integral has many nice properties, such as linearity, but for this thesis, the important property is that the Itô integral has zero mean.

Using the Itô integral, we can define Itô processes, which are processes of the form

$$X_t = X_0 + \int_0^t \mu(s, X_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s)\,\mathrm{d}B_s, \tag{3.1}$$

where X_0 is deterministic, and $\mu\colon [0,\infty) \times \mathbb{R} \to \mathbb{R}$ and $\sigma\colon [0,\infty) \times \mathbb{R} \to \mathbb{R}$ are functions such that

$$|\mu(t,x)| + |\sigma(t,x)| \lesssim (1 + |x|),$$
$$|\mu(t,x) - \mu(t,y)| + |\sigma(t,x) - \sigma(t,y)| \lesssim |x - y|,$$

for all $t \in [0,\infty)$ and $x, y \in \mathbb{R}$ (Øksendal, 2003, Section 5.2). In other words, the coefficients satisfy a linear growth condition and are globally Lipschitz continuous in x uniformly in t. A shorthand for Equation (3.1) is the *stochastic differential equation*

$$\mathrm{d}X_t = \mu(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}B_t.$$

The ordinary rules of calculus do not hold for the Itô integral. In particular, the chain rule is replaced by Itô's formula, which states that for a function $f \in C^2(\mathbb{R})$,

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\,\mathrm{d}B_s + \frac{1}{2}\int_0^t f''(X_s)\,\mathrm{d}s.$$

In many applications, one values instead to have the usual chain rule, and this is achieved by the *Stratonovich integral*,

$$\int_0^t X_s \circ \mathrm{d}B_s,$$

and is defined (again, omitting its precise construction) as the limit of sums of the form

$$\sum_{i=0}^{n-1} \frac{1}{2}(X_{t_i} + X_{t_{i+1}})(B_{t_{i+1}} - B_{t_i}).$$

Again, these converge in mean-square sense. The Stratonovich chain rule is just

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) \circ \mathrm{d}B_s,$$

the usual chain rule. While the standard chain rule means that it is easier to work with the Stratonovich integral, it does not have as nice properties as the Itô integral. For instance, it does not have zero mean. In some cases, however, the Stratonovich integral is preferred. For example, while beyond the scope of this thesis, it is a natural question to ask how one can generalize stochastic analysis to manifolds. A description of stochastic calculus on manifolds is given in for instance (Hsu, 2002). In essence, the less awkward chain rule of the Stratonovich integral makes the generalization to manifolds much more straightforward.

3.2 Stochastic Lie–Poisson Systems

In Example 2.13, we considered a class of systems governed by equations of the form

$$\dot{\xi} = \mathrm{ad}^*_{d\mathcal{H}_\xi} \xi,$$

where ξ is an element of the dual of a Lie algebra $\mathfrak{h}$, and $\mathcal{H}$ is a Hamiltonian on $\mathfrak{h}^*$. We assume that $\mathfrak{h}$ is J-quadratic, meaning that there is a matrix J such that all $A \in \mathfrak{h}$ satisfy

$$A^*J + JA = 0$$

with $J^* = \pm J$ and $J^2 = cI_n$, for $c \in \mathbb{R}$ non-zero, where I_n denotes the $n \times n$ identity matrix. Note that if $A \in \mathfrak{h}$, then the conjugate transpose A^* satisfies

$$AJ + JA^* = \pm(JA^* + AJ)^* = 0.$$

Moreover, we assume that the group of $\mathfrak{h}$ is a compact, simply connected subgroup of $\mathrm{GL}(n, \mathbb{C})$.

As noted in (Modin and Viviani, 2019b) and in Paper III, this assumption (under an appropriate choice of duality pairing) implies in particular that the Lie–Poisson system is an isospectral system of the form

$$\dot{\xi} = [\nabla\mathcal{H}^*, \xi].$$

A way to make this system stochastic is to add *transport noise*, i.e., to set $\mathcal{H} = \mathcal{H}_0$ and introduce M noise Hamiltonians $\mathcal{H}_k \colon \mathfrak{h}^* \to \mathbb{R}$, $k = 1, \dots, M$, and consider the system

$$\mathrm{d}\xi = [\nabla\mathcal{H}_0^*, \xi]\,\mathrm{d}t + \sum_{k=1}^{M} [\nabla\mathcal{H}_k^*, \xi] \circ \mathrm{d}B_t^k,$$

where B_t^k are independent Brownian motions. This type of system is known as a *stochastic Lie–Poisson system*, and has been studied extensively in the literature, both from the perspective of numerical analysis but also for their use in the modelling of physical systems; see for instance the papers (Bréhier et al., 2023; Hong et al., 2021; Liao and Wang, 2005; Luesink et al., 2024; Holm, 2015). Their numerical integration is the focus of Paper III of this thesis.

Stochastic Lie–Poisson systems are interesting for many reasons. Firstly, it evolves on the same coadjoint orbits and has the same Casimirs as the deterministic system. Furthermore, its coefficients typically do not satisfy the usual Lipschitz conditions, so the system is not amenable to standard existence and uniqueness results. Instead, its geometric properties can be used to show that the system has a unique solution; see (Bréhier et al., 2023, Proposition 1).

Example 3.2: The rigid body part III: stochastic

Let us revisit the rigid body from Example 2.17, but now with added noise. Recall that the system evolves on (the dual of) the Lie algebra $\mathfrak{so}(3) \cong \mathbb{R}^3$, but that it remains constrained to the sphere with radius determined by the length of the angular momentum vector at the initial time. What would happen if we added noise to the system? If we add additive Itô noise, i.e., for some $\Sigma \in \mathbb{R}^{3\times 3}$ and $B_t = (B_t^1, B_t^2, B_t^3)$ independent Brownian motions, we get the system

$$\mathrm{d}L_t = L_t \times I^{-1} L_t \,\mathrm{d}t + \Sigma \,\mathrm{d}B_t.$$

In the deterministic case, the Casimir $\frac{1}{2} L_t^\top L_t = \frac{1}{2}\|L_t\|_{\mathbb{R}^3}^2$ is a preserved quantity that constrains the system to the sphere. By Itô's formula, we have

$$\begin{aligned}\mathrm{d}(\|L_t\|_{\mathbb{R}^3}^2) &= 2\langle L_t, L_t \times I^{-1} L_t\rangle_{\mathbb{R}^3}\,\mathrm{d}t + 2L_t^\top \Sigma\,\mathrm{d}B_t + \mathrm{Tr}(\Sigma\Sigma^\top)\,\mathrm{d}t \\ &= 2L_t^\top \Sigma\,\mathrm{d}B_t + \mathrm{Tr}(\Sigma\Sigma^\top)\,\mathrm{d}t,\end{aligned}$$

so in this case,

$$\mathbb{E}\,\|L_t\|_{\mathbb{R}^3}^2 = \|L_0\|_{\mathbb{R}^3}^2 + \mathrm{Tr}(\Sigma\Sigma^\top)t,$$

meaning that the Casimir is not preserved, and we drift off the sphere. If, however, we add Stratonovich transport noise, we get

$$\mathrm{d}L_t = L_t \times I^{-1} L_t \,\mathrm{d}t + L_t \times \nabla_{L_t}\mathcal{H}(L_t) \circ \mathrm{d}B_t,$$

and in this case, the Casimir is preserved. Indeed, the Stratonovich

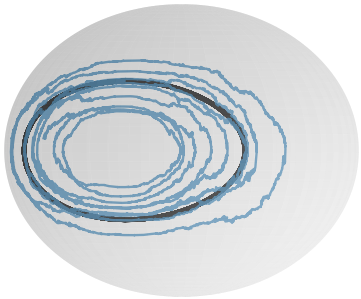


Figure 3.1: A realization of the stochastic rigid body system in blue, contrasted with the deterministic path in black.

chain rule gives that

$$\begin{aligned} \mathrm{d}(\|L_t\|^2_{\mathbb{R}^3}) = 2\langle L_t, L_t \times I^{-1} L_t \rangle_{\mathbb{R}^3} \, \mathrm{d}t \\ + 2\langle L_t, L_t \times \nabla_{L_t} \mathcal{H}(L_t) \rangle_{\mathbb{R}^3} \circ \mathrm{d}B_t = 0. \end{aligned}$$

A realization of the system is shown in Figure 3.1, computed using the method of Paper III. The system remains on the sphere, but due to the noise, the behavior is different from the deterministic case.

Moreover, to numerically integrate this class of systems in a manner that preserves interesting properties such as coadjoint orbits, several methods exist; see for instance the papers (Bréhier et al., 2023; Hong et al., 2021; Liao and Wang, 2005; Luesink et al., 2024).

3.3 Structured Noise on Surfaces

In Section 3.1 and Section 3.2, we saw stochastic processes (taking values in either $\mathbb{R}$ or in a Lie algebra $\mathfrak{h}$) driven by scalar Brownian motions. In the definition of a stochastic process, there is nothing that says that the index set T has to be one-dimensional. Indeed, a stochastic process can be indexed by any set, and in particular, it can be indexed by a spatial domain such as a manifold. In the literature, a stochastic process indexed by a spatial domain is known as a *random field* to distinguish it from other stochastic processes. The generation of random fields has been the main motivation of the work presented in Papers I and II of this thesis.

In the following, let M be a Riemannian manifold. Formally, we view a random field $\mathcal{Z}$ as a collection of random variables indexed by the points in the manifold, i.e., as a $\mathcal{F} \otimes \mathcal{B}(M)$-measurable function $\mathcal{Z} \colon \Omega \times M \to \mathbb{R}$.

Here $\mathcal{B}(M)$ is the Borel σ-algebra on M, i.e., the σ-algebra consisting of all sets that can be generated by open sets in M. Later in this thesis, we view random fields as random variables taking values in a suitable function space on M. To make the latter view precise, we need to introduce said function spaces, which is why we postpone the discussion to Section 3.4.

A random field is said to be *second-order* if, for all $x \in M$, the random variable $\mathcal{Z}(x)$ has finite second moment. The mean function of $\mathcal{Z}$ is given by

$$\mu(x) = \mathbb{E}[\mathcal{Z}(x)] = \int_\Omega \mathcal{Z}(x) \, \mathrm{d}\, \mathbb{P}\,.$$

If $\mu(x) = 0$ then $\mathcal{Z}$ is *centered*. The covariance function $C\colon M \times M \to \mathbb{R}$ is defined by

$$C(x, y) = \mathbb{E}[(\mathcal{Z}(x) - \mu(x))(\mathcal{Z}(y) - \mu(y))].$$

The covariance function is a measure of how much the random variables $\mathcal{Z}(x)$ and $\mathcal{Z}(y)$ are correlated.

Further, $\mathcal{Z}$ is said to be *Gaussian* if, for all $x_1, \ldots, x_n \in M$ and coefficients $a_1, \ldots, a_n \in \mathbb{R}$, the random variable $\sum_{i=1}^n a_i \mathcal{Z}(x_i)$ is Gaussian. A Gaussian field is completely determined by its mean and covariance functions.

Finally, if H is a Lie group and M is an H-homogeneous space (i.e., H acts transitively on M), then we say that $\mathcal{Z}$ is *isotropic* with respect to the action of H if, for all $k \in \mathbb{N}$, $x_1, \ldots, x_k \in M$, and $h \in H$, the multivariate random variable

$$(\mathcal{Z}(h \cdot x_1), \mathcal{Z}(h \cdot x_2), \ldots, \mathcal{Z}(h \cdot x_n))$$

has the same law as

$$(\mathcal{Z}(x_1), \mathcal{Z}(x_2), \ldots, \mathcal{Z}(x_n)).$$

As a shorthand, we say that $\mathcal{Z}$ is *H-isotropic*. A classical example is that of a *stationary Gaussian random field* on $\mathbb{R}^d$.

Example 3.3: Isotropy I: the Euclidean space

The Euclidean space of dimension d is a Lie group, and acts transitively on itself. A Gaussian random field $\mathcal{Z}$ (with finite second moment) on $\mathbb{R}^d$ is *stationary* if its distribution is invariant under translations, i.e., for all $k \in \mathbb{N}$ and $x_1, \ldots, x_k, y \in \mathbb{R}^d$,

$$(\mathcal{Z}(x_1), \ldots, \mathcal{Z}(x_k))$$

has the same law as

$$(\mathcal{Z}(x_1 + y), \ldots, \mathcal{Z}(x_k + y)).$$

This has two immediate consequences. Firstly,

$$\mu(x) = \mathbb{E}[\mathcal{Z}(x)] = \mathbb{E}[\mathcal{Z}(x+y)] = \mu(x+y),$$

so the mean function of $\mathcal{Z}$ is constant (meaning that we can assume that it is zero), and secondly, its covariance function satisfies

$$C(x_1, x_2) = \mathbb{E}[\mathcal{Z}(x_1)\mathcal{Z}(x_2)] = \mathbb{E}[\mathcal{Z}(x_1 - x_2)\mathcal{Z}(0)] = C(x_1 - x_2, 0),$$

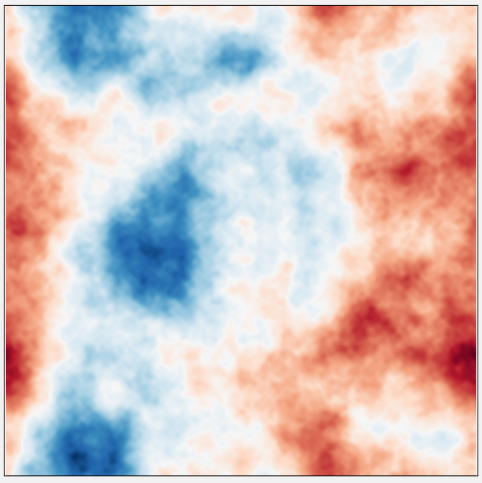
meaning that the covariance function only depends on the difference of the points.
Moreover, we can ask what happens if the random field is also isotropic with respect to the rotation group $\mathrm{SO}(d)$, i.e., we consider the group of translations and rotations, the special Euclidean group $\mathrm{SE}(d)$, that acts transitively on $\mathbb{R}^d$. An element of $\mathrm{SE}(d)$ is a pair (R, y), where $R \in \mathrm{SO}(d)$ and $y \in \mathbb{R}^d$, and the action is given by $(R, y) \cdot x = Rx + y$. Thus, that $\mathcal{Z}$ is $\mathrm{SE}(d)$-isotropic means, in addition to being invariant under translations, that

$$(\mathcal{Z}(Rx_1), \ldots, \mathcal{Z}(Rx_k))$$

has the same law as

$$(\mathcal{Z}(x_1), \ldots, \mathcal{Z}(x_k)),$$

for all $k \in \mathbb{N}$, $x_1, \ldots, x_k \in \mathbb{R}^d$, and $R \in \mathrm{SO}(d)$. An example of a field (on a periodic square) is shown below.



A field that is $\mathrm{SE}(d)$-isotropic is called an *isotropic stationary Gaussian random field* on $\mathbb{R}^d$, and in this case, the mean function is constant, and the covariance function depends only on the distance between the

points. Indeed, by first using the invariance under translations, we have that

$$C(x_1, x_2) = C(x_1 - x_2, 0),$$

and then using the invariance under rotations, we have that

$$C(x_1 - x_2, 0) = C(R(x_1 - x_2), 0)$$

meaning that C is constant on the sphere of radius $\|x_1 - x_2\|_{\mathbb{R}^d}$, implying that the covariance function only depends on the distance between the points. Thus, to specify an isotropic stationary Gaussian random field on $\mathbb{R}^d$, it is enough to specify a positive definite function depending only on a scalar, i.e., the distance.

It is clear that the properties of the random field depend on the appropriate choice of group. To drive home the point, consider the sphere $\mathbb{S}^2$.

Example 3.4: Isotropy II: the sphere

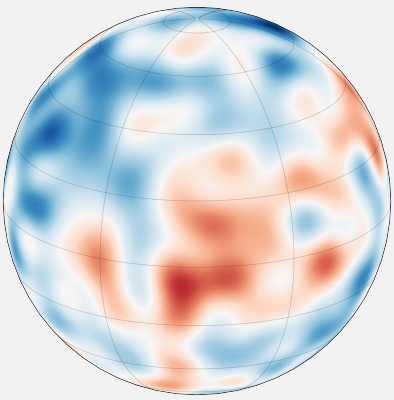
Consider the case when the Riemannian manifold M is the unit sphere in $\mathbb{R}^3$, i.e., $M = \mathbb{S}^2$. The sphere is an SO(3)-homogeneous space, so we say that a Gaussian random field $\mathcal{Z}$ on $\mathbb{S}^2$ is isotropic if, for all $k \in \mathbb{N}$, $x_1, \ldots, x_k \in \mathbb{S}^2$ and $T \in \mathrm{SO}(3)$, the multivariate random variable

$$(\mathcal{Z}(Tx_1), \mathcal{Z}(Tx_2), \ldots, \mathcal{Z}(Tx_n))$$

has the same law as

$$(\mathcal{Z}(x_1), \mathcal{Z}(x_2), \ldots, \mathcal{Z}(x_n)).$$

A sample of an isotropic Gaussian random field is illustrated below.



Note now that for any $x, y \in \mathbb{S}^2$, there is a unique $T \in \mathrm{SO}(3)$ such

that $Tx = y$. Thus, the mean function satisfies

$$\mu(x) = \mathbb{E}[\mathcal{Z}(x)] = \mathbb{E}[\mathcal{Z}(Tx)] = \mu(Tx) = \mu(y),$$

so it is constant. Further, by the invariance under rotations, we have that

$$C(x, y) = C(Tx, Ty)$$

for all $T \in \mathrm{SO}(3)$, so the covariance function can only depend on quantities that are invariant under rotations, in other words, on the distance between x and y.

In both cases, the appropriate choice of symmetry group leads to a constant mean function and a covariance function that only depends on the distance between the points.

Honing in on the Euclidean case, we can ask how to generate $\mathrm{SE}(d)$-isotropic Gaussian random fields on $\mathbb{R}^d$. A Gaussian random field on $\mathbb{R}^d$ is completely determined by its mean and covariance functions, and a popular choice of covariance function is the Matérn covariance function, defined as

$$C(\|h\|_{\mathbb{R}^d}) = \frac{2^{1-\nu}\sigma^2}{\Gamma(\nu)} K_\nu(\kappa\|h\|_{\mathbb{R}^d})(\kappa\|h\|_{\mathbb{R}^d})^\nu, \tag{3.2}$$

where $\nu > 0$ is a smoothness parameter, $\sigma^2 > 0$ is the variance, $\alpha > 0$ is a scale parameter, K_ν is the modified Bessel function of the second kind, and Γ is the Gamma function. The Matérn covariance function was originally introduced by the Swedish statistician Bertil Matérn, originally for forestry applications (Matérn, 1960), but has since found applications in several fields, from meteorology to machine learning (Guttorp and Gneiting, 2006; Handcock and Wallis, 1994; Porcu et al., 2024). Fields with Matérn covariance are known as *Whittle–Matérn fields*.

If we now want to generate a Whittle–Matérn field on the sphere, a natural first attempt is to replace the Euclidean distance with the geodesic distance on the sphere, i.e., to replace $\|h\|_{\mathbb{R}^3}$ with $d(x, y) = \arccos(\langle x, y\rangle_{\mathbb{R}^3})$ in Equation (3.2). This, however, does not always lead to a valid covariance function, as noted by Gneiting (2013). Therefore, to generate a Whittle–Matérn field on the sphere (or any other Riemannian manifold), we need to take a different approach, which we introduce in the next section.

3.4 Elliptic SPDEs on Manifolds

Whittle (1963) showed that random fields with Matérn covariance functions

are solutions to the stochastic partial differential equation

$$(\kappa^2 - \Delta)^\beta \mathcal{Z} = \mathcal{W},$$

where κ is a positive constant, Δ is the Laplace operator on $\mathbb{R}^d$, $\beta > d/4$ is a positive constant, and $\mathcal{W}$ is a white noise. The parameters β and σ^2 are related to the parameters of the Matérn covariance function (3.2) by

$$\nu = 2\beta - d/2, \quad \sigma^2 = \frac{\Gamma(\nu)}{\Gamma(2\beta)(4\pi)^{d/2}\kappa^{4\beta}}.$$

Following Lindgren et al. (2011), the equivalent SPDE on a compact manifold allows us to define Whittle–Matérn random fields on the manifold. It is immediately clear that some machinery is needed to do this. In this section, we touch on some necessary elements of the theory of partial differential equations. We refer to (Taylor, 2011) for an in-depth treatment of the subject.

Consider a compact, oriented Riemannian manifold (M, g), without boundary. The first step is to define the equivalent of the Laplace operator on M. This is the Laplace–Beltrami operator Δ_M, which is, analogously to the Laplace operator on $\mathbb{R}^d$, defined as the divergence of the gradient, i.e.,

$$\Delta_M f = \operatorname{div}(\nabla_M f),$$

where the divergence of a vector field v is the function $\operatorname{div} v$ given by

$$d(\iota_v \mu_g) = \operatorname{div} v \mu_g.$$

In local coordinates, the Laplace–Beltrami operator is given by

$$\Delta_M f = \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x_i} \left(\sqrt{|g|} g^{ij} \frac{\partial f}{\partial x_j} \right),$$

where $|g|$ is the determinant of the metric tensor and (g^{ij}) are the elements of the inverse of the metric tensor.

When working with elliptic stochastic partial differential equations, it is convenient to view the solutions as random variables taking values in Hilbert spaces known as *Sobolev spaces*. We denote by $L^2(M)$ the space of square-integrable functions on M, equipped with the inner product

$$\langle f, g \rangle_{L^2(M)} = \int_M f g \, \mu_g,$$

where μ_g is the volume form on M induced by the metric g. Sobolev spaces on M can be defined by *Bessel potentials*, i.e., operators of the form $\mathsf{B}_r = (I - \Delta_M)^{-r/2}$ The Sobolev space $H^r(M)$ with $r > 0$ is defined as

$$H^r(M) = \mathsf{B}_r L^2(M),$$

where the corresponding norm is given by $\|f\|_{H^r(M)} = \|\mathsf{B}_{-r} f\|_{L^2(M)}$. Sobolev spaces with $r < 0$ are defined as the spaces of distributions generated by

$$H^{-r}(M) = \left\{ u = \mathsf{B}_{-k} v, v \in H^{2k+r}(M) \right\},$$

where k is the smallest positive integer such that $2k + r > 0$. More details on Sobolev spaces on manifolds defined using Bessel potentials can be found in (Herrmann et al., 2018; Strichartz, 1983; Triebel, 1985).

A random field $\mathcal{Z}$ on M is an $L^2(M)$-valued random variable, i.e., as a $\mathcal{F} \otimes \mathcal{B}(L^2(M))$-measurable function $\mathcal{Z} \colon \Omega \to L^2(M)$. The generalization of Whittle–Matérn fields on $\mathbb{R}^d$ to M is given by the solution to the stochastic partial differential equation

$$(\kappa^2 - \Delta_M)^\beta \mathcal{Z} = \mathcal{W}. \tag{3.3}$$

Here $\mathcal{W}$ denotes a white noise on M, i.e., a zero-mean Gaussian random field understood as a random variable in $H^r(M)$, where $r < -\dim(M)/2$, such that for any $\phi, \psi \in L^2(M)$,

- $\langle \mathcal{W}, \phi \rangle_{L^2(M)}$ is Gaussian with mean zero, and
- $\mathbb{E}[\langle \mathcal{W}, \phi \rangle_{L^2(M)} \langle \mathcal{W}, \psi \rangle_{L^2(M)}] = \langle \phi, \psi \rangle_{L^2(M)}$.

By the spectral theorem (Taylor, 2011, Proposition 6.6.), there are eigenpairs (λ_i, ϕ_i) of the Laplace–Beltrami operator, i.e., $\Delta_M \phi_i = \lambda_i \phi_i$, such that the eigenfunctions form an orthonormal basis of $L^2(M)$. This allows for an easy definition of the operator $(\kappa^2 - \Delta_M)^\beta$ as

$$(\kappa^2 - \Delta_M)^\beta f = \sum_{i=1}^{\infty} (\kappa^2 - \lambda_i)^\beta \langle f, \phi_i \rangle_{L^2(M)} \phi_i.$$

The numerical solution of Equation (3.3) is the focus of Paper I of this thesis.

Equation (3.3) immediately allows for two generalizations. The first is to consider more general functions of the differential operator, i.e., to consider the equation

$$\mathcal{Z} = \zeta(\Delta_M) \mathcal{W}, \tag{3.4}$$

where $\zeta(\Delta_M)$ is defined by

$$\zeta(\Delta_M) f = \sum_{i=1}^{\infty} \zeta(\lambda_i) \langle f, \phi_i \rangle_{L^2(M)} \phi_i.$$

Note that the function ζ must be chosen such that the series converges in $L^2(M)$, and that since ζ appears also on the right-hand side, we do not require that ζ is invertible. The class of random fields defined by Equation (3.4) was

studied in (Lang and Pereira, 2023). A second generalization is to consider fields defined by more general differential operators, i.e., to consider the equation

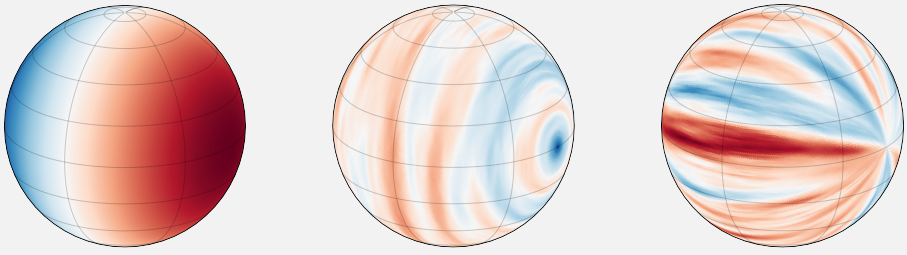
$$\mathcal{Z} = \zeta(\mathcal{L})\mathcal{W}, \tag{3.5}$$

where $\mathcal{L}$ is an elliptic self-adjoint differential operator on M; see for instance (Dziuk and Elliott, 2013, Equation 3.4) for an example on surfaces. Using Equation (3.5), we can obtain random fields that behave differently in different parts of the manifold. Fields defined by Equation (3.5) are studied in Paper II of this thesis.

By varying the coefficients of the differential operator, one can model different types of random fields and obtain a large class of random fields on manifolds.

Example 3.5: Spatially varying fields on the sphere

To illustrate the class of fields that can be generated by Equation (3.5), consider the sphere $\mathbb{S}^2$, since it is a domain where explicit computations are easily available. We consider the differential operator $\mathcal{L} = \operatorname{div}(\mathcal{D}\nabla_{\mathbb{S}^2}) + V$, where $\mathcal{D}(x)\colon T_x\mathbb{S}^2 \to T_x\mathbb{S}^2$ is a positive definite self-adjoint operator, and $V\colon \mathbb{S}^2 \to \mathbb{R}^+$ is a potential function. The potential function can be used for instance to scale the field in different parts of the domain, but in this example, we set $V = 10$ and focus on the operator $\mathcal{D}$ that models anisotropy in the field. As an example, we select $\mathcal{D}(x)v = \alpha_1(\nabla f \cdot v)\nabla f + \alpha_2(X_f \cdot v)X_f$, where $f\colon \mathbb{S}^2 \to \mathbb{R}$ is a smooth function, X_f is the Hamiltonian vector field of f, and $\alpha_1, \alpha_2 > 0$ are constants. If $\alpha_2 \ll \alpha_1$, the field is elongated orthogonally to the level sets of f, and if $\alpha_1 \ll \alpha_2$, the field is elongated along the level sets of f. We select $f(x) = x_2$, where $x = (x_1, x_2, x_3)$, and $\zeta(z) = z^{-1}$. We consider both $(\alpha_1 = 10, \alpha_2 = 1/10)$ (elongated orthogonally to the level sets of f), and $(\alpha_1 = 1/10, \alpha_2 = 10)$ (elongated along the level sets of f). Below we show one realization of each case, generated using the method of Paper II along with the function f.



Function | Tangential field | Orthogonal field

3.5 Numerical Methods for Elliptic SPDEs on Surfaces

We now briefly describe two approaches to the numerical solution of elliptic SPDEs on surfaces, namely *spectral methods* and *finite element methods*. Finite element methods are the main computational tools in Papers I and II of this thesis. In the following, let M be a compact, oriented surface embedded without boundary in $\mathbb{R}^3$. Spectral methods are based on a truncation of the Fourier series expansion of the solution in terms of an eigenbasis of the differential operator $\mathcal{L}$. Finite element methods, on the other hand, approximate the solution to PDEs using linear combinations of functions in some finite-dimensional function space defined on a triangulation of the surface. While both spectral and finite element methods use linear combinations of basis functions, a key distinction lies in their support: spectral methods are *global*, while finite elements are *local*.

To be more concrete, consider the deterministic elliptic PDE

$$\mathcal{L}u = f,$$

where $\mathcal{L}$ is an elliptic differential operator on M as defined in Section 3.4 and f is a function on M. FEM methods rely on the weak formulation of the PDE, that is, to find $u \in H^1(M)$ such that

$$a(u, v) = \langle f, v \rangle_{L^2(M)}, \tag{3.6}$$

for all $v \in H^1(M)$, where $a : H^1(M) \times H^1(M) \to \mathbb{R}$ is the bilinear form associated with the $\mathcal{L}$.

To apply FEM on surfaces, we approximate M by a polyhedral surface M_h obtained via triangulation. Specifically, we define

$$M_h = \bigcup_{T_j \in \mathcal{T}} T_j,$$

where $\mathcal{T}$ is a collection of non-degenerate triangles with vertices on M. We assume the that for any two triangles $T_i, T_j \in \mathcal{T}$, the intersection $T_i \cap T_j$ is either the empty set, a common vertex, or a common edge.

If M is a Euclidean domain, that is, flat, one has that $M_h = M$, and we can proceed by defining a finite-dimensional subspace of $H^1(M)$ consisting of piecewise linear functions on the mesh. The weak formulation (3.6) then reduces to a linear system of equations on this subspace. For more details on FEM in the Euclidean case, see (Brenner and Scott, 2008; Larsson and Thomée, 2008; Strang and Fix, 2008).

In the more general case when M is a surface, the situation is more complicated since $M_h \neq M$; see Figure 3.2.

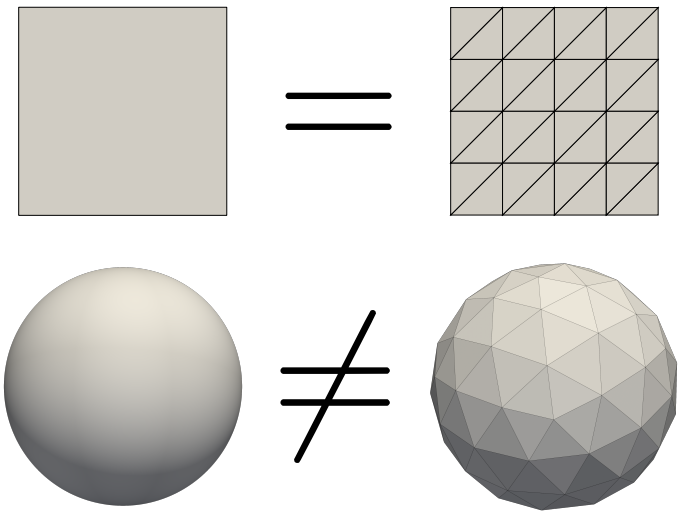


Figure 3.2: Illustration of the difficulty of working with FEM on surfaces. A Euclidean domain can be discretized by flat triangles, but a curved domain cannot.

Before carrying on, let us briefly discuss spectral methods. These have benefits over finite element methods, such as they:

(i) may converge faster,

(ii) require no discretization of the spatial domain, and

(iii) are easier to implement if an explicit eigenbasis is known.

Moreover, spectral methods are in a sense intrinsic to the surface as the eigenbasis is determined by the geometry of the surface and not by its embedding in $\mathbb{R}^3$. In contrast, the FEM methods considered in this thesis are extrinsic.

However, spectral methods also have their drawbacks. As they are global methods, they can be a poor choice for problems with strong local features. More critically, spectral methods require a tractable eigenbasis, which is only available in special cases. To illustrate this, consider Equation (3.4). If an eigenbasis (λ_i, ϕ_i) to Δ_M is explicitly known, then the solution $\mathcal{Z}$ can be written as

$$\mathcal{Z} = \zeta(\mathcal{L})\mathcal{W} = \sum_{i=1}^{\infty} \zeta(\lambda_i) W_i \phi_i,$$

where W_i are independent standard Gaussian random variables. Truncating

the series after N terms gives the approximation

$$\mathcal{Z}_N = \sum_{i=1}^{N} \zeta(\lambda_i) W_i \phi_i.$$

Under the assumption that $\zeta(\lambda)$ decays as $\lambda^{-\alpha}$ for some $\alpha > 0$, the mean-square error of the approximation satisfies

$$\mathbb{E}[\|\mathcal{Z} - \mathcal{Z}_N\|^2_{L^2(M)}] = \sum_{i=N+1}^{\infty} \zeta(\lambda_i)^2 \leq \sum_{i=N+1}^{\infty} \lambda_i^{-2\alpha}.$$

By Weyl's law, the eigenvalues grow as $\lambda_i \sim i^{2/\dim(M)}$ (Jost, 2011, Equation 3.2.24), so under the assumption that $\alpha > \dim(M)/4$, the error decays as

$$\begin{aligned}\mathbb{E}[\|\mathcal{Z} - \mathcal{Z}_N\|^2_{L^2(M)}] &\leq \sum_{i=N+1}^{\infty} i^{-4\alpha/\dim(M)} \leq \int_N^{\infty} x^{-4\alpha/\dim(M)} \,\mathrm{d}x \\ &= \frac{1}{-4\alpha/\dim(M) + 1} N^{-4\alpha/\dim(M)+1}.\end{aligned}$$

However, the eigenbasis is not known in general. A notable exception is the Laplace–Beltrami operator on the sphere $\mathbb{S}^2$, whose eigenbasis is given by the spherical harmonic functions. A conceivable remedy in the general case would be to use finite element methods to approximate the eigenbasis, but this is not feasible in practice, as one cannot hope to obtain a sufficiently good approximation of the eigenbasis (Boffi, 2010). Instead, one can use the finite element method to approximate the solution directly. This is the motivation of Paper II of this thesis.

As in the Euclidean setting, the finite element method on surfaces approximates solutions to partial differential equations by solving the weak formulation (3.6) on a finite-dimensional function space. However, on surfaces, the finite element space is a subspace of $H^1(M_h)$, not of $H^1(M)$. This complicates the error analysis, as functions in $H^1(M)$ and $H^1(M_h)$ are not directly comparable.

One solution to this problem was proposed by Dziuk (1988), who introduced the *surface finite element method (SFEM)*. The method solves the issue by firstly selecting an appropriate coordinate system to work in, and by then providing a suitable bijection between the surface and its discretization that allows one to define the *lift* of functions on the mesh to the surface. That is, a function u_h defined on M_h can be mapped to a function u_h^ℓ defined on M. This allows the error analysis to be carried out. Nonetheless, the method is not without drawbacks. The discretization of the surface introduces additional error terms that are not present in the Euclidean case. Moreover, a proper definition of SFEM requires a significant amount of machinery. For

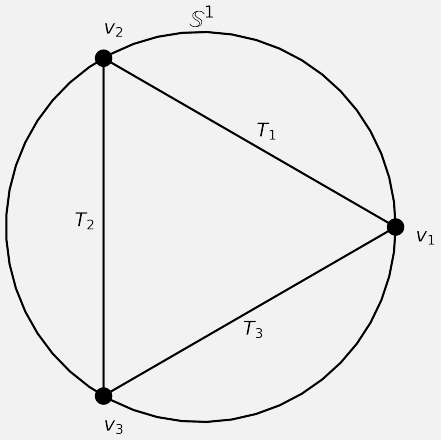
a thorough treatment of the subject, we refer to (Dziuk and Elliott, 2013). In the interest of clarity, we instead consider a brief 1-dimensional example.

Example 3.6: SFEM on the circle

Let us consider the Helmholtz equation on the circle $\mathbb{S}^1$,

$$u(\theta) - \Delta_{\mathbb{S}^1} u(\theta) = f(\theta), \tag{3.7}$$

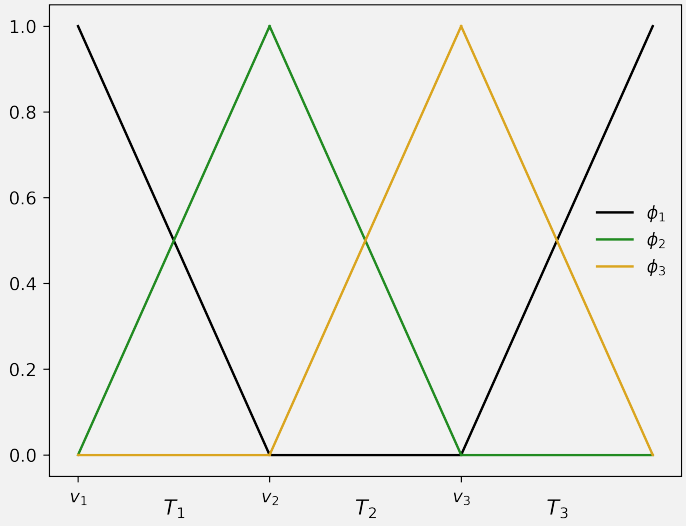
where θ is the angle on the circle, and $f(\theta) = \cos(\theta)$. The solution to Equation (3.7) is $u(\theta) = \cos(\theta)/2$. We study the approximation of the equation on a discretization of the circle, that is, a triangulation by straight line segments. We consider the case of three mesh points $\theta_1 = 0$, $\theta_2 = 2\pi/3$ and $\theta_3 = 4\pi/3$, i.e., the points with coordinates $v_1 = (1, 0)$, $v_2 = (-1/2, \sqrt{3}/2)$ and $v_3 = (-1/2, -\sqrt{3}/2)$. The length of each line segment is $\sqrt{3}$, which also is the mesh size. Denote the line segments T_1, T_2 and T_3. The triangulation of the circle is thus the union of the line segments, that is, the triangle with vertices at the mesh points, and is denoted by $\mathbb{S}^1_h$ (see below).



On $\mathbb{S}^1_h$, we define the finite element space S_h as the set of piecewise linear functions on each T_i,

$$S_h = \{\phi_h \in C(\mathbb{S}^1_h) | \phi_h|_{T_i} \in P^1(T_i) \text{ for each } T_i, i = 1, 2, 3\},$$

where $P^1(T_i)$ is the space of linear functions on the triangle T_i. The space S_h is spanned by the basis functions ϕ_1, ϕ_2, and ϕ_3: these are defined by $\phi_i(v_j) = \delta_{ij}$ for $i, j = 1, 2, 3$.
Every function in S_h can be written as a linear combination of the basis functions by $\phi_h = \sum_{i=1}^3 \beta_i \phi_i$ for some $\beta_i \in \mathbb{R}$. See below for an illustration of the basis functions.



The weak formulation of Equation (3.7) is to find $u \in H^1(\mathbb{S}^1)$ such that

$$\int_{\mathbb{S}^1} (uv + \nabla_{\mathbb{S}^1} u \cdot \nabla_{\mathbb{S}^1} v) \, \mathrm{d}x = \int_{\mathbb{S}^1} fv \, \mathrm{d}x,$$

for all $v \in H^1(\mathbb{S}^1)$. To write down the weak formulation on the mesh, we need to determine how to compute the gradient on $\mathbb{S}^1_h$. In the context of surface finite elements, this is done by computing the gradient in the embedded space (in this case $\mathbb{R}^2$) and then projecting it back to the mesh, i.e., $\nabla_{\mathbb{S}^1_h} = P_h \nabla$, where the ij-th component of P_h is given by $(P_h)_{ij} = \delta_{ij} - (\nu_h)_i (\nu_h)_j$. Here, ν_h is the normal to $\mathbb{S}^1_h$. Given some approximation $F_h \colon \mathbb{S}^1_h \to \mathbb{R}$ of the function $f \colon \mathbb{S}^1 \to \mathbb{R}$, the weak formulation on $\mathbb{S}^1_h$ is to find $u_h \in S_h$ such that

$$\int_{\mathbb{S}^1_h} \left(u_h v_h + \nabla_{\mathbb{S}^1_h} u_h \cdot \nabla_{\mathbb{S}^1_h} v_h \right) \mathrm{d}x_h = \int_{\mathbb{S}^1_h} F_h v_h \, \mathrm{d}x_h,$$

for all $v_h \in S_h$. The weak form on S_h is equivalent to the linear system that determines the coefficients $\beta_i, i = 1, 2, 3$ in the expansion of u_h in terms of the basis functions. This linear system is given by

$$\mathbf{I}\beta + \mathbf{S}\beta = \mathbf{F}_h,$$

where $[\mathbf{I}]_{ij} = \int_{\mathbb{S}^1_h} \phi_i \phi_j \, \mathrm{d}x_h$, $[\mathbf{S}]_{ij} = \int_{\mathbb{S}^1_h} \nabla_{\mathbb{S}^1_h} \phi_i \cdot \nabla_{\mathbb{S}^1_h} \phi_j \, \mathrm{d}x_h$, $\beta = (\beta_1, \beta_2, \beta_3)$ and $[\mathbf{F}_h]_i = \int_{\mathbb{S}^1_h} F_h \phi_i \, \mathrm{d}x_h$. The matrices $\mathbf{I}$ and $\mathbf{S}$ can be computed explicitly,

$$\mathbf{I} = \frac{h}{6} \begin{bmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{bmatrix}, \quad \mathbf{S} = \frac{1}{h} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}.$$

To approximate the function f on the mesh, we use piecewise linear interpolation: we set $F_h(v_i) = f(\theta_i)$, $i = 1, 2, 3$ and linearly interpolate between the points, along the line segments to obtain,

$$F_h|_{T_1}(t) = \frac{2h - 3t}{2h},\ F_h|_{T_2}(t) = -\frac{1}{2},\ F_h|_{T_3}(t) = \frac{3t - h}{2h},$$

for $t \in [0, h]$. The right-hand side of the linear system is given by

$$\mathbf{F}_h = \begin{bmatrix} \sum_{i=1}^{3} \int_{T_i} F_h|_{T_i}(t)\phi_1|_{T_i}(t)\,\mathrm{d}t \\ \sum_{i=1}^{3} \int_{T_i} F_h|_{T_i}(t)\phi_2|_{T_i}(t)\,\mathrm{d}t \\ \sum_{i=1}^{3} \int_{T_i} F_h|_{T_i}(t)\phi_3|_{T_i}(t)\,\mathrm{d}t \end{bmatrix} = \begin{bmatrix} h/2 \\ -h/4 \\ -h/4 \end{bmatrix}.$$

The solution to the linear system is then given by

$$\beta = \begin{bmatrix} \frac{h^2}{h^2+6} \\ -\frac{h^2}{2(h^2+6)} \\ -\frac{h^2}{2(h^2+6)} \end{bmatrix},$$

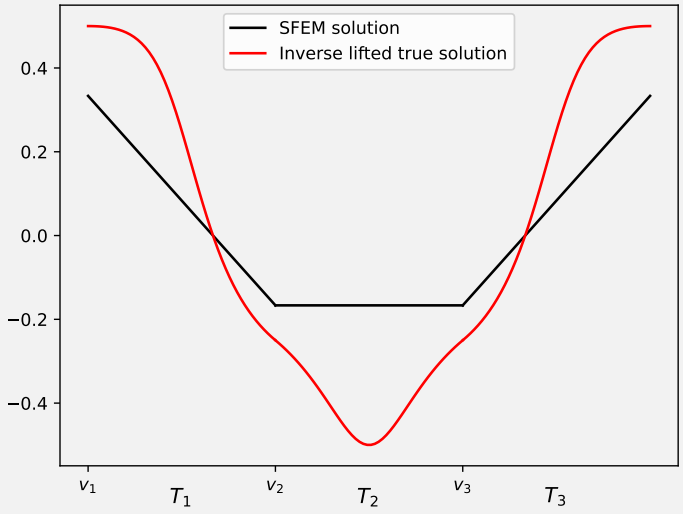
and the FEM approximation of u is given by $u_h = \sum_{i=1}^{3} \beta_i \phi_i$. To compare u_h with u, we need to move one of the functions to the other space, i.e., we must lift u_h to u or project u to u_h. In SFEM, u_h can be moved to $\mathbb{S}^1$ by lifting it from $\mathbb{S}_h^1$ along the normal of $\mathbb{S}^1$. The starting point for the lifting operator is the isomorphism $p : \mathbb{S}_h^1 \to \mathbb{S}^1$ given by

$$p(x) = \frac{x}{\|x\|_{\mathbb{R}^2}}.$$

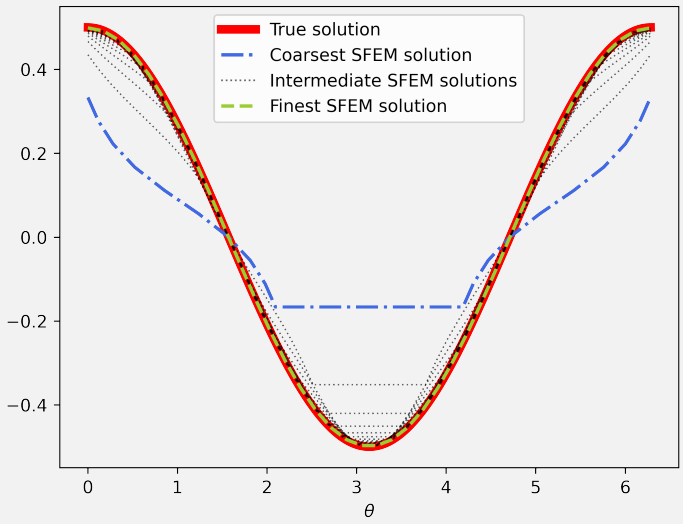
The lift of a function $\eta \colon \mathbb{S}_h^1 \to \mathbb{R}$ is the function $\eta^\ell \colon \mathbb{S}^1 \to \mathbb{R}$ given by

$$\eta^\ell(x) = \eta \circ p^{-1}(x).$$

When, as in this case, one is working explicitly, it can be easier to instead send the true function to the mesh. The inverse lift of the function f evaluated at a point $p \in \mathbb{S}_h^1$ is given by $f(\theta_p)$, where θ_p is the angle of the point p expressed in polar coordinates. The approximate solution u_h and the inverse lifted true solution are shown below.



As expected, this approximation is not very good, as the mesh is very coarse. To improve the approximation, one can refine the mesh, i.e., add more points to the mesh, and then repeat the procedure. Below, we show the behavior of the approximation as the mesh is refined.



Note that in the above figure, we plot the solution as well as its approximations on the interval $[0, 2\pi]$, i.e., on the circle. This is to facilitate the comparison of the solutions, but it also showcases another interesting feature of SFEM. Consider the coarsest solution. On S_h^1, it is piecewise linear, but when lifted to $\mathbb{S}^1$, it is not piecewise linear. This is due to the fact that the normal to the circle is not constant.

Chapter 4: Matching Problems and Hydrodynamics

Having covered the mathematical and mechanical background, we now turn to matching problems and hydrodynamics. Many concepts presented in this chapter can be understood as examples of mechanical systems evolving as geodesics on Lie groups. In other words, we are considering generalizations of the rigid body as presented in Example 2.17. This generalization was first made by Arnold (1966), who showed that the Euler equations for an incompressible fluid can be understood as a geodesic on the group of volume-preserving diffeomorphisms. In this chapter, we first briefly discuss how the set of diffeomorphisms can be understood as a group and how this group separates into different subgroups. Next, we touch on the geometric understanding of ideal hydrodynamics, before moving on to *shape matching*. To conclude the theoretical part of the chapter, we discuss the optimal transport problem and how it can be understood in the same context as the other problems presented in this chapter. Finally, we discuss some aspects of numerics for these kinds of problems.

4.1 The Diffeomorphism Group

In this section, we briefly introduce the diffeomorphism group and its structure. For more details, see (Hamilton, 1982; Khesin and Wendt, 2009; Schmeding, 2022).

Let (M, g) be a smooth, oriented and compact Riemannian manifold. Consider two diffeomorphisms ϕ, ψ of M. We can make some observations about their composition.

1. The composition of two diffeomorphisms is a diffeomorphism, i.e., $\phi \circ \psi$ is a diffeomorphism.

2. The inverse of a diffeomorphism is a diffeomorphism, i.e., ϕ^{-1} is a diffeomorphism.

3. There is an identity diffeomorphism, namely the identity map $e(x) = x$ for all $x \in M$.

4. It holds that $\psi \circ \psi^{-1} = e$.

Thus, the set of all diffeomorphisms of M is a group, denoted $\mathrm{Diff}(M)$.

In fact, a more can be said: $\mathrm{Diff}(M)$ can be interpreted as an infinite-dimensional Lie group (Leslie, 1967). A detailed treatment of infinite-dimensional Lie groups is beyond the scope of this thesis, but briefly, an infinite-dimensional manifold is modeled on an infinite-dimensional space rather than a finite-dimensional one. The group of diffeomorphisms is a *Fréchet manifold*, that is, a manifold that locally looks like a Fréchet space. To make it a Lie group, the inversion and composition maps must be smooth. The Lie algebra of $\mathrm{Diff}(M)$ is given by $\mathfrak{X}(M)$.

It is possible to equip $\mathrm{Diff}(M)$ with a Riemannian metric of the form

$$\langle \dot{\gamma}, \dot{\gamma} \rangle_\gamma = \int_M v \cdot Lv \, \mu_g, \tag{4.1}$$

where γ is a smooth curve taking values in $\mathrm{Diff}(M)$, $L \colon \mathfrak{X}(M) \to \mathfrak{X}^*(M)$ is an invertible elliptic differential operator acting on vector fields and $v = \dot{\gamma} \circ \gamma^{-1}$. One example of such an operator is $L = (1 - \Delta_M)^k$ for some $k \in \mathbb{N} \cup \{0\}$. The metric (4.1) is right-invariant as well as *weak*, meaning that the metric does not provide an isomorphism between tangent spaces and cotangent spaces, but only an injective map.

The invariance means that the Euler–Arnold framework can be applied to the diffeomorphism group. In Section 2.5 the metric is left-invariant. However, a right-invariant metric can be made left-invariant by switching the order of the composition of the diffeomorphisms, and the only difference is that the sign changes.

An important subgroup of $\mathrm{Diff}(M)$ is the group of volume-preserving diffeomorphisms, denoted $\mathrm{SDiff}(M)$. This means that the determinant of the Jacobian of the diffeomorphism is equal to one. The Lie algebra of $\mathrm{SDiff}(M)$ is given by the space of divergence-free vector fields, denoted $\mathfrak{X}_{\mathrm{div}}(M)$. An arbitrary smooth vector field can be decomposed into a divergence-free part and a gradient part by $v = v_{\mathrm{div}} + \nabla f$, where $v_{\mathrm{div}} \in \mathfrak{X}_{\mathrm{div}}(M)$ for some $f \in C^\infty(M)$. Moreover, the divergence-free vector fields are orthogonal to the gradient fields with respect to the L^2 inner product on M.

4.2 Spherical Ideal Hydrodynamics

The motion of an incompressible fluid on a Riemannian manifold (M, g) is governed by the Euler equation,

$$\dot{v} + \nabla_v v = -\nabla p, \tag{4.2}$$

where v denotes the velocity field and p the pressure. In (Arnold, 1966), it was shown that the Euler equations (4.2) arise similarly to the rigid body equations, namely as a geodesic on a Lie group.

Consider a fluid occupying M. The fluid is incompressible, meaning that the volume of a reference fluid element is preserved by the flow. We track the motion of the fluid by a mapping $\phi\colon M \times [0,1] \to M$. This mapping must be a diffeomorphism to prevent any non-physical behavior of the fluid, such as discontinuities or particles passing through each other. Moreover, it has to be volume-preserving, i.e., be an element of $\mathrm{SDiff}(M)$. The velocity of a fluid element is then given by $\dot{\phi}$, and its kinetic energy is given by

$$\mathcal{L}(\phi, \dot{\phi}) = \frac{1}{2} \langle \dot{\phi}, \dot{\phi} \rangle_\phi = \frac{1}{2} \int_M v \cdot v \, \mu_g,$$

where μ_g is the volume form induced by the metric g, $v = \dot{\phi} \circ \phi^{-1}$ is the velocity field of the fluid and $\langle \cdot, \cdot \rangle_.$ is the metric on $\mathrm{SDiff}(M)$ given by (4.1) with $L = \mathrm{Id}$.

A curve that extremizes the kinetic energy is a geodesic on $\mathrm{SDiff}(M)$, and one can directly compute the minimizer by calculus of variations, that is, by taking variations of ϕ. See for instance (Modin, 2019) for details on this approach. Alternatively, one can note that due to the right-invariance of the metric, the equations of motion are the Euler–Arnold equations. We identify the Lie algebra of divergence-free vector fields with its dual by the inner product $\frac{1}{2} \int_M v \cdot v \, \mu_g$ and consequently the Euler–Arnold equation becomes

$$\dot{v} - \mathrm{ad}^*_v \, v = 0.$$

Thus, to determine the equations of motion, it suffices to identify what ad^* is in this context. It turns out that

$$\langle \mathrm{ad}^*_v \, v, w \rangle = \int_M \nabla_v v \cdot w \, \mu_g.$$

However, we cannot conclude that $\mathrm{ad}^*_v \, v = \nabla_v v$ since $\nabla_v v$ is not necessarily divergence-free, and we are testing it against a divergence-free vector field. Therefore, the identification of $\mathrm{ad}^*_v \, v$ with $\nabla_v v$ is only valid up to a gradient field. That is, $\mathrm{ad}^*_v \, v = \nabla_v v + \nabla p$ for some $p \in C^\infty(M)$ (Arnold, 1989, Appendix B), (Marsden and Ratiu, 1999, Section 1.5). Thus, we recover the Euler equations (4.2).

In the following, we focus on the case $M = \mathbb{S}^2$. While we could consider a larger class of manifolds, known as *Kähler manifolds*, of which the sphere is an example, we stick to the sphere for brevity, as the arguments are similar to the general case but much more transparent. Let $J\colon T\mathbb{S}^2 \to T\mathbb{S}^2$ be the mapping that rotates tangent vectors by an angle of $\pi/2$. Interestingly, that

the Hamiltonian vector field of the function f is related to the gradient of f by

$$X_f = -J\nabla f.$$

This is a consequence of the fact that the sphere is a Kähler manifold, i.e., that J connects the symplectic form with the metric.

As discussed above, the Euler equations can be interpreted as an Euler–Arnold equation on $\mathrm{SDiff}(\mathbb{S}^2)$. It also turns out that the Euler equations also can be written as a Lie–Poisson system on $(C_0^\infty(\mathbb{S}^2), \{\cdot,\cdot\})$, where

$$\{\psi_1, \psi_2\} = J\nabla\psi_1 \cdot \nabla\psi_2.$$

To see this, first note that a smooth function $\psi \in C^\infty(\mathbb{S}^2)$ is mapped to a divergence-free vector field by $\psi \mapsto -J\nabla\psi$. This mapping is not an isomorphism; for example, all constant functions are mapped to the zero vector field. However, if we quotient out the constants, we obtain an isomorphism between $C^\infty(\mathbb{S}^2)/\mathbb{R}$ and $\mathfrak{X}_{\mathrm{div}}(\mathbb{S}^2)$. Moreover, this is a Lie algebra anti-homomorphism, since

$$-X_{\{\psi_1,\psi_2\}} = [X_{\psi_1}, X_{\psi_2}].$$

A Lie–Poisson system evolves on the dual of the Lie algebra. To find the dual of $C^\infty(\mathbb{S}^2)/\mathbb{R}$, note first that $C^\infty(\mathbb{S}^2)^* \cong C^\infty(\mathbb{S}^2)$ using the L^2 pairing $\langle\cdot,\cdot\rangle_{L^2(\mathbb{S}^2)}$. With this identification, the dual of $C^\infty(\mathbb{S}^2)/\mathbb{R}$ is the annihilator of the constant functions, that is, the set of functions satisfying

$$\int_{\mathbb{S}^2} \psi\, \mu_g = 0.$$

This means that the dual is the space $C_0^\infty(\mathbb{S}^2)$. The general form of a Lie–Poisson system on $C^\infty(\mathbb{S}^2)$ is

$$\dot\omega + \mathrm{ad}^*_{\frac{\delta\mathcal{H}}{\delta\omega}}\,\omega = 0,$$

for some Hamiltonian $\mathcal{H}\colon C_0^\infty(\mathbb{S}^2) \to \mathbb{R}$. On $\mathbb{S}^2$, ad^* is given by $\mathrm{ad}^*_\psi\,\omega = -\{\psi,\omega\}$. It thus remains only to determine the Hamiltonian of the Euler equations. First, note that $v = -J\nabla\psi$ for some ψ. Using the kinetic energy as the Hamiltonian, we have

$$\mathcal{H}(\omega) = \frac{1}{2}\int_{\mathbb{S}^2} v\cdot v\,\mu_g = \frac{1}{2}\int_{\mathbb{S}^2}\langle J\nabla\psi, J\nabla\psi\rangle\,\mu_g = -\frac{1}{2}\int_{\mathbb{S}^2}\psi\Delta_{\mathbb{S}^2}\psi\,\mu_g.$$

Hence, $\Delta_{\mathbb{S}^2}\psi = \omega$ and we obtain the *vorticity formulation* of the Euler equations on $\mathbb{S}^2$:

$$\dot\omega + \{\psi,\omega\} = 0, \quad \Delta_{\mathbb{S}^2}\psi = \omega. \tag{4.3}$$

An infinite number of conserved quantities, known as Casimirs, exist for Equation (4.3). These are given by integrals of the form

$$C_k = \int_{\mathbb{S}^2} \psi^k \, \mu_g,$$

for $k \in \mathbb{N}$. Solutions of Euler equations evolve on the level sets of the Casimirs. The second Casimir (enstrophy) plays an important role in fluid behavior (see for instance (Kraichnan, 1967)). It has also been conjectured that higher-order Casimirs are important for the formation of large-scale vorticity structures as well (Abramov and Majda, 2003; Newton, 2015).

4.3 Shape Analysis

Shape analysis, the mathematical study of how shapes can be compared, is another area of mathematics that can be viewed through the lens of geodesics on the diffeomorphism group. To compare shapes, we must define a meaningful notion of distance between shapes. Shapes—at this stage still loosely defined—are nonlinear objects, and there is no canonical way to compare them. Shape analysis addresses this problem by constructing shape matching frameworks that are flexible enough to be applicable in many settings.

The origins of shape matching go back to D'Arcy Thompson in (Thompson, 1992), while its modern form was developed by Grenander (1994). Thorough treatments of the topic can be found in (Bruveris and Holm, 2015; Younes, 2010). The basic idea, as illustrated in Figure 4.1, is to find a suitable *warp* that takes one shape to another in an energy-minimizing way. One can then compare two shapes by measuring the energy of the warp that takes one shape to the other. This has found applications in for instance medical image analysis (Bistoquet et al., 2008; Bruveris and Holm, 2015; Ceritoglu et al., 2013; Risser et al., 2013; Qiu et al., 2009).

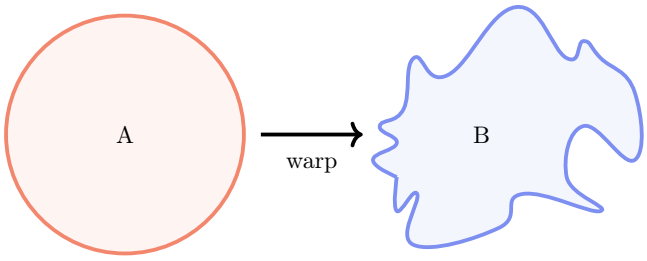


Figure 4.1: The idea of shape matching is to move A to B in an energy minimizing manner. From (Jansson, 2022).

In more detail, the goal is to identify a set of suitable shapes (often a metric space) $\mathcal{S}$ and a Lie group H acting on $\mathcal{S}$. The deformation of shapes

by elements of H is defined as the action of H on $\mathcal{S}$. In many cases, the group H is the group of diffeomorphisms of M, i.e., $\mathrm{Diff}(M)$, where M is the underlying spatial domain of the shapes. A diffeomorphism $\phi \in \mathrm{Diff}(M)$ warps the shape $A \in \mathcal{S}$ by deforming the underlying domain M by ϕ. The action depends on the choice of shape space, which varies depending on the context.

Example 4.1: Shapes

A few examples of shape spaces are

Smooth functions: Images can be modeled as smooth functions on a domain. For image matching, the shape space is $C^\infty(M)$ and a diffeomorphism ϕ warps the shape $I \in C^\infty(M)$ by the action $\phi \cdot I = I \circ \phi^{-1}$.

Landmarks: Landmarks are points on M used to describe shapes. In this case, the shape space is M^N where N is the number of landmarks. A diffeomorphism ϕ warps the shape $L = (L_1, \dots, L_N) \in M^N$ by $\phi \cdot L = (\phi(L_1), \dots, \phi(L_N))$.

Densities: Densities can also be understood as shapes. In this case, the shape space is the set of smooth functions on M that are positive and integrate to one. A diffeomorphism ϕ warps the shape f by $\phi \cdot f = f \circ \phi^{-1} \det(D\phi)$, where $D\phi$ is the Jacobian of ϕ.

In the following, we focus on the case when H is the group of diffeomorphisms. The goal of shape matching is to find a diffeomorphism that minimizes an appropriate distance between two shapes. For example, if $\mathcal{S} = C^\infty(M)$, the distance between two shapes $A, B \in C^\infty(M)$ can be measured by the L^2-distance, i.e.,

$$d(A, B) = \|A - B\|_{L^2(M)} = \left(\int_M |A(x) - B(x)|^2 \, \mu_g \right)^{1/2}.$$

A naive first approach would be to just minimize the distance between the shapes, i.e., to find a diffeomorphism ϕ that minimizes $d(\phi \cdot A, B)$. This is known as *greedy matching*. However, we cannot assume that A and B are in the same orbit of the action of $\mathrm{Diff}(M)$; there may not exist a diffeomorphism ϕ such that $\phi \cdot A = B$. As a result, greedy matching will run indefinitely and generate increasingly complex warps of the underlying space. To address this, a regularization term can be introduced to penalize the complexity of the warp. The Riemannian metric on $\mathrm{Diff}(M)$ (4.1) can be used to regularize the

matching problem by the term

$$\frac{1}{2\sigma}\int_0^1 \langle \dot\gamma(t), \dot\gamma(t)\rangle_{\gamma(t)}\,\mathrm{d}t = \frac{1}{2\sigma}\int_0^1 \int_M v(t)\cdot Lv(t)\,\mu_g \mathrm{d}t, \tag{4.4}$$

where $\sigma > 0$ is a regularization parameter and v is the time-dependent vector field that generates the warp by the differential equation $\dot\gamma(t) = v(t)\circ\gamma(t), \gamma(0) = e$. Thus, the energy (or cost) of matching the shapes A and B is given by the functional

$$E(v) = d(\gamma(1)\cdot A, B) + \frac{1}{2\sigma}\int_0^1 \int_M v(t)\cdot Lv(t)\mu_g \mathrm{d}t. \tag{4.5}$$

Interestingly, the inclusion of the regularization terms leads to a variational (or mechanical) understanding of the problem of minimizing E in Equation (4.5) over curves $v : [0,1] \to \mathfrak{X}(M)$, that is to say, a reduction to a dynamical formulation. To see this, note first that the matching term $d(\gamma(1)\cdot A, B)$ depends only on the *final* time point of the curve $t \mapsto \gamma(t)$. Hence, an optimal curve $t \mapsto v(t)$ must follow the dynamics determined by the action functional that consists solely of the regularization term (4.4). Due to the right-invariance of the metric, the Euler–Arnold framework can be applied to the problem, as with the rigid body and the Euler equations. Indeed, by Marsden and Ratiu (1999, Theorem 13.5.3), the optimal curve $t \mapsto v(t)$ must satisfy the Euler–Arnold equation (2.19) where $m = Lv$. As with the Euler equations, the derivation comes down to determining what $\mathrm{ad}^*_v\, m$ is on the diffeomorphism group. For an arbitrary vector field $w \in \mathfrak{X}(M)$, we have that

$$\begin{aligned}\langle \mathrm{ad}^*_v\, m, w\rangle &= -\langle m, [v,w]\rangle = -\langle m, \nabla_v w - \nabla_w v\rangle \\ &= -\int_M m\cdot\nabla_v w - m\cdot\nabla_w v\,\mu_g.\end{aligned}$$

Metric compatibility implies that $m\cdot\nabla_v w = v(m\cdot w) - w\cdot\nabla_v m$. The divergence satisfies that for any vector field v and smooth function f, $\mathrm{div}(fv) = f\,\mathrm{div}\, v + v(f)$, so $m\cdot\nabla_v w = \mathrm{div}((m\cdot w)v) - w\cdot m\,\mathrm{div}\, v - w\cdot\nabla_v m$, and thus

$$\begin{aligned}\int_M m\cdot\nabla_v w\,\mu_g &= \int_M (\mathrm{div}((m\cdot w)v) - w\cdot m\,\mathrm{div}\, v - w\cdot\nabla_v m)\ \mu_g \\ &= -\int_M (w\cdot m\,\mathrm{div}\, v + w\cdot\nabla_v m)\ \mu_g\end{aligned}$$

where the final step follows from the divergence theorem on Riemannian manifolds (Lee, 2012, Theorem 16.32). Finally, we define $\nabla^\top_m v$ by

$$\int_M \nabla^\top_m v\cdot w\,\mu_g = \int_M m\cdot\nabla_w v\,\mu_g,$$

so we have

$$\langle \operatorname{ad}_v^* m, w\rangle = \langle \nabla_m^\top v + \nabla_v m + m \operatorname{div} v, w\rangle,$$

and we see that the Euler–Arnold equation (2.19) becomes

$$\dot{m} + \nabla_m^\top v + \nabla_v m + m \operatorname{div} v = 0, \quad m = Lv. \tag{4.6}$$

These equations are known as the EPDiff equations. An illustration of shape matching is provided in Figure 4.2. The EPDiff equations (4.6) generates a curve of diffeomorphisms $\gamma\colon [0,1] \to \operatorname{Diff}(M)$ that is a geodesic in $\operatorname{Diff}(M)$. Since $\psi \in \operatorname{Diff}(M)$ acts on a shape $A \in \mathcal{S}$ by $\psi.A$, γ descends to a flow on the *orbit* of A, given by

$$\operatorname{Orb}(A) \coloneqq \{\psi.A\colon \psi \in \operatorname{Diff}(M)\} \subset \mathcal{S}.$$

The goal of shape matching is to find the element in the orbit of A that is as close to the target B as possible. Typically, $B \notin \operatorname{Orb}(A)$, so the matching will not be exact. However, this is precisely the point: by only being able to match approximately, the problem is regularized compared to greedy matching.

Furthermore, consider the set of diffeomorphisms that leaves A unchanged,

$$\operatorname{Diff}_A(M) \coloneqq \{\psi \in \operatorname{Diff}(M) : \psi.A = A\} \subset \operatorname{Diff}(M).$$

Note that if $\psi, \varphi \in \operatorname{Diff}_A(M)$, then $(\psi \circ \varphi).A = \psi.\varphi.A = A$, and $\psi^{-1}.A = \psi^{-1}.\psi.A = (\psi^{-1} \circ \psi).A = A$. Thus, $\operatorname{Diff}_A(M)$ is a subgroup of $\operatorname{Diff}(M)$. The orbit can be understood as the quotient set $\operatorname{Diff}(M)/\operatorname{Diff}_A(M)$.

The EPDiff equations (4.6) provide a starting point for concrete matching algorithms: begin with an initial guess of the initial value $v(0)$ of v, then solve the EPDiff equations (4.6) to determine $v(t)$ for all $t \in [0,1]$. With this curve, we compute $\gamma(1)$ and evaluate the modified energy

$$\tilde{E}(v) = d(\gamma(1) \cdot A, B) + \frac{1}{2\sigma} \int_M v(0) \cdot Lv(0)\, \mu_g.$$

This simplified expression for the energy (4.5) is valid since the regularization energy is conserved along the dynamics of the EPDiff equations (4.6). Indeed,

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_M v \cdot m\, \mu_g = 2 \int_M v \cdot \dot{m}\, \mu_g = 2 \int_M m \cdot \operatorname{ad}_v v\, \mu_g = 0,$$

since $\operatorname{ad}_v v = 0$. The initial guess of $v(0)$ is then updated by computing the gradient of the energy $\tilde{E}$ with respect to $v(0)$ and taking a step in the negative gradient direction, as detailed in (Beg et al., 2005). Of course, the EPDiff equations are partial differential equations, so a spatio-temporal discretization is in practice needed to solve them.

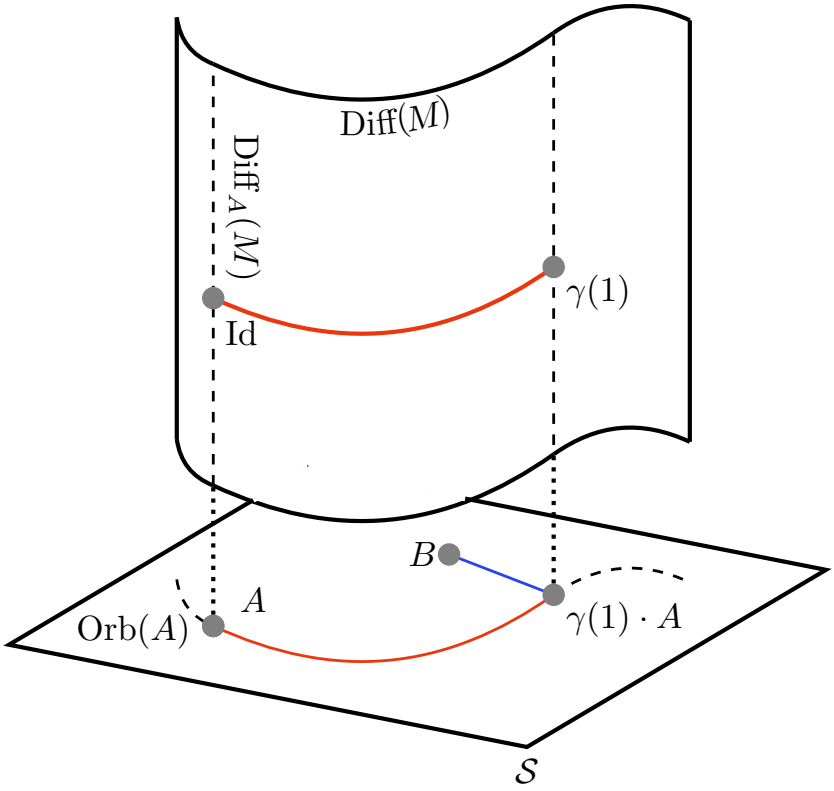


Figure 4.2: Illustration of the geometric structure of shape matching. The EPDiff equations induce a curve in $\mathrm{Diff}(M)$ that moves the initial shape $A \in \mathcal{S}$ along its orbit to get A as close as possible to the target shape B. From (Jansson, 2022).

4.4 Optimal Transport

Optimal transport is another type of matching problem, concerned with finding the most efficient way to transport one distribution of mass to another. The classical example is the problem of moving a pile of earth into a hole in the ground in the most efficient way. More specifically, given two probability measures μ and ν on measurable spaces X and Y, the optimal transport problem is to find a measurable map $T\colon X \to Y$ that minimizes the cost

$$\int_X c(x, T(x)) \, \mathrm{d}\mu(x),$$

while satisfying the constraint that $\nu(B) = \mu(T^{-1}(B))$ for all measurable sets $B \subset Y$. Here $c\colon X \times Y \to \mathbb{R}$ is the distance function.

This problem, known as the Monge problem, has a rich and well-developed theory, see for instance (Villani, 2009), which is beyond the scope of this thesis. Instead, we will briefly consider the fluid-mechanical interpretation of the optimal transport problem, known as the Benamou–Brenier approach (Benamou and Brenier, 2000), as it is closely related to the Euler equations and

the diffeomorphism group. Let $\mathrm{Dens}(M) = \{f \in C^\infty(M) : f > 0, \int_M f\,\mu_g = 1\}$ be the set of densities on M. This space is, as with diffeomorphism group, a Fréchet manifold (Hamilton, 1982, Chapter 4). Its tangent spaces are given by smooth functions that integrates to zero, i.e., $T_\rho \mathrm{Dens}(M) = \{f \in C^\infty(M) : \int_M f\,\mu_g = 0\}$. We can equip $\mathrm{Dens}(M)$ with a Riemannian metric, the Otto metric, given by

$$\langle \dot\rho, \dot\rho \rangle_\rho = \int_M |\nabla S|^2 \rho\,\mu_g,$$

where S is given by $\dot\rho + \mathrm{div}(\rho \nabla S) = 0$. The optimal transport problem can be understood as a Lagrangian variational problem of finding a path $\rho : [0,1] \to \mathrm{Dens}(M)$ that minimizes the action

$$\frac{1}{2}\int_0^1 \langle \dot\rho_t, \dot\rho_t \rangle_{\rho_t}\,\mathrm{d}t.$$

The optimal transportation map is given by the flow of the vector field $-\nabla S_t(x)$ at time $t = 1$, where S_t is the solution to the continuity equation $\dot\rho_t + \mathrm{div}(\rho_t \nabla S_t) = 0$. Moreover, the Otto metric arises naturally. When the cost is quadratic, the Otto metric arises as the *reduction* of an L^2 metric on the diffeomorphism group $\mathrm{Diff}(M)$ to the space of densities $\mathrm{Dens}(M)$, as detailed in for instance (Khesin and Wendt, 2009; Modin, 2017; Otto, 2001). This reduction is carried out by a certain projection mapping, which provides a one-to-one correspondence between geodesics on $\mathrm{Dens}(M)$ and a certain class of geodesics on $\mathrm{Diff}(M)$ (Khesin and Wendt, 2009, Corollary A.5:5.7). This connection not only provides an interesting geometric point of view of the optimal transport problem, but also provides a way to study the geometry of certain matrix decompositions, as demonstrated in (Modin, 2017). The matrix flows discussed in (Modin, 2017) served as the foundation for the work presented in Paper IV.

4.5 Geometry-Preserving Spatial Discretization

In the previous section, we have seen a few different types of problems that can all be understood as geodesic problems on Lie groups. However, their motivations in applications are all quite different. For instance, the Euler equations have a clear physical interpretation, with several conserved quantities that are important for the dynamics of the system. In geophysical fluid dynamics, the long-term behavior of the Euler equations is of interest, and its behavior is in those cases often determined by the geometric properties of the equations. Thus, when simulating the Euler equations or other fluid dynamical equations, over long timescales, it is important to preserve these geometric properties, i.e., to use geometric numerical methods. When solving equations

arising in shape matching, the primary objective is to find an optimal warp that transforms one shape into another. Given that the simulation horizon is often relatively short, it can be argued that less sophisticated numerical methods may suffice for these problems. For instance, finite differences or finite elements in space combined with an explicit time-stepping scheme can be good enough.

For the Euler equations, we need structure preservation in both space and time. The structure preserving spatial discretization we consider here is based on the idea of *Berezin–Toeplitz quantization* (Hoppe, 1989; Hoppe and Yau, 1998). A detailed description of this numerical method, known as *Zeitlin's method*, introduced in (Zeitlin, 1991, 2004), can be found in (Modin and Viviani, 2024). The method applies to a relatively small class of manifolds, but one that includes the important examples such as the sphere and the flat torus. The idea is to take the vorticity equation (4.3) and replace it with a Lie–Poisson system on $\mathfrak{su}^*(N)$, the algebra of skew-Hermitian matrices, for some $N \in \mathbb{N}$. This system is given by

$$\dot{W} + \frac{1}{\hbar}[P, W] = 0, \qquad \Delta_N W = P,$$

where $P, W \in \mathfrak{su}(N)$ and $\Delta_N \colon \mathfrak{su}(N) \to \mathfrak{su}(N)$ is a quantized version of the Laplacian. Zeitlin's method is interesting in its own right, and has a beautiful geometric description, as detailed in (Modin and Viviani, 2024). However, from the perspective of traditional numerical analysis, Zeitlin's method displays relatively poor performance. It is proven in (Gallagher, 2002) (for the toroidal case) and (Modin and Viviani, 2024) (for the spherical case) that Zeitlin's method converges, for a fixed final time, as $1/N$, but with a constant depending exponentially on the final time. Thus, for long-time simulations, the error may grow large.

From a geometric numerical perspective, however, the method is still useful, as it preserves the geometric properties of the Euler equations in the sense that

1. the discrete system is a Lie–Poisson system, just as the continuous system, and

2. there are N independent Casimirs, of the form $\operatorname{tr}(W^k)$, $k = 1, \dots, N$.

These properties are important for the long-term behavior of the system, and thus, the method is useful for long-time simulations of the Euler equations. Moreover, Arnold (1989, Appendix B) noted that the instability—the fact that two similar initial conditions of the Euler equations can lead to diverging solutions—can be understood by the fact that the curvature of the diffeomorphism group is negative. Modin and Perrot (2024) showed that Zeitlin's method correctly approximates the curvature of the diffeomorphism group,

meaning that Zeitlin's method displays the same type of stability as the Euler equations if N is large enough.

In time, a Lie–Poisson integrator such as the isospectral midpoint method of Modin and Viviani (2019b) should be used, to ensure the preservation of the above geometric properties. Simulations with Zeitlin's method provide physically plausible behavior such as the conservation of enstrophy over long timescales. This contrasts with more traditional numerical methods, such as the spectral method proposed by Dritschel et al. (2015).

Returning briefly to shape matching (or optimal transport, for that matter), the situation, as noted earlier, differs somewhat. The primary goal is to find an optimal warp that transforms one shape into another, and the long-term behavior of the system is generally not of concern. Nevertheless, it is sometimes useful to obtain spatially discrete systems in a way that capture the desired geometric properties. In Paper VI, the shapes are chains of molecules, and we wish to act on them only by rigid motions, i.e., elements of the rotation group. To achieve this, it is possible to restrict the set of available deformations to only include the Lie group of rotations. The shape matching framework still applies in this case, but, not only are we guaranteed to obtain rigid deformations, the problem is also automatically spatially discrete, i.e., there is only need for a temporal discretization. In this sense, it can also be viewed as a geometric numerical method.

Chapter 5: Summary of Included Papers

Having discussed some background and motivation for the included papers, we now provide a brief summary of each paper.

5.1 Paper I

We study a finite element-based solution method for the stochastic elliptic differential equation

$$(\kappa^2 - \Delta_{\mathbb{S}^2})^\beta u = \mathcal{W},$$

where $\kappa > 0$ is a constant, $\Delta_{\mathbb{S}^2}$ is the Laplace–Beltrami operator on the sphere $\mathbb{S}^2$, $\beta > 1/2$, and $\mathcal{W}$ is white noise on $\mathbb{S}^2$. This is exactly the Whittle–Matérn random field on the sphere, as described in Section 3.3 and Section 3.4. A sample realization of u is pictured in Section 3.3. The solution of the equation is a random field on $\mathbb{S}^2$ that, following (Lindgren et al., 2011), is the equivalent of a Whittle–Matérn random field on $\mathbb{R}^2$. The main contribution of the paper is the derivation of a surface finite element method for the equation, building on the methodology of Dziuk and Elliott (2013). The operator in question is fractional, and the proposed solution is to split the equation into two parts: a non-fractional part and a fractional part. By letting $\beta = \lfloor \beta \rfloor + \{\beta\}$, where $\lfloor \beta \rfloor \in \mathbb{N}$ denotes the integer part of β and $\{\beta\} \in [0, 1)$ the fractional part, the equation is rewritten as a system of equations given by the recursion

$$(\kappa^2 - \Delta_{\mathbb{S}^2})u^i = u^{i-1}$$

for $i = 1, \ldots, \lfloor \beta \rfloor$ with $u^0 = \mathcal{W}$. If $\{\beta\} = 0$, we set $u = u^{\lfloor \beta \rfloor}$, otherwise, we solve the equation $u = (\kappa^2 - \Delta_{\mathbb{S}^2})^{\{\beta\}} u^{\lfloor \beta \rfloor}$. This fractional operator is, following Bonito and Pasciak (2015), approximated using a sinc quadrature rule, in which the inverse fractional operator is approximated by the sum of the solution of several elliptic problems. In other words, the price to pay to get rid of the fractional operator is that several elliptic problems need to be solved. We prove that the method converges in mean square, and provide some numerical illustrations of this fact.

5.2 Paper II

We study the numerical approximation of random fields defined on a compact, oriented and boundaryless smooth surface or curve M that is given by

$$\mathcal{Z} = \zeta(\mathcal{L})\mathcal{W},$$

where $\mathcal{W}$ is white noise on M. This equation is described in more detail in Section 3.4.

Here, $\mathcal{L}$ is the elliptic differential operator associated to the bilinear form

$$a(u,v) = \int_M \mathcal{D}(x)\nabla_M u(x) \cdot \nabla_M v(x)\ \mathrm{d}x + \int_M \mathcal{V}(x)u(x)v(x)\ \mathrm{d}x,$$

where $\mathcal{D}(x)\colon T_xM \to T_xM$ is a positive definite, symmetric operator, and $\mathcal{V}(x)\colon T_xM \to \mathbb{R}^+$ is a potential function. Finally, $\zeta\colon \mathbb{R} \to \mathbb{R}$ is a smooth function with sufficiently fast decay at infinity, to ensure that $\mathcal{Z}$ is regular enough. This model is similar to the one studied in (Lang and Pereira, 2023), in which $\mathcal{L}$ is the Laplace–Beltrami operator on M.

Using the eigenpairs (λ_i, e_i) of $\mathcal{L}$, the random field can be written as

$$\mathcal{Z} = \sum_{i=1}^{\infty} \zeta(\lambda_i) W_i e_i,$$

where W_i are independent standard Gaussian random variables. However, as the eigenpairs of $\mathcal{L}$ are unknown in general, we propose to use surface finite elements as in (Dziuk and Elliott, 2013) to approximate $\mathcal{Z}$. This is done by defining a discrete version of $\mathcal{L}$ defined on a triangulation of M that defines the approximate surface M_h, and then using the eigenpairs (Λ_i^h, E_i^h) of the discrete operator to approximate $\mathcal{Z}$ on M_h by

$$\mathsf{Z}_h = \sum_{i=1}^{N_h} \zeta(\Lambda_i^h) W_i^h E_i^h,$$

where W_i^h are independent standard Gaussian random variables. In order to avoid relying on the eigenpairs of the discrete operator, we prove that we can express Z_h using the nodal basis of the finite element space, i.e., that there are Gaussian weights Z_i such that

$$\mathsf{Z}_h = \sum_{i=1}^{N_h} Z_i \psi_i,$$

where ψ_i are the nodal basis functions of the finite element space $S_h \subset H^1(M_h)$. The weights Z_i are correlated, but the correlation matrix can be expressed in terms of the finite element matrices. The key contribution of the paper is to prove that the root-mean-square error between Z_h and $\mathcal{Z}$ converges to zero as the mesh size h goes to zero. The rate depends on the smoothness of $\mathcal{Z}$.

5.3 Paper III

We consider the numerical integration of stochastic Lie–Poisson systems on the class of J-quadratic Lie algebras $\mathfrak{h}$, i.e., equations of the form

$$\dot{\xi} = [\nabla \mathcal{H}_0(X_t)^*, \xi] + \sum_{k=1}^{M} [\nabla \mathcal{H}_k(X_t)^*, \xi] \circ \mathrm{d}B_t^k,$$

where $\xi \in \mathfrak{h}^*$, $\mathcal{H}_i \colon \mathfrak{h}^* \to \mathbb{R}$ and B_t^k are independent Brownian motions. These systems were discussed in more detail in Section 3.2. The goal is to derive integrators that preserve the geometric properties of the system, i.e., the Lie–Poisson structure and the Casimirs. To do this, we adapt the integrator for deterministic systems that was suggested by Modin and Viviani (2019b). In essence, the goal is to lift the system to the cotangent bundle of the Lie group of $\mathfrak{h}$ using the momentum map μ, on which the system "unreduces" to a canonical stochastic Hamiltonian system. This system is then integrated using the implicit midpoint method. Apart from deriving the integrator, the key contribution of the paper is to first prove the desired preservation properties, but also the convergence of the method. The theoretical methods used in the paper are all based on the momentum map framework of Section 2.4. The convergence proof essentially entails proving that the error analysis of the proposed method is, by the momentum map, inherited from the error analysis of the implicit midpoint method, which is well-studied in the literature, see for instance Milstein and Tretyakov (2004). The method and its convergence, is illustrated on several examples, including the rigid body appearing in Section 3.2 and the Zeitlin discretization of the Euler equations, briefly discussed in Section 4.5.

5.4 Paper IV

We consider landmark matching, as described in Section 4.3, modified in two ways. The first modification is that the vector field warping the initial landmarks is parametrized by a set $\mathcal{U}$ of parameters. The second modification, inspired by Chen and Öktem (2018), is that the matching term includes the case where the target landmarks $z_1, \ldots, z_m$ are in a metric space N that may be different from the original manifold M. The main reason for introducing these modifications is that they allow us to connect deep learning and landmark matching. This is possible by viewing residual neural networks as temporal discretizations of time-continuous control problems, as in (Celledoni et al., 2021; Li et al., 2018), and by describing how the control problems can be interpreted as high-dimensional landmark matching problems.

The main result of the paper lies in describing the geometric structure of the modified problem. The parametrization of vector fields determines a subset of $\mathfrak{X}(M)$ which is not a Lie subalgebra of $\mathfrak{X}(M)$. The matching problem is considered as a nonlinear control system where the space of controls is given by $\mathcal{U}$. Inspired by Younes et al. (2020), we refer to this version of landmark matching as *sub-Riemannian landmark matching*.

It is shown that a dynamic formulation of sub-Riemannian landmark matching can be derived using an Euler–Arnold argument, just as in the non-modified case. In more detail, an equation governing the evolution of the time-dependent control variable u is derived. We describe how the dynamic formulation allows for matching to be performed by a shooting algorithm, meaning that it is only the initial state of the control u that needs to be determined. Finally, some numerical results illustrate the algorithms.

5.5 Paper V

Paper V is concerned with proving the convergence of one of the matrix flows appearing in the context of the optimal transport of Gaussian distributions. The flow was proposed in Modin (2017), namely the vertical flow

$$\dot{B} = -\operatorname{Proj}_{\mathrm{Ver}} \nabla_{\mathcal{G}} J(B), B(0) = A,$$

where $J = \operatorname{Tr}(\Sigma_0(I - A^T)(I - A^T))$ for some positive definite and symmetric matrix Σ_0. Here, $\operatorname{Proj}_{\mathrm{Ver}}$ refers to the projection onto the fibers of the principal $\mathrm{O}(n, \Sigma_0)$-bundle over $P(n)$,

$$\begin{array}{ccc} \mathrm{GL}(n) & \longleftarrow & \mathrm{O}(n, \Sigma_0) \\ \downarrow & & \\ P(n) & & \end{array} \tag{5.1}$$

where $P(n)$ is the space of positive definite matrices and $\mathrm{O}(n, \Sigma_0)$ is the group of orthogonal matrices with respect to the metric Σ_0. The main result of the paper is to prove that the flow converges to the matrix obtained in the polar decomposition of A, i.e., the unique $P \in P(n)$ and $Q \in \mathrm{O}(, \Sigma_0)$ such that $A = PQ$. The techniques used in the paper relies on the geometric interpretation of the Monge problem briefly discussed in Section 4.4.

5.6 Paper VI

In this paper, we consider a shape matching problem that arises in the context of imaging of proteins. Cryogenic electron microscopy (cryo-EM) is a technique that allows for imaging of proteins at atomic resolution. Since

low doses of radiation are used, the resulting images are very noisy, so the reconstruction of the 3D structure is a challenging problem. A few simplifying assumptions are made: the type of protein is known, and its global orientation is known. Further, the protein is simplified so that we model it as a rigid chain of positions, i.e., we only keep track of the positions of certain atoms in the *backbone* of the protein. Due to these simplifications, we can view the conformation reconstruction problem as an *indirect* shape matching problem, in which a prior configuration of the protein (modelled as a set of relative positions in $\mathbb{R}^3$) is deformed to match the observed data (modelled as a function on $L^2(\mathbb{R}^2)$). This results in a deformed prior, which should match the conformation that resulted in the observed data. In practice, a vector v of N positions, i.e., $v \in \mathbb{R}^{3N}$ is compared to an observation $f \in L^2(\mathbb{R}^2)$ by means of a forward operator $\mathcal{F}\colon \mathbb{R}^{3N} \to L^2(\mathbb{R}^2)$, i.e., the similarity score between f and v is given by

$$d^2(f, v) = \|\mathcal{F}(v) - f\|^2_{L^2(\mathbb{R}^2)}.$$

In the words of Section 4.3, the shape space is thus $\mathbb{R}^{3N}$.

To ensure that the deformations are rigid, we restrict the deformations to the N-fold product of the group of rotations $\mathrm{SO}(3)$, i.e., $H = \mathrm{SO}(3)^N$, where each factor acts on a triplet of positions. The main contribution of the paper is to show that the protein conformation reconstruction problem is amenable to the shape matching framework, and to derive implementable gradient-based shape matching methods for the problem.

The method is illustrated on synthetic data in various numerical experiments.

5.7 Paper VII

We study the numerical signature of blow-up in hydrodynamical equations. Whether hydrodynamical equations blow up in finite time is usually hard to determine analytically, so numerical methods are often used to study the problem. For instance, for the three-dimensional Euler equations, numerics by Luo and Hou (2014) partially guided Chen and Hou (2021); Elgindi (2021). However, as noted by Fefferman (2022), it is often the case that numerical simulations might seemingly indicate blow-up, but that the blow-up is an artifact of the numerical method. This paper concerns the numerical signature of blow-up in the context of Zeitlin's method, the merits of which were discussed in Section 4.5. We study the complexified Euler equations, suggested by Sverak (2018) as a model equation since these equations, at least in the toroidal case, are not well-posed (Albritton and Ogden, 2023). The

complexified Euler equations are given by

$$\partial_t u + \nabla_{\overline{u}} u + \nabla_u \overline{u} = -\nabla p, \quad \operatorname{div} u = 0,$$

where u is a time-dependent complexified vector field, and p is a complexified pressure. The paper first contains a derivation of the complexified vorticity equations, as a Lie–Poisson system on the algebra $(C^\infty(\mathbb{S}^2, \mathbb{C}), \{\cdot,\cdot\}_{\mathbb{C}})$, where $\{\cdot,\cdot\}_{\mathbb{C}}$ is the complexified Lie–Poisson bracket. Then, Zeitlin's method is extended to the complexified case, and careful numerical experiments are performed to argue that

1. the observed blow-up is not an artifact of the numerical method, and
2. the time to blow-up decreases as the spatial resolution increases.

To argue for the first point, we show that the Casimirs of the complexified vorticity equations are conserved up to machine precision and that the reversibility of the spatially discrete system is preserved, i.e., by running the simulation backwards in time the initial condition is recovered. Finally, we exclude the possibility that the blow-up is due to the temporal discretization by varying the time step size, and see that the time to blow-up remains, for a fixed spatial resolution, within a certain range. To argue for the second point, we repeat the experiments with varying spatial resolution, and see that the time to blow-up decreases as the spatial resolution increases. Finally, we identify a possible signature of blow-up: we observe a linear relationship between the spatial resolution and the growth rate of the norms: the higher the spatial resolution, the faster the norms grow.

Bibliography

R. V. Abramov and A. J. Majda, Statistically relevant conserved quantities for truncated quasigeostrophic flow, *Proceedings of the National Academy of Sciences* **100** (2003), 3841–3846.

D. Albritton and W. J. Ogden, Remarks on the complex Euler equations, *Communications on Pure and Applied Analysis* **23** (2023), 1407–1422.

V. I. Arnold, Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits, *Annales de l'institut Fourier* **16** (1966), 319–361.

V. I. Arnold, *Mathematical Methods of Classical Mechanics*, vol. 60 of *Graduate Texts in Mathematics*, Springer New York, 1989.

F. Beg, M. Miller, A. Trouvé, and L. Younes, Computing large deformation metric mappings via geodesic flows of diffeomorphisms, *International Journal of Computer Vision* **61** (2005), 139–157.

J.-D. Benamou and Y. Brenier, A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem, *Numerische Mathematik* **84** (2000), 375–393.

A. Bistoquet, J. Oshinski, and O. Skrinjar, Myocardial deformation recovery from cine MRI using a nearly incompressible biventricular model, *Medical Image Analysis* **12** (2008), 69–85.

D. Boffi, Finite element approximation of eigenvalue problems, *Acta Numerica* **19** (2010), 1–120.

A. Bonito and J. E. Pasciak, Numerical approximation of fractional powers of elliptic operators, *Mathematics of Compututation* **84** (2015), 2083–2110.

C.-E. Bréhier, D. Cohen, and T. Jahnke, Splitting integrators for stochastic Lie–Poisson systems, *Mathematics of Computation* **92** (2023), 2167–2216.

S. C. Brenner and L. R. Scott, *Mathematical Theory of Finite Element Methods*, vol. 15, 3rd ed., Springer, 2008.

M. Bruveris and D. D. Holm, Geometry of image registration: The diffeomorphism group and momentum maps, *Geometry, Mechanics, and Dynamics: The Legacy of Jerry Marsden*, vol. 73 of *Fields Institute Communications*, chap. 2, p. 19–56, Springer New York, 2015.

E. Celledoni, M. J. Ehrhardt, C. Etmann, R. I. McLachlan, B. Owren, C.-B. Schönlieb, and F. Sherry, Structure-preserving deep learning, *European Journal of Applied Mathematics* **32** (2021), 888–936.

C. Ceritoglu, X. Tang, M. Chow, D. Hadjiabadi, D. Shah, T. Brown, M. H. Burhanullah, H. Trinh, J. T. Hsu, K. A. Ament, D. Crocetti, S. Mori, S. H. Mostofsky, S. Yantis, M. I. Miller, and J. T. Ratnanather, Computational analysis of LDDMM for brain mapping, *Frontiers in Neuroscience* **7** (2013).

C. Chen and O. Öktem, Indirect image registration with large diffeomorphic deformations, *SIAM Journal of Imaging Science* **11** (2018), 575–617.

J. Chen and T. Y. Hou, Finite time blowup of 2D Boussinesq and 3D Euler equations with $C^{1,\alpha}$ velocity and boundary, *Communications in Mathematical Physics* **383** (2021), 1559–1667.

D. G. Dritschel, W. Qi, and J. B. Marston, On the late-time behaviour of a bounded, inviscid two-dimensional flow, *Journal of Fluid Mechanics* **783** (2015), 1–22.

G. Dziuk, Finite elements for the Beltrami operator on arbitrary surfaces, *Partial Differential Equations and Calculus of Variations*, vol. 1357 of *Lecture Notes in Math.*, chap. 6, pp. 142–155, Springer Berlin Heidelberg, 1988.

G. Dziuk and C. M. Elliott, Finite element methods for surface PDEs, *Acta Numerica* **22** (2013), 289–396.

T. M. Elgindi, Finite-time singularity formation for $C^{1,\alpha}$ solutions to the incompressible Euler equations on $\mathbb{R}^3$, *Annals of Mathematics* **194** (2021), 647–727.

C. L. Fefferman, Existence and smoothness of the Navier–Stokes equation, `https://www.claymath.org/wp-content/uploads/2022/06/navierstokes.pdf`, 2022.

J. P. Fortney, *A Visual Introduction to Differential Forms and Calculus on Manifolds*, Birkhäuser Cham, 2018.

I. Gallagher, Mathematical analysis of a structure-preserving approximation of the bidimensional vorticity equation, *Numerische Mathematik* **91** (2002), 223–236.

T. Gneiting, Strictly and non-strictly positive definite functions on spheres, *Bernoulli* **19** (2013), 1327–1349.

H. H. Goldstine, *A History of Numerical Analysis from the 16th through the 19th Century*, vol. 2 of *Studies in the History of Mathematics and Physical Sciences*, Springer New York, 1977.

J. F. Grcar, John von Neumann's analysis of Gaussian elimination and the origins of modern numerical analysis, *SIAM Review* **53** (2011), 607–682.

U. Grenander, *General Pattern Theory: A Mathematical Study of Regular Structures*, Oxford University Press, 1994.

P. Guttorp and T. Gneiting, Studies in the history of probability and statistics XLIX: On the Matérn correlation family, *Biometrika* **93** (2006), 989–995.

R. S. Hamilton, The inverse function theorem of Nash and Moser, *Bulletin of the American Mathematical Society* **7** (1982), 65–222.

M. S. Handcock and J. R. Wallis, An approach to statistical spatial-temporal modeling of meteorological fields, *Journal of the American Statistical Association* **89** (1994), 368–378.

L. Herrmann, A. Lang, and C. Schwab, Numerical analysis of lognormal diffusions on the sphere, *Stochastics and Partial Differential Equations: Analysis and Computations* **6** (2018), 1–44.

D. D. Holm, Variational principles for stochastic fluid dynamics, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **471** (2015), 20140963.

J. Hong, J. Ruan, L. Sun, and L. Wang, Structure-preserving numerical methods for stochastic Poisson systems, *Communications in Computational Physics* **29** (2021), 802–830.

J. Hoppe, Diffeomorphism groups, quantization, and SU(∞), *International Journal of Modern Physics A* **04** (1989), 5235–5248.

J. Hoppe and S.-T. Yau, Some properties of matrix harmonics on $\mathbb{S}^2$, *Communications in Mathematical Physics* **195** (1998), 67–77.

E. P. Hsu, *Stochastic Analysis on Manifolds*, vol. 38 of *Graduate studies in mathematics*, American Mathematical Society, 2002.

E. Jansson, *Geometric Discretization in Shape Analysis*, Licentiate thesis, Chalmers University of Technology, Göteborg, Sweden, 2022.

J. Jost, *Riemannian Geometry and Geometric Analysis*, Universitext, Springer Berlin Heidelberg, 2011.

I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, vol. 113 of *Graduate Texts in Mathematics*, Springer New York, 1998.

B. Khesin and R. Wendt, *The Geometry of Infinite-Dimensional Groups*, Springer Berlin Heidelberg, 2009.

A. A. Kirillov, *Lectures on the Orbit Method*, vol. 64 of *Graduate Studies in Mathematics*, American Mathematical Society, 2004.

R. H. Kraichnan, Inertial ranges in two-dimensional turbulence, *The Physics of Fluids* **10** (1967), 1417–1423.

A. Lang and M. Pereira, Galerkin–Chebyshev approximation of Gaussian random fields on compact Riemannian manifolds, *BIT Numerical Mathematics* **63** (2023), 51.

S. Larsson and V. Thomée, *Partial Differential Equations with Numerical Methods*, vol. 45 of *Texts in applied mathematics*, 1 ed., Springer Berlin Heidelberg, 2008.

J. M. Lee, *Introduction to Smooth Manifolds*, vol. 218 of *Graduate Texts in Mathematics*, 2nd ed., Springer New York, 2012.

J. M. Lee, *Introduction to Riemannian Manifolds*, vol. 176 of *Graduate Texts in Mathematics*, 2nd ed., Springer Cham, 2018.

J. Leslie, On a differential structure for the group of diffeomorphisms, *Topology* **6** (1967), 263–271.

Q. Li, L. Chen, T. Cheng, and W. E, Maximum principle based algorithms for deep learning, *Journal of Machine Learning Research* **18** (2018), 165:1–165:29.

M. Liao and L. Wang, Motion of a rigid body under random perturbation, *Electronic Communications in Probability* **10** (2005), 235–243.

F. Lindgren, H. Rue, and J. Lindström, An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach, *Journal of the Royal Statistical Society, Series B (Statistical Methodology)* **73** (2011), 423–498.

E. Luesink, S. Ephrati, P. Cifani, and B. Geurts, Casimir preserving stochastic Lie–Poisson integrators, *Advances in Continuous and Discrete Models* **2024** (2024), 1.

G. Luo and T. Y. Hou, Toward the finite-time blowup of the 3D axisymmetric Euler equations: A numerical investigation, *Multiscale Modeling and Simulation* **12** (2014), 1722–1776.

J. E. Marsden and T. Ratiu, *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems*, vol. 17 of *Texts in Applied Mathematics*, 2nd ed., Springer New York, 1999.

B. Matérn, *Spatial Variation: Stochastic Models and Their Application to Some Problems in Forest Surveys and Other Sampling Investigations*, Ph.D. thesis, Statens Skogsforskningsinstitut, 1960.

G. N. Milstein and M. V. Tretyakov, *Stochastic Numerics for Mathematical Physics*, Scientific Computation, Springer Berlin Heidelberg, 2004.

K. Modin, Geometry of matrix decompositions seen through optimal transport and information geometry, *Journal of Geometric Mechanics* **9** (2017), 335–390.

K. Modin, Geometric hydrodynamics: from Euler, to Poincaré, to Arnold, 2019, `arXiv:1910.03301`.

K. Modin and M. Perrot, Eulerian and Lagrangian stability in Zeitlin's model of hydrodynamics, *Communications in Mathematical Physics* **405** (2024).

K. Modin and M. Viviani, A Casimir preserving scheme for long-time simulation of spherical ideal hydrodynamics, *Journal of Fluid Mechanics* **884** (2019a).

K. Modin and M. Viviani, Lie–Poisson methods for isospectral flows, *Foundations of Computational Mathematics* **20** (2019b), 889–921.

K. Modin and M. Viviani, Two-dimensional fluids via matrix hydrodynamics, arXiv:2405.14282, 2024.

P. J. Morrison and J. M. Greene, Noncanonical Hamiltonian density formulation of hydrodynamics and ideal magnetohydrodynamics, *Physical Review Letters* **45** (1980), 790–794.

P. K. Newton, The fate of random initial vorticity distributions for two-dimensional Euler equations on a sphere, *Journal of Fluid Mechanics* **786** (2015), 1–4.

F. Otto, The geometry of dissipative evolution equations: the porous medium equation, *Communications in Partial Differential Equations* **26** (2001), 101–174.

H. Poincaré, Sur une forme nouvelle des équations de la mécanique, *Comptes Rendus de l'Académie des Sciences* **132** (1901), 369–371.

E. Porcu, M. Bevilacqua, R. Schaback, and C. J. Oates, The Matérn model: A journey through statistics, numerical analysis and machine learning, *Statistical Science* **39** (2024), 469–492.

A. Qiu, C. Fennema-Notestine, A. M. Dale, and M. I. Miller, Regional shape abnormalities in mild cognitive impairment and Alzheimer's disease, *Neuroimage* **45** (2009), 656–661.

L. Risser, F.-X. Vialard, H. Y. Baluwala, and J. A. Schnabel, Piecewise-diffeomorphic image registration: Application to the motion estimation between 3D CT lung images with sliding conditions, *Medical Image Analysis* **17** (2013), 182–193.

A. Schmeding, *An Introduction to Infinite-Dimensional Differential Geometry*, vol. 202 of *Cambridge Studies in Advanced Mathematics*, Cambridge University Press, 2022.

G. R. Stibitz and J. A. Larrivee, *Mathematics and Computers*, 1st ed., McGraw-Hill, New York, 1957.

G. Strang and G. Fix, *An Analysis of the Finite Element Method*, Wellesley-Cambridge Press, 2008.

R. S. Strichartz, Analysis of the Laplacian on the complete Riemannian manifold, *Journal of Functional Analysis* **52** (1983), 48–79.

V. Sverak, On certain models in the PDE theory of fluid flows, *Journées équations aux dérivées partielles* (2018), 1–26.

M. E. Taylor, *Partial Differential Equations I: Basic Theory*, vol. 115 of *Applied Mathematical Sciences*, Springer New York, 2011.

D. Thompson, *On Growth and Form*, Cambridge University Press, 1992.

H. Triebel, Spaces of Besov–Hardy–Sobolev type on complete Riemannian manifolds, *Arkiv för Matematik* **24** (1985), 299–337.

C. Villani, *Optimal Transport: Old and New*, vol. 338 of *Grundlehren der mathematischen Wissenschaften*, Springer Berlin Heidelberg, 2009.

J. von Neumann and H. H. Goldstine, Numerical inverting of matrices of high order, *Bulletin of the American Mathematical Society* **53** (1947), 1021–1099.

G. Wanner, Kepler, Newton and numerical analysis, *Acta Numerica* **19** (2010), 561–598.

P. Whittle, Stochastic processes in several dimensions, *Bull. Inst. Int. Stat.* **40** (1963), 974–994.

L. Younes, *Shapes and Diffeomorphisms*, vol. 171 of *Applied Mathematical Sciences*, Springer Berlin Heidelberg, 2010.

L. Younes, B. Gris, and A. Trouvé, Sub-Riemannian methods in shape analysis, *Handbook of Variational Methods for Nonlinear Geometric Data*, chap. 17, pp. 463–495, Springer Cham, 2020.

V. Zeitlin, Finite-mode analogs of 2D ideal hydrodynamics: Coadjoint orbits and local canonical structure, *Physica D: Nonlinear Phenomena* **49** (1991), 353–362.

V. Zeitlin, Self-consistent finite-mode approximations for the hydrodynamics of an incompressible fluid on nonrotating and rotating spheres, *Physical Review Letters* **93** (2004).

B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*, Universitext, 6th ed., Springer Berlin Heidelberg, 2003.