

# Magnetohydrodynamics Simulation via Discrete Exterior Calculus

Mark Gillespie

June 8, 2018

# Acknowledgements

First, I would like to thank my advisor Peter Schröder. I feel incredibly lucky to have the opportunity to work with you. In addition to teaching me an amazing amount about computer science in the last year, you showed me how deep concepts from math and physics can play a central role in computer science research. I am inspired by the subtle connections that you find between seemingly-unrelated ideas. I am grateful for the endless patience you showed me as I got stuck time and time again, and for all of the advice you have given me, both about my research and about academic life in general.

I would like to thank Mathieu Desbrun. You have taught me an incredible amount about computer science and mathematics. Perhaps more importantly, you taught me that I don't have to be too serious, and that research can be a fun and funny endeavor. Furthermore, I am deeply grateful for all the advice you gave me throughout my process of applying to grad school; it made the difficult process much more manageable.

I would like to thank Alan Barr. You showed me that there is much more to computer graphics than simply making pictures (although you taught me plenty about making pictures too). I have really enjoyed all of our conversations over the past three years, about everything from splines to spaceflight to *Bedtime for Bonzo*. I appreciate all of the advice you have given me as I have navigated the world of Caltech and tried to figure out what I want to do with my life. It has been indispensable.

I would like to thank Ulrich Pinkall, Felix Knöppel, and Albert Chern. I am amazed by how much math I learned while talking to you over just a few weeks this summer. It was inspirational to see you leverage your deep understanding of geometry to gain insight into so many varied topics.

And, of course, I would like to thank my family, who always encouraged my love of learning and my love of art. Without you, I could never have made it here.

# Abstract

Electrically conducting fluids appear throughout nature, from the center of the Earth to the surface of the Sun. Simulations of magnetohydrodynamics (MHD) in the earth's core have provided key insight into the earth's changing magnetic field. And understanding the interplay between magnetic fields and fluid flow in the Sun is essential to understanding solar phenomena such as sun spots and coronal mass ejection. We present an integrator for ideal MHD in a two-dimensional domain with boundary and prove that the integrator preserves total energy and cross helicity.

# Contents

<b>Introduction</b>	<b>vii</b>
<b>I Background</b>	<b>1</b>
<b>1 Mathematical Preliminaries</b>	<b>2</b>
1.1 Manifolds . . . . .	2
1.2 Tangent Spaces . . . . .	3
1.3 Vector Bundles . . . . .	7
1.4 Differential Forms . . . . .	8
1.5 Lie Derivative . . . . .	15
1.6 Jet Bundles . . . . .	16
<b>2 Physical Preliminaries</b>	<b>18</b>
2.1 Variational Mechanics . . . . .	18
2.1.1 Motivation . . . . .	18
2.1.2 Lagrangian Mechanics . . . . .	18
2.1.3 Lagrangian Mechanics and Differential Geometry . . . . .	21
2.1.4 Euler Lagrange Equations . . . . .	22
2.1.5 Lagrangian Flow . . . . .	24
2.1.6 Conservation of the Symplectic Form . . . . .	24
2.1.7 Noether's Theorem . . . . .	25
2.2 Variational Field Theories . . . . .	25
2.2.1 Introduction . . . . .	25
2.2.2 Geometric Setup . . . . .	26
2.3 Maxwell's Equations . . . . .	27

2.3.1	Maxwell's Equations in Vector Calculus . . . . .	27
2.3.2	Working in Spacetime . . . . .	29
2.3.3	Covariant formulation of Maxwell's equations . . . . .	30
2.3.4	Ideal Fluids . . . . .	31
2.4	Magnetohydrodynamics . . . . .	32
2.4.1	Equations of Motion . . . . .	32
2.4.2	Conservation Laws . . . . .	33
2.4.3	Magnetic Helicity . . . . .	34
2.4.4	2D MHD . . . . .	35
2.5	Formal Lagrangians . . . . .	37
<b>3</b>	<b>Computational Preliminaries</b>	<b>39</b>
3.1	Introduction . . . . .	39
3.2	Discrete Exterior Calculus . . . . .	39
3.2.1	Discrete Differential Forms . . . . .	39
3.2.2	Discrete Exterior Derivative . . . . .	40
3.2.3	Discrete Hodge Star . . . . .	41
3.2.4	Discrete Inner Product . . . . .	42
3.2.5	Discrete Codifferential . . . . .	43
3.2.6	Discrete Wedge Product . . . . .	43
3.2.7	Explicit Expressions on a Regular Grid . . . . .	43
3.3	Simulation . . . . .	44
3.3.1	Euler Integration . . . . .	45
3.3.2	Discrete Lagrangian Mechanics . . . . .	45
<b>II</b>	<b>Results</b>	<b>47</b>
<b>4</b>	<b>Discretization</b>	<b>48</b>
4.1	Discrete Interior Product . . . . .	48
4.2	Solving the Discrete Equations . . . . .	49
4.3	Grids with Boundary . . . . .	49
<b>5</b>	<b>Conservation Behavior</b>	<b>51</b>

5.1	Energy . . . . .	51
5.2	Cross-Helicity . . . . .	52
<b>6</b>	<b>Computational Examples</b>	<b>53</b>
6.1	Alfvén Wave . . . . .	53
6.2	MHD with boundary . . . . .	53
6.2.1	Plume . . . . .	54
	<b>Appendices</b>	<b>65</b>
<b>A</b>	<b>Lagrange Multiplier Example: Pendulum</b>	<b>66</b>
<b>B</b>	<b>Exterior Calculus Identities</b>	<b>68</b>
B.1	Conversion between forms on the plane and in $\mathbb{R}^3$ . . . . .	69
<b>C</b>	<b>Formal Lagrangians</b>	<b>71</b>
C.1	Example of a mechanical system . . . . .	71
<b>D</b>	<b>Discrete Interior Product</b>	<b>73</b>
<b>E</b>	<b>Derivation of the Jacobian</b>	<b>75</b>
E.1	Velocity Equation . . . . .	75
E.2	Magnetic 1-Form Equation . . . . .	76
E.3	Divergence-Free Constraint . . . . .	77
<b>F</b>	<b>Assorted Vector and Matrix Derivatives</b>	<b>78</b>
F.1	Vector Derivatives . . . . .	78
F.2	Matrix Derivatives . . . . .	78

# Introduction

Magnetohydrodynamics(MHD) is the study of plasma and other conducting fluids. In everyday life, we do not often encounter plasma. But the physics of plasma governs countless important natural phenomena in our universe. The interstellar medium, the dispersed matter between different solar systems in a galaxy, is one such example. The interstellar medium forms enormous structures, hundreds of light years across, whose dynamics are governed by the laws of MHD. The behavior of the interstellar medium is an essential part of the formation of new stars [1]



Figure 1: View of the central 50 pc ( 150 light years) of the Galactic center showing ionized gas (traced by Paschen alpha emission observed by Hubble Space Telescope), hot plasma (traced in the X-ray by the Chandra X-ray Observatory) and warm dust (traced by mid-infrared radiation by the Spitzer Space Telescope).[2]

MHD also plays a key role in the behavior of the sun. The surface of the sun is a giant sea of plasma which has its own fascinating dynamics. Strong magnetic fields can build up in this plasma, which fling enormous amounts of energy and matter out of the sun in solar flares and coronal mass ejections. A solar flare can release  $10^{32}$  ergs of energy in a matter of hours. These solar flares and mass ejections are hugely complicated phenomena whose behavior depends on the interactions between twisted magnetic fields and the fluids flowing around on the sun's surface[3].

Closer to home, MHD also describes the behavior of the earth's core. At the center of the earth lies a solid iron inner core, surrounded by a liquid iron outer core. The earth's

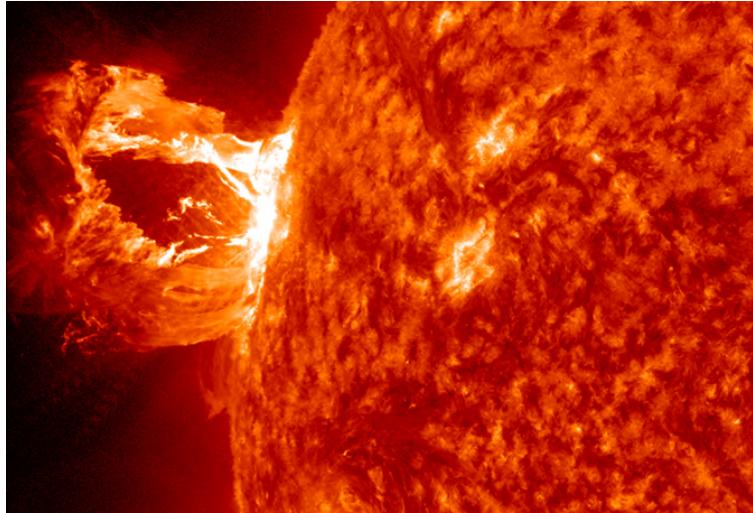


Figure 2: Coronal mass ejection photographed by NASA [4]

rotation causes various currents and whirlpools in the outer core, which in turn generate the Earth's magnetic field. MHD is essential to understanding how this liquid core behaves, which allows us to study how the Earth's magnetic field changes over time. In particular, MHD simulations allow scientists to study reversals of the magnetic field, the long periods of time over which the Earth's magnetic north and south poles switch.

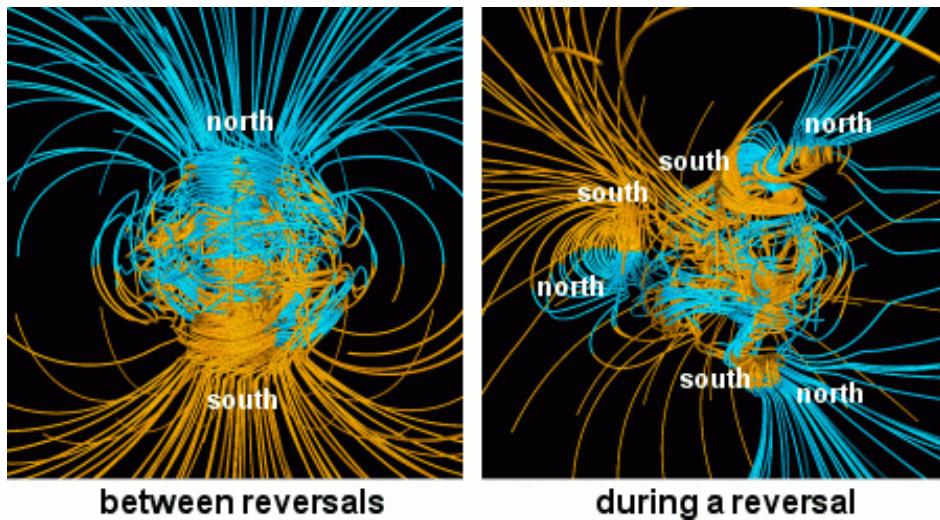


Figure 3: Simulation of the Earth's magnetic field [5]

In addition to being useful for pure scientific applications, MHD simulation is an essential tool in the design of nuclear fusion reactors. Fusion reactors make use of plasma contained in strong magnetic fields, so understanding and simulating the behavior of plasma in magnetic fields is an important part of research into fusion reactors [6].

Because of the numerous interesting applications of MHD, there is significant interest in numerical simulations of MHD. In particular, many researchers are interested in structure-preserving integrators for MHD [7], [8], [9], [10], [11]. Structure-preserving integrators are a

class of simulation methods which use deep principles and theorems from physics to derive simulations which preserve important properties of the relevant physical systems, and thus promise to generate physically-plausible long-term behavior of the simulation.

In this thesis, we will develop the mathematical machinery necessary required to understand the integrator presented by Kraus and Maj in [10]. Then, we will prove some conservation laws for the integrator, and extend the integrator to 2D systems with boundaries.



# Part I

## Background

# Chapter 1

## Mathematical Preliminaries

In this chapter, we present the mathematical tools of differential geometry. All the material (and much more) can be found in [12] or [13].

### 1.1 Manifolds

Differential calculus gives us powerful tools for understanding functions  $\mathbb{R}^n \rightarrow \mathbb{R}^m$ . However, we frequently want to study functions on more general spaces. For one example, consider a pendulum swinging in two dimensions. Such a pendulum can be described by its angle, which we can identify with a point on the unit circle  $S^1$ . So to study the physics of such a pendulum, we need to understand the behavior of functions to and from the circle.

Luckily for us, differential calculus only depends on the local behavior of functions. When taking the derivative of a function at a point, only needs to know the function's values infinitesimally close to that point. So we can actually apply differential calculus to study functions on all sorts of spaces, as long as they locally “look like”  $\mathbb{R}^n$ . The concept of a manifold formalizes this notion of a space which “looks like”  $\mathbb{R}^n$ .

A *chart* of a manifold  $M$  is a homeomorphism  $\varphi : U \rightarrow V \subseteq \mathbb{R}^n$  from a neighborhood in  $M$  to a subset  $V$  of  $\mathbb{R}^n$ . A chart tells us precisely how the neighborhood  $U$  looks like part of  $\mathbb{R}^n$ . An *atlas* is a collection of charts whose domains cover  $M$  where the charts all satisfy a technical compatibility condition. We call an atlas maximal if it is not properly contained in any other atlas (i.e. we could not add any more charts without violating the compatibility condition).

**Definition 1.1.1.** A (smooth) manifold is a Hausdorff, second-countable topological space equipped with a maximal atlas.

## 1.2 Tangent Spaces

Manifolds have a lot of nice structure, but they are nowhere near as structured as  $\mathbb{R}^n$  itself. Notably,  $\mathbb{R}^n$  has an exceedingly useful linear structure which general manifolds lack. Because linearity is so useful, it is often helpful to consider linear approximations to manifolds. For manifolds immersed in Euclidean space, we can obtain a linear approximation to a point  $p$  in the manifold by considering the hyperplane tangent to the manifold at  $p$ . This definition of linear approximations provides useful intuition about how linear approximations to manifolds behave. In fact, we usually call linear approximations to manifolds *tangent spaces* because they behave like the tangent planes to manifolds immersed in Euclidean space.

Now that we have some idea of how tangent spaces should behave, we need to actually define tangent spaces for abstract manifolds. This is a fairly tricky task. Indeed, there are several different constructions that are commonly used. I'll sketch two of them here. The constructions are a little bit abstract, so it is helpful to think about simple examples when following them. In particular,  $\mathbb{R}^n$  itself is a smooth manifold, and its linear approximation should be  $\mathbb{R}^n$  itself. When thinking about the construction of abstract tangent spaces, it is helpful to think about why these constructions yield  $\mathbb{R}^n$  as its own tangent space.

Now, let  $M^n$  be an  $n$ -dimensional manifold. So far, we just know that manifolds are particularly nice topological spaces. So there are really only two tools we have to understand  $M$  are functions into  $M$  or functions out of  $M$ . As it turns out, we can use either to construct tangent spaces to  $M$ . First, we'll construct tangent spaces using functions into  $M$ .

### Tangent vectors as equivalence classes of curves

To get a sense of how this construction should work, consider the case of the two-dimensional sphere  $S^2$  embedded in  $\mathbb{R}^3$  as the unit sphere. If we take some curve  $f : \mathbb{R} \rightarrow S^2 \subseteq \mathbb{R}^3$ , then that curve's velocity vector at a point  $p$  is tangent to the curve, and thus lies in the tangent plane to  $S^2$  at  $p$ . It seems reasonable that every tangent vector to the sphere is the velocity vector for some curve. Now, we just need to determine when two curves determine the same velocity vector. Suppose two curves have different velocity vectors at  $p$ . Then their velocity vectors must differ in some coordinate. Now, consider the projection of the two curves onto that axis. Their velocity vectors must differ at the point corresponding to  $p$ . Conversely, if we get the same tangent vector no matter which axis we project onto, then the two curves must have the same tangent vector. We can summarize this idea by saying that two curves  $\gamma_1, \gamma_2 : I \rightarrow S^2$  agree to first order at time  $t$  if  $\gamma_1(t) = \gamma_2(t)$  and  $(f \circ \gamma_1)'(t) = (f \circ \gamma_2)'(t)$  for any differentiable real-valued function  $f$  defined on a neighborhood of  $\gamma_1(t)$ .

With this example in mind, we can now define the tangent space to an abstract manifold. We can simply define a tangent vector to a point  $p \in M$  to be an equivalence class of curves passing through  $p$  subject to the equivalence relation defined above.

This definition of tangent spaces gives a geometrically-intuitive interpretation of tangent vectors. But the fact that one can equip this set of tangent vectors with a linear structure is somewhat surprising. The definition of addition on these vectors is somewhat involved.

## Tangent vectors as differential operators

The second construction of tangent spaces is more abstract and its geometric nature is less obvious. However, it has the advantage of giving our tangent vectors a straightforward linear structure. And it leads to some useful, but non-obvious, interpretations of tangent vectors which we will make use of later.

For this construction, we'll consider smooth real-valued functions on our manifold. Let's again consider the example of  $S^2 \subseteq \mathbb{R}^3$ . Given a real-valued function  $f : S^2 \rightarrow \mathbb{R}$  and a tangent vector  $v$  to  $S^2$  at  $p \in S^2$ , we can take a directional derivative  $D_v f(p)$  of  $f$  at  $p$  in the direction of  $v$ . Furthermore, if we are given the directional derivative operator, we can uniquely determine the corresponding vector  $v$  by testing out the differential operator on various functions  $f : S^2 \rightarrow \mathbb{R}$  (this is not obvious to me, but it turns out to be true).

This definition is fairly straightforward to extend to abstract manifolds. We begin with formalizing the idea of a first order differential operator. In particular, we define *derivations* on the algebra of real-valued functions  $C^\infty(M)$ .

**Definition 1.2.1.** Given a point  $p \in M$ , a *derivation at  $p$*  is an  $\mathbb{R}$ -linear map  $d : C^\infty(M) \rightarrow \mathbb{R}$  which satisfies the Leibniz rule

$$d(fg) = f(p) \cdot dg + df \cdot g(p) \quad (1.1)$$

Now, we simply define the tangent space at  $p$  of  $M$  to be the space of derivations at  $p$ . We denote this tangent space  $T_p M$ . From the definition of a derivation, we see that derivations have the structure of a real vector space. This allows us to use the tangent space as a linearization of our manifold near  $p$ . One can check that for a manifold of dimension  $n$ , the tangent space to any point will be an  $n$ -dimensional vector space.

## Pushforwards

Now, we have a nice way of taking linear approximations to manifolds near a point. In fact, we can go further. We can use these linear approximations to linearly approximate smooth functions between smooth manifolds! Given a smooth map  $\varphi : M \rightarrow N$  which sends  $m \in M$  to  $n \in N$ , we will construct a map called the *pushforward* or *differential* of  $\varphi$  at  $m$  which goes from  $T_m M$  to  $T_n N$ . This map is denoted  $\varphi_{*,m} : T_m M \rightarrow T_n N$ , or  $d\varphi_m : T_m M \rightarrow T_n N$ .

**Aside:** We will construct this map using the definition of the tangent space as a the space of derivations. One can also define the pushforward using the other definition of tangent vectors. Both are fairly simple.

Let  $v \in T_m M$  be a differential operator in  $T_m M$ . We want to use  $\varphi$  to turn it into a differential operator in  $T_n N$ . Let  $f : N \rightarrow \mathbb{R}$ . We want to apply  $v$  to  $f$ , but  $f$  is defined

on  $N$ , not  $M$ . We can fix this by precomposing  $f$  with  $\varphi$ . Then we obtain a function  $f \circ \varphi : M \rightarrow \mathbb{R}$ , and we can apply  $v$ . So we get a functional  $\varphi_{*,m}v : C^\infty(N) \rightarrow \mathbb{R}$  given by

$$\varphi_{*,m}v(f) := v(f \circ \varphi) \quad (1.2)$$

One can check that this is a derivation on  $N$  at  $n = \varphi(m)$ . It is clear that  $\varphi_{*,m}$  is a linear map. Thus,  $\varphi_{*,m}$  defines a linear map from  $T_m M$  to  $T_n N$ .

If we work in coordinates, the pushforward is the Jacobian of  $f$ .

### Local Coordinates on the Tangent Space

Let  $M$  be an  $n$ -dimensional manifold. Let  $p \in M$  be a point whose neighborhood  $U$  has a coordinate chart. We'll denote this chart by  $\varphi : V \rightarrow U$  where  $V \subseteq \mathbb{R}^n$ . For simplicity, assume that  $\varphi(0) = p$ . It turns out that not only does  $\varphi$  give us coordinates on  $U$ , it also give us a convenient basis for  $T_p M$ ! Recall that  $\mathbb{R}^n$  is its own tangent space at every point. In particular,  $T_0 \mathbb{R}^n \cong \mathbb{R}^n$ , and has a standard basis  $\{e_1, \dots, e_n\}$ . We can use the pushforward  $\varphi_{*,0}$  to send these basis vectors into  $T_p M$ . Since  $\varphi$  is injective,  $\{\varphi_{*,0}(e_1), \dots, \varphi_{*,0}(e_n)\}$  must be linearly independent in  $T_p M$ . Since  $T_p M$  is  $n$ -dimensional, they form a basis for  $T_p M$ .

Because we identify tangent vectors with differential operators, the standard basis of  $T_0 \mathbb{R}^n$  is often denoted  $\{\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}\}$  or simply  $\{\partial_1, \dots, \partial_n\}$ . The induced basis for  $T_p M$  is also frequently denoted by  $\{\frac{\partial}{\partial x_i}\}$  or  $\{\partial_i\}$ .

### Tangent Bundles

So far, we've focused on constructing tangent spaces to individual points on an abstract manifold. But viewing the tangent spaces at different points as entirely separate objects misses some of the structure that we see in the tangent planes to  $S^2 \subseteq \mathbb{R}^3$ . The tangent planes to an immersed manifold “vary smoothly” in some sense as you move around the manifold. We would like to capture this idea, and relate the separate tangent spaces of an abstract manifold to each other.

We can relate the separate tangent spaces together using a construction called the *tangent bundle*, denoted  $TM$  for a manifold  $M$ . As a set,  $TM$  is just the disjoint union of the tangent spaces to every point in  $M$ .

$$TM := \bigsqcup_{p \in M} T_p M \quad (1.3)$$

But  $TM$  has far more structure than just a set. We can actually turn it into a manifold! The charts on  $M$  induce charts on  $TM$  in a very natural way. Suppose  $\varphi : V \rightarrow U \subseteq M$  is a chart mapping a neighborhood  $V \subseteq \mathbb{R}^n$  to a neighborhood  $U$  in  $M$ . Recall that for any point  $x \in V$ ,  $\varphi_{*,x}$  is an isomorphism  $T_x \mathbb{R}^n \xrightarrow{\sim} T_{\varphi(x)} M$ . Since all of the tangent spaces to  $\mathbb{R}^n$  are canonically isomorphic to  $\mathbb{R}^n$  itself, we see that for each point  $x \in V$ , we have a map  $\varphi_{*,x} : \mathbb{R}^n \xrightarrow{\sim} T_{\varphi(x)} M$ . So we can define local coordinates on the neighborhood  $TU \subseteq TM$  defined as  $TU := \bigsqcup_{u \in U} T_u M$  using the chart

$$\begin{aligned} (\varphi, \varphi_{*,\cdot}) : V \times \mathbb{R}^n &\rightarrow TU \\ (x, v) &\mapsto (\varphi(x), \varphi_{*,x}(v)) \end{aligned} \quad (1.4)$$

One can check that these charts satisfy all of the conditions required to make  $TM$  into a smooth manifold.

A useful feature of tangent bundles is that we have a projection  $\pi : TM \rightarrow M$  given by sending any tangent vector to its basepoint.

Given a smooth map  $\varphi : M \rightarrow N$ , we can define the pushforward of  $\varphi$  as a map  $\varphi_* : TM \rightarrow TN$ . As one might expect,  $\varphi_*$  restricted to a particular tangent space  $T_p M$  is simply our old pushforward  $\varphi_{*,p}$ . One can check that the pushforward of a smooth map is itself a smooth map between the manifolds  $TM$  and  $TN$ .

**Aside:** There is a nice category-theoretic interpretation of this pushforward. We have a category of smooth manifolds **Man** whose objects are finite-dimensional smooth manifolds and whose morphisms are smooth maps. The existence of the pushforward tells us that operation of taking tangent bundles is functorial. That is to say, we have a functor  $\mathcal{T} : \mathbf{Man} \rightarrow \mathbf{Man}$  which takes the manifold  $M$  to  $TM$ , and takes the morphism  $\varphi : M \rightarrow N$  to  $\varphi_* : TM \rightarrow TN$ . People sometimes write  $\varphi_*$  as  $T\varphi$ , in which case our functor takes  $M$  to  $TM$  and  $\varphi$  to  $T\varphi$ .

Not only is  $\mathcal{T}$  a functor, but it is actually a monad! As the oft-cited definition goes, a monad is a monoid in the category of endofunctors. This means that a monad is a functor  $\mathcal{F} : \mathcal{C} \rightarrow \mathcal{C}$  from a category  $\mathcal{C}$  to itself along with a natural transformation from the identity functor to  $\mathcal{F}$ , and a natural transformation from  $\mathcal{F} \circ \mathcal{F}$  to  $\mathcal{F}$ , subject to some compatibility conditions.

Every tangent bundle has a zero section, which gives us a canonical map  $M \rightarrow TM$  for any manifold  $M$ . One can check that this gives a natural transformation from the identity functor to  $\mathcal{T}$ . And for any tangent bundle  $TM$ , one obtains a canonical map  $T(TM) \rightarrow TM$  by taking the pushforward  $\pi_* : T(TM) \rightarrow TM$  of the canonical projection  $\pi : TM \rightarrow M$ . Again, one can check that this defines a natural transformation from  $\mathcal{T} \circ \mathcal{T}$  to  $\mathcal{T}$ .

## Smooth Vector Fields

Often, we care not only about individual vectors on a manifold, but about *vector fields*, which assign a vector to each point on the manifold. Using the machinery we have built up so far, it is easy to define a smooth vector field.

**Definition 1.2.2.** A *smooth vector field* on a manifold  $M$  is a section of the tangent bundle  $TM$ . Explicitly, this means that a vector field is a map  $v : M \rightarrow TM$  such that  $\pi \circ v = \text{id}_M$  where  $\pi : TM \rightarrow M$  is the projection described above.

We denote the space of sections of  $TM$  by  $\Gamma(TM)$ . This is precisely the space of smooth vector fields on  $M$ .

### Riemannian Metrics

A Riemannian metric is a smooth assignment of an inner product to each tangent space  $T_x M$  of a manifold  $M$ . The precise meaning of smooth in this context will be clarified in the next section using the notion of a vector bundle.

A metric is useful because it allows us to talk about lengths of tangent vectors and angles between tangent vectors. This in turn allows us to talk about lengths of curves on the manifold, and angles between curves.

### Flows

One useful concept associated with vector fields is the idea of a flow. Given a vector field  $V \in \Gamma(TM)$ , we can think of  $V$  as a velocity field on  $M$ . A natural question to ask is where a particle would end up if it moved along this velocity field. This gives us a family of functions  $\varphi_t : M \rightarrow M$ , where  $\varphi_t(x)$  is where the particle starting at  $x$  ends up after flowing along the velocity field for time  $t$ . Given any smooth vector field, we can define a flow (at least on small neighborhoods of  $M$ ), for some time interval  $(-\epsilon, \epsilon)$ .

### Lie Bracket

The definition of vector fields as first-order differential operators gives us an interesting operation on vector fields: we can compose them. This results in a second-order differential operator, which is not in general a vector field. However, it turns out that because mixed partial derivatives commute, the commutator of two vector fields must be another first-order differential operator (because the second-order parts cancel out). Thus, the commutator gives us an  $\mathbb{R}$ -bilinear operation which takes in two vector fields and returns another vector field. This operation is called the Lie bracket, and is denoted  $[X, Y]$ .

## 1.3 Vector Bundles

Earlier, we defined the tangent bundle  $TM$  to a manifold  $M$ , which attached to each point  $p \in M$  its tangent space  $T_p M$ . Often, it is convenient to consider attaching other vector spaces to each point on a manifold. This construction is called a *vector bundle*.

As with every other object we have considered so far, vector bundles are characterized by how they appear locally. Recall that we constructed charts for our tangent bundle of the form

$$(\varphi, \varphi_*) : V \times \mathbb{R}^n \rightarrow TM \tag{1.5}$$

These charts show that  $TM$  “locally looks like” the cartesian product of  $M$  with the vector space  $\mathbb{R}^n$ . Vector bundles capture this idea of “local cartesian products” between a manifold and a vector space.

In general, a vector bundle consists of a manifold  $E$  called the *total space* equipped with a projection onto another manifold  $B$  called the base space. We require that for every  $b \in B$ , the preimage of  $b$  under  $\pi$  is a copy of some fixed vector space  $V$ .  $\pi^{-1}(b)$  is called the *fiber of  $E$  over  $b$* . Since all of the fibers must be isomorphic vector spaces, we frequently refer to “the fiber”  $V$ .

To enforce the constraint the vector bundles must locally look like cartesian products, we require that for every  $b \in B$ , there is a neighborhood  $b \ni U \subseteq B$  such that  $\pi^{-1}(U)$  is homeomorphic to  $U \times V$  in a way that respects the linear structure on  $V$ . These maps are called *local trivializations* of the fiber bundle. If  $E$  is globally equal to  $B \times V$ , we say that  $E$  is a trivial bundle.

The vector bundle  $\pi : E \rightarrow B$  with fiber  $V$  is frequently denoted

$$\begin{array}{ccc} V & \hookrightarrow & E \\ & & \downarrow \pi \\ & & B \end{array}$$

Many operations on vector spaces can also be performed on vector bundles over a fixed manifold  $M$  by performing the operations on each fiber. Let  $E, F$  be two vector bundles over  $M$ .

**Direct sum** We define  $E \oplus F$  to be the vector bundle whose fiber over  $x \in M$  is  $E_x \oplus F_x$ .

**Tensor product** Similarly, we define  $E \otimes F$  to be the vector bundle whose fiber over  $x \in M$  is  $E_x \otimes F_x$ .

**Dual** We define  $E^*$  to be the vector bundle whose fiber over  $x \in M$  is  $(E_x)^*$ , the dual space to  $E_x$ .

Using these operations, we can build some interesting vector bundles out of the tangent bundle. Taking the dual of the tangent bundle, we get a vector bundle over  $M$  whose fiber over  $x$  is the space of all linear functionals on  $T_x M$ . We call this vector bundle the *cotangent bundle* and denote it by  $T^* M$ . Now, by taking tensor products of  $TM$  and  $T^* M$ , we can obtain *tensor bundles*, whose fibers are spaces of tensors.

Tensor bundles give us a convenient language to use when talking about smooth functions which take in vector fields or output vector fields. For example, we can view a metric  $g$  as a fiberwise nondegenerate and positive definite section of the tensor bundle  $T^* M \otimes T^* M$ .

## 1.4 Differential Forms

So far, we have focused on the application of differential calculus to manifolds. But it turns out that we can use the mathematical machinery we have defined so far to extend integral calculus to manifolds as well! We can perform integration on manifolds using *differential*

*forms.* At a high level, differential forms are just “mathematical objects that you can integrate”. The  $dx$  in the usual integral  $\int_a^b f(x) dx$  can be interpreted as a differential form, and indeed differential forms often take the form  $dx$ .

Intuitively, a differential form measures an “infinitesimal area” or “infinitesimal volume”. The idea is that to integrate over a curve, we cut the curve into infinitesimal pieces, then use our differential form to measure the length of each piece, and then sum all of the lengths together. To integrate over a surface, we first cut the surface into infinitesimal pieces, then use a differential form to measure the area of each infinitesimal piece, and sum them all together. We could imagine performing a similar procedure for higher-dimensional domains of integration.

Now, we need to formalize this idea of “infinitesimal volume”. Measuring volumes of curves objects is complicated. But we know that if we look closely at a manifold  $M$ , it looks like  $\mathbb{R}^n$ . In fact, the tangent space  $T_p M$  is a linear approximation to  $M$ , and tells us precisely how  $M$  looks like  $\mathbb{R}^n$  when we are very close to  $p$ . So instead of thinking about infinitesimal volumes in  $M$ , we can instead think about infinitesimal volumes in  $\mathbb{R}^n$ . We can specify a  $k$ -dimensional region by looking at the parallelepiped determined by  $k$  vectors. The parallelepiped’s volume is given by the determinant of these  $k$  vectors. So if we want to measure generalized infinitesimal volumes, we can take inspiration from the determinant. One notable feature of the determinant is that it is multilinear - if you double the length of one side of a parallelepiped, the volume doubles. Also, the determinant does not simply measure volume. Instead, it measures *signed volume*. If you swap the order of two vectors, you swap the sign of the determinant.

Using these two properties, we can create functions to measure infinitesimal  $k$ -dimensional volumes. First, we will work in some vector space  $V$ , and once we understand volume-measurement in  $V$ , we can apply the idea to manifolds. Like the determinant, we want our volume function to be multilinear function that takes in  $k$  vectors and returns a real scalar. That is to say, we need a multilinear function  $V^k \rightarrow \mathbb{R}$ . But not every such multilinear map is an acceptable volume measurement. We also want our map to be alternating. We denote the set of all alternating, multilinear maps from  $V^k \rightarrow \mathbb{R}$  by  $\Lambda^k(V)$ . Note that we can add together alternating, multilinear maps to obtain other alternating, multilinear maps. So  $\Lambda^k(V)$  is a vector space.

We can use this idea to measure infinitesimal  $k$ -dimensional volumes on manifolds. Given a manifold  $M$ , we can measure infinitesimal  $k$ -dimensional volumes at a point  $p$  using an element of  $\Lambda^k(T_p M)$ . As we did earlier with tangent spaces, we can construct a vector bundle over  $M$  by attaching to each point  $p \in M$  the vector space  $\Lambda^k(T_p M)$ . We denote this vector bundle  $\Lambda^k(TM)$ . Finally, we come to the definition of a differential  $k$ -form.

**Definition 1.4.1.** A *differential  $k$ -form* is a smooth section of the vector bundle  $\Lambda^k(TM)$ . We denote the set of all differential  $k$ -forms on  $M$  by  $\Omega^k(M)$ .

Intuitively, a differential  $k$ -form gives you a way of measuring infinitesimal volumes at each point  $p \in M$ , and it gives you these volume-measuring functions in a smooth way.

## Pullbacks

Just like we can push forward vector fields along a smooth map  $f : M \rightarrow N$ , we can pull back a differential form from  $N$  to  $M$ . We construct a pullback map  $f^* : \Omega^k(N) \rightarrow \Omega^k(M)$  as follows: suppose  $\omega \in \Omega^k(N)$ . We can use  $\omega$  to define a  $k$ -form on  $M$  by defining

$$(f^*\omega)(v_1, \dots, v_k) := \omega(f_*(v_1), \dots, f_*(v_k)) \quad (1.6)$$

**Aside:** This construction feels similar to our definition of the pushforward. Recall that to define the pushforward of a vector field, we view the vector field as a differential operator on functions, and define

$$(f_*v)(g) := v(g \circ f) \quad (1.7)$$

We can think of function composition as a pullback of functions. Given a real-valued function  $g : N \rightarrow \mathbb{R}$ , we can pull  $g$  back along  $f$  to obtain the function  $g \circ f : M \rightarrow \mathbb{R}$ . So we could write  $g \circ f$  as  $f^*g$ . If we do so, then our defining equation for the pushforward becomes

$$(f_*v)(g) := v(f^*g) \quad (1.8)$$

And this is exactly the same as our definition for the pullback of differential forms, except we have swapped upper and lower stars!

## 0-Forms

Given a vector space  $V$ ,  $\Lambda^0(V)$  is the space of multilinear real-valued functions which take in 0 arguments. Such a function must just be constant, so  $\Lambda^0(V)$  is simply  $\mathbb{R}$ . Thus, a 0-form on  $M$  is simply a smooth assignment of a real number to each point  $p \in M$ , which is just a smooth map  $M \rightarrow \mathbb{R}$ . So we can identify  $\Omega^0(M)$  with  $C^\infty(M)$ , the space of real-valued functions on  $M$ .

## 1-Forms

A function of 1 argument is trivially alternating, and it is multilinear as long as it is linear. So  $\Lambda^1(V)$  is the space of linear maps  $V \rightarrow \mathbb{R}$ . This is simply the dual space  $V^*$ . Thus,  $\Omega^1(M)$  is a vector bundle whose fiber above each point  $p$  is the dual space  $(T_p M)^*$ . So  $\Omega^1(M)$  is just the cotangent bundle  $T^*M$ .

The space of 1-forms is more complicated than the space of 0-forms, but it is still easy to find examples of 1-forms. Many important examples are constructed from scalar functions. Let  $f \in C^\infty(M)$ . Then  $df : TM \rightarrow T\mathbb{R}$ . But since the tangent space to  $\mathbb{R}$  at any point is canonically isomorphic to  $\mathbb{R}$ , we can instead view  $df$  and a function  $df : TM \rightarrow \mathbb{R}$ . This makes  $df$  into a 1-form.

Let  $M$  be an  $n$ -dimensional manifold. Given a chart for a neighborhood  $U \subseteq M$ , we have  $n$  coordinate functions  $x^1, \dots, x^n : U \rightarrow \mathbb{R}$ . Taking the differentials of these coordinate functions gives us a basis  $\{dx^1, \dots, dx^n\}$  for  $\Omega^1(U)$ , the space of 1-forms of  $M$  restricted to  $U$ . This basis interacts nicely with our induced basis on the space of local vector fields. Recall that a chart also gives us a basis  $\{\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}\}$  for local vector fields on  $U$ . These two induced bases are dual to each other, in the sense that

$$dx^i \left( \frac{\partial}{\partial x^j} \right) = \delta_{ij} \quad (1.9)$$

Another important source of 1-forms is smooth vector fields. Suppose we have a Riemannian metric  $g$ .  $g$  allows us to identify each tangent space with its dual

$$g_p : T_p M \xrightarrow{\sim} T_p^* M \quad (1.10)$$

Thus, our metric gives us a map  $\Gamma(TM) \rightarrow \Omega^1(M)$  from the space of smooth vector fields to the space of 1-forms. Traditionally, this map is denoted by the symbol  $\flat$  (“flat”). This is because in the usual coordinate notation, components of vectors are denoted with upper indices, and components of 1-forms are written with lower indices, so the  $\flat$  map lowers the indices. Writing this map explicitly, we have

$$X^\flat(Y) := \langle X, Y \rangle \quad (1.11)$$

Our metric also induces an inverse map  $\Omega^1(M) \rightarrow \Gamma(TM)$ , called  $\sharp$  (“sharp”) because it raises indices. This map is given by

$$\langle \omega^\sharp, Y \rangle := \omega(Y) \quad (1.12)$$

Together,  $\flat$  and  $\sharp$  are referred to as the musical isomorphisms.

## The Wedge Product

One nice feature of integrals is that we can compute integrals over multidimensional regions as iterated integrals each over a lower-dimensional regions. For example,

$$\int_{[0,1] \times [0,1]} x^2 + y^2 \, dA = \int_0^1 \int_0^1 x^2 + y^2 \, dx dy \quad (1.13)$$

This fact hints that we should be able to “multiply together” lower-dimensional differential forms to obtain higher-dimensional differential forms. It turns out that we can!

The product of differential forms is called the *wedge product*. There are several normalization conventions for the wedge product, but the basic idea is the same in any case. We will define the wedge product of forms  $\omega \in \Omega^m(V), \eta \in \Omega^n(V)$  to be the form  $\omega \wedge \eta \in \Omega^{m+n}(M)$  such that

$$(\omega \wedge \eta)(v_1, \dots, v_{m+n}) := \frac{1}{m!n!} \sum_{\sigma \in S_{m+n}} \text{sgn}(\sigma) \omega(v_{\sigma(1)}, \dots, v_{\sigma(m)}) \eta(v_{\sigma(m+1)}, \dots, v_{\sigma(m+n)}) \quad (1.14)$$

where  $S_{m+n}$  is the symmetric group of permutations on  $m + n$  elements.

The expression for the wedge product is messy, but there is a nice idea behind it. The wedge product needs to produce an alternating function of  $m + n$  arguments given an alternating function of  $m$  arguments and an alternating function of  $n$  arguments. The simplest way in which one might attempt to define  $\omega \wedge \eta$  is to plug in the first  $m$  arguments to  $\omega$  and the last  $n$  arguments to  $\eta$ . But this might not be alternating. To fix this problem, we just sum over all possible permutations of the arguments, using the appropriate signs.

## The Exterior Derivative

Earlier, we saw that one can take the derivative of a 0-form  $f$  to obtain a 1-form  $df$ . It turns out that we can generalize this procedure to obtain the *exterior derivative*. For any  $k$ , the exterior derivative gives us a map  $d : \Omega^k(M) \rightarrow \Omega^{k+1}(M)$  defined as

$$d(f(x) dx^{i_1} \wedge \cdots \wedge dx^{i_k}) = df \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \quad (1.15)$$

The exterior derivative satisfies several useful properties. Like all good derivatives, it satisfies a Leibniz rule. For all  $\omega \in \Omega^m(M), \eta \in \Omega^n(M)$ , we have

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^m \omega \wedge d\eta \quad (1.16)$$

**Definition 1.4.2.** A differential form  $\omega$  is called *closed* if  $d\omega = 0$ .

**Definition 1.4.3.** A differential form  $\omega$  is called *exact* if there exists some other differential form  $\eta$  such that  $\omega = d\eta$ .

It turns out that exterior derivative squares to 0 (i.e.  $d(d\omega) = 0$  for all differential forms  $\omega$ ). We can restate this fact by saying that all exact forms are closed.

## Stokes' Theorem

The exterior derivative of a differential form satisfies a useful identity, called *Stokes' Theorem*. Stokes' theorem says that for any differential  $k$ -form  $\omega \in \Omega^k(M)$ , and any  $(k+1)$ -dimensional submanifold (possibly with boundary)  $S \subseteq M$ , we have

$$\int_{\partial S} \omega = \int_S d\omega \quad (1.17)$$

where  $\partial S$  denotes the boundary of  $S$ . Integrating a differential form along a region's boundary is the same as integrating the differential form's derivative over the entire region. Stokes' theorem generalizes several theorems from multivariable calculus. For example, consider Green's theorem

$$\oint_{\partial S} F \cdot dr = \iint_S \nabla \times F \cdot \hat{n} dS \quad (1.18)$$

We'll explore exactly how to derive Green's theorem from Stokes' theorem later. For now, we will just observe that they look very similar: both relate the integral of a quantity over a region's boundary to the integral of the quantity's derivative over the region.

## The Interior Product

A differential form is a function that takes in vector fields. One natural operation to consider on functions is *partial function application*. Given a differential form  $\omega \in \Omega^k(M)$  and a vector field  $X \in \Gamma(TM)$ , we can construct a function that takes in  $k-1$  vector fields by just plugging  $X$  into  $\omega$  as the first input. We denote the new function  $\iota_X \omega$  and define it to be

$$\begin{aligned}\iota_X \omega : \Gamma(TM)^{k-1} &\rightarrow \mathbb{R} \\ \iota_X \omega : (v_1, \dots, v_{k-1}) &\mapsto \omega(X, v_1, \dots, v_{k-1})\end{aligned}\tag{1.19}$$

Since  $\omega$  is multilinear and alternating,  $\iota_X \omega$  must be multilinear and alternating as well. So  $\iota_X$  defines a map  $\Omega^k(M) \rightarrow \Omega^{k-1}(M)$ . We call  $\iota_X \omega$  the *interior product* of  $\omega$  with  $X$ .

## The Hodge Inner Product

A metric on a vector space  $V$  induces a metric on the dual space  $V^*$ , since the metric defines an isomorphism  $V \xrightarrow{\sim} V^*$ . Thus, a metric on  $M$  gives us a symmetric, nondegenerate, bilinear map

$$\Omega^1(M) \otimes \Omega^1(M) \rightarrow \Omega^0(M)\tag{1.20}$$

defined by applying the metric to the 1-forms on each fiber. We can use this product to define a product on  $k$ -forms, noting that (at least locally) the space of  $k$ -forms has a basis given by wedge products of a basis of 1-forms.

Then, we can define the product of two  $k$ -forms as

$$\begin{aligned}\Omega^k(M) \otimes \Omega^k(M) &\rightarrow \Omega^0(M) = C^\infty(M) \\ a_1 \wedge \cdots \wedge a_k, b_1 \wedge \cdots \wedge b_k &\mapsto \langle a, b \rangle := \det(\langle a_i, b_j \rangle)_{ij}\end{aligned}\tag{1.21}$$

If  $M$  is compact, we can integrate this function to obtain a scalar. This defines an inner product on  $\Omega^k(M)$ , called the *Hodge inner product*, which we denote with double angle brackets

$$\langle\langle a, b \rangle\rangle := \int_M \langle a, b \rangle \mu\tag{1.22}$$

where  $\mu$  is the volume form induced by the metric.

Note the distinction between single and double brackets on the inner product.  $\langle a, b \rangle$  is the scalar function on  $M$  given by the pointwise inner product of the forms  $a$  and  $b$ , whereas  $\langle\langle a, b \rangle\rangle$  is a single scalar.

We denote the norm induced by this inner product  $\|\eta\|$ . Again, we have a single vs double bar distinction.  $|\eta|$  is a scalar function on  $M$ , whereas  $\|\eta\|$  is a single scalar.

## The Hodge Star

Let  $V$  be an  $n$ -dimensional vector space. Then

$$\dim \Lambda^k V = \binom{n}{k} = \binom{n}{n-k} = \dim \Lambda^{n-k} V\tag{1.23}$$

Thus,  $\Lambda^k V \cong \Lambda^{n-k} V$ . But we do not necessarily have a canonical isomorphism between them.

A metric on  $M$  gives us an isomorphism between  $\Omega^k(M)$  and  $\Omega^{n-k}(M)$ . This isomorphism is called the *Hodge star*.

**Definition 1.4.4.** The Hodge star is the unique linear map

$$\star : \Omega^k(M) \rightarrow \Omega^{n-k}(M) \quad (1.24)$$

that obeys the identity

$$\alpha \wedge \star \beta = \langle \alpha, \beta \rangle \mu \quad (1.25)$$

where  $\mu$  is the volume form induced by the metric

Note that with this definition, we can write the Hodge inner product defined in the last section as

$$\langle\!\langle \alpha, \beta \rangle\!\rangle = \int_M \alpha \wedge \star \beta \quad (1.26)$$

## Multivariable Calculus

If our manifold has a metric, we can relate this machinery of differential forms back to vector calculus in a nice way. The musical isomorphisms allow us to translate between vector fields and 1-forms.

Let  $f \in C^\infty(M), v \in \Gamma(TM)$ . It turns out that we have the following identities

$$\nabla f = (df)^\sharp \quad (1.27)$$

$$\nabla \times v = (\star^{-1} dv^\flat)^\sharp \quad (1.28)$$

$$\nabla \cdot v = \star^{-1} d \star v^\flat \quad (1.29)$$

It is instructive to prove these identities by working in coordinates.

We can use these identities to translate facts about differential forms into the language of multivariable calculus and vice versa. For example, specializing to the case of 1-forms, the fact that  $d^2\omega = 0$  translates to the vector identity

$$\nabla \cdot \nabla \times v = 0 \quad (1.30)$$

One important example is the *Laplacian*. Recall that the Laplacian of a real-valued function  $f$  is defined to be

$$\Delta f := \nabla \cdot \nabla f \quad (1.31)$$

Translating to differential forms, this becomes

$$\Delta f = \star^{-1} d \star df \quad (1.32)$$

*codifferential* We call

$$\delta := \star^{-1} d \star \quad (1.33)$$

the *codifferential*. In terms of the codifferential, we can write the Laplacian of a scalar function as  $\Delta f = \delta df$ . Whereas our original Laplacian was only defined on scalar functions,  $\delta d$  is an operator that can be applied to any differential form. We could define  $\Delta = \delta d$  to be a Laplacian on differential forms.

% todo[inline]why is the Laplacian not just  $\delta d$ ? It turns out that it is better to opt for the more symmetric operator

$$\Delta := \delta d + d\delta \quad (1.34)$$

We call  $\delta d + d\delta$  the *Laplace-de Rham operator*

## 1.5 Lie Derivative

So far, we have defined several types of fields on manifolds. The simplest type of field is a scalar field, which is just a real-valued function on  $M$ . There are also vector fields, differential  $k$ -forms, and general sections of vector bundles. Often, we want to understand how these fields vary across the manifold. For scalar fields, we know that we can take a directional derivative at a point using a tangent vector, or we can use a vector field to differentiate the whole scalar field. We also have the exterior derivative of a differential form is also a type of derivative, but it is fairly different - it captures some general sense of how much the form is changing at a point, but does not tell us about how the form changes in a given direction. We do not yet know how to take directional derivatives of non-scalar fields. It turns out that there are several notions of derivative for these more complicated fields. One notion of differentiation is given by the *Lie derivative*.

First, we will consider the case of vector fields. The key difficulty in differentiating vector fields is that tangent vectors to different points live in separate tangent spaces, and we do not have a canonical way of moving vectors between different tangent spaces. For an  $n$ -dimensional manifold, every tangent space  $T_p M$  is  $n$ -dimensional, so all of the tangent spaces are isomorphic. However, there is no *canonical* isomorphism between any two tangent spaces, so there is no canonical way of comparing vectors in different tangent spaces.

The Lie derivative solves this problem in an interesting way. Suppose we have two vector fields  $X, Y \in \Gamma(TM)$  and a point  $p \in M$  at which we want to differentiate  $Y$  along  $X$ . We saw earlier that we can (at least locally) integrate the vector field  $X$  to obtain a flow  $\varphi_t : U \rightarrow M$  on a neighborhood of  $p$  such that this flow is a diffeomorphism onto its image. This flow gives us maps which take nearby points of  $M$  to  $p$ , and we can use these maps to identify tangent spaces nearby  $p$  to  $T_p M$  itself. Explicitly, we define the Lie derivative of  $Y$  along  $X$  at a point  $p$  to be

$$(\mathcal{L}_Y X)_p := \lim_{h \rightarrow 0} \frac{(\varphi_h^{-1})_*(Y_{\varphi_h(p)}) - Y_p}{h} \quad (1.35)$$

**Aside:** It turns out that the Lie derivative of vector fields is closely related to the

Lie bracket which we defined earlier. In fact

$$\mathfrak{L}_X Y = [X, Y] \quad (1.36)$$

We can define the Lie derivative of a differential form  $\omega$  along a vector field  $X$  at a point  $p$  in a very similar way. The only difference is that Lie derivatives pull back along maps instead of pushing forwards. So we define

$$\mathfrak{L}_X \omega := \lim_{h \rightarrow 0} \frac{(\varphi_h)^*(\omega_{\varphi_h(p)}) - \omega_p}{h} \quad (1.37)$$

### Cartan's Magic Formula

There is a convenient expression for the Lie derivative of a differential form in terms of the interior product and exterior derivative. Because it is so useful, it is called *Cartan's magic formula*

$$\mathfrak{L}_X \omega = \iota_X d\omega + d\iota_X \omega \quad (1.38)$$

## 1.6 Jet Bundles

In this section, we follow the treatment of jet bundles found in [14] and [15]. Later on, it will be convenient to define a sort of tangent space for arbitrary vector bundles. Let  $Y$  be a vector bundle over  $X$  with fiber  $V$ .

$$\begin{array}{ccc} F & \hookrightarrow & Y \\ & & \downarrow \pi \\ & & X \end{array}$$

The fiber over a point  $x \in X$  is denoted  $F_x$ . Let  $\varphi_1, \varphi_2$  be sections of  $Y$ . We say that  $\varphi_1$  and  $\varphi_2$  agree to first order at  $x$  if  $\varphi_1(x) = \varphi_2(x)$  and  $(\varphi_1)_*, (\varphi_2)_*$  are equal (as linear maps  $T_x X \rightarrow T_{\varphi_1(x)} Y$ ). We define the first jet bundle of  $Y$ , denoted  $J^1 Y$ , to be the space of all (local) sections modulo this equivalence relation.

There is another characterization of  $J^1 Y$  which is often helpful. We can identify  $J^1 Y$  with the vector bundle over  $Y$  whose fiber over  $y \in F_x$  is the space  $L(T_x X, V_y Y)$  where

$$V_y Y := \ker \pi_* = \{v \in T_y Y \mid \pi_*(v) = 0\} \quad (1.39)$$

is the fiber above  $y$  in the *vertical subbundle*  $VY \subseteq TY$  of vectors parallel to the fibers of  $Y$ .

**Aside:**  $J^1Y$  is not quite a vector bundle; instead it is an *affine bundle*. An affine bundle is like a vector bundle, except our fibers are now *affine spaces* instead of vector spaces.

Recall that an affine space is “a vector space where we have forgotten the identity”. Jet bundles are affine spaces rather than vector spaces, because different charts disagree about what it means for the pushforward of a section to be zero.

Given a section  $\varphi : X \rightarrow Y$ , its differential  $T_x\varphi : T_xX \rightarrow T_{\varphi(x)}Y$  is an element of the fiber  $J^1Y_{\varphi(x)}$ . So the map  $x \mapsto T_x\varphi$  defines a section of  $J^1Y$ . We call this section  $j^1\varphi$  and call it *the first jet prolongation of  $\varphi$* . In coordinates,

$$j^i\varphi : x^\mu \mapsto (x^\mu, \varphi^A(x^\mu), \partial_\nu \varphi^A(x^\mu)) \quad (1.40)$$

# Chapter 2

## Physical Preliminaries

### 2.1 Variational Mechanics

#### 2.1.1 Motivation

Newton's laws give us a picture of the world that is often intuitive, but it can be hard to see underlying structures of a physical system when viewing it from a Newtonian perspective. It often appears mysterious that physical systems conserve quantities such as energy or momentum. Lagrangian mechanics offers a dramatically different perspective on physics, and one of the benefits of this new perspective is that it makes the reasons for many conservation laws much easier to understand. This is useful to understand when building simulations, because if we understand where conservation laws come from, then we have a better chance of designing algorithms which preserve the conservation laws in the discrete setting.

#### 2.1.2 Lagrangian Mechanics

In Lagrangian Mechanics, we describe a physical system by defining a *Lagrangian*, a real-valued function that takes in a system's position and velocity. The Lagrangian lets us compute how a system evolves over time. Frequently, we take our Lagrangian to be kinetic energy minus potential energy.

$$\mathcal{L}(q, \dot{q}) = \frac{1}{2}m\dot{q}^2 - V(q) \tag{2.1}$$

We define the system's *action* by integrating the Lagrangian over time

$$S[q] := \int_{t_0}^{t_1} \mathcal{L}(q(t), \dot{q}(t)) dt \tag{2.2}$$

The action is a function which takes in trajectories of the system and returns real numbers. *Hamilton's Principle of Least Action* tells us that the trajectory the system takes is

a stationary point of the action. To find the stationary points, we imagine an infinitesimal variation  $\delta q$  of the path  $q$ , which leaves the endpoints fixed. The resulting variation of the action is

$$\delta S = \int_{t_0}^{t_1} \frac{\partial \mathcal{L}}{\partial q} \delta q + \frac{\partial \mathcal{L}}{\partial \dot{q}} \delta \dot{q} dt \quad (2.3)$$

It's a little bit unclear what  $\delta \dot{q}$  is supposed to mean, but we can resolve this problem by integrating by parts to move the time derivative off of the variation. Doing so, we find that

$$\delta S = \int_{t_0}^{t_1} \frac{\partial \mathcal{L}}{\partial q} \delta q - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} \delta q dt \quad (2.4)$$

$$= \int_{t_0}^{t_1} \left( \frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} \right) \delta q dt \quad (2.5)$$

At a stationary point of the action,  $\delta S$  needs to be zero for any  $\delta q$ . If we want the integral to be 0, we must have

$$\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} = 0 \quad (2.6)$$

at every time  $t \in [t_0, t_1]$ . Equation 2.6 is known as the Euler-Lagrange equation. This equation lets us determine how our physical system evolves over time.

### Example: Point Particle in a Gravitational Field

Consider a particle of mass  $m$  moving around in space near a stationary planet of mass  $M$ . Denote the position of the particle at time  $t$  by  $q(t) \in \mathbb{R}^3$ . For convenience, let's use a coordinate system centered at the planet's center. Then we can write the particle's gravitational potential energy as  $-\frac{GMm}{|q|}$ . The particle's kinetic energy is, as usual, given by  $\frac{1}{2}m|\dot{q}|^2$ . Our Lagrangian is thus

$$\mathcal{L}(q, \dot{q}) = \frac{1}{2}m|\dot{q}|^2 + \frac{GMm}{|q|} \quad (2.7)$$

The Euler-Lagrange equations tell us that

$$\frac{\partial \mathcal{L}}{\partial q} = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} \quad (2.8)$$

In this case,

$$\frac{\partial \mathcal{L}}{\partial q} = -\frac{GMmq}{|q|^3} \quad (2.9)$$

(where we have used equation F.2) to differentiate  $\frac{1}{|q|}$ ) Furthermore,

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} = \frac{d}{dt} m\dot{q} = m\ddot{q} \quad (2.10)$$

Thus, in this case the Euler-Lagrange equations tell us that

$$m\ddot{q} = -\frac{GMmq}{|q|^3} \quad (2.11)$$

which is just the standard equation for gravitational attraction.

### Example: Point Particle in a Potential

It's no coincidence that the Euler-Lagrange equations recover the equations of motion that you would get from Newton's laws in the previous problem. (Indeed, it would be bad if we got a different answer, as Newton's Laws do a good job at predicting the movement of planets). Here, we will show that for a particle moving in any potential, Lagrangian mechanics reduces to Newton's law. Again, consider a particle whose trajectory is given by  $q(t)$ . Suppose that the particle moves subject to some potential  $V(q)$ . Then, we can again write our Lagrangian

$$\mathcal{L}(q, \dot{q}) = \frac{1}{2}m|\dot{q}|^2 - V(q) \quad (2.12)$$

In this case, the Euler-Lagrange equations tell us that

$$m\ddot{q} = -\frac{\partial V}{\partial q} \quad (2.13)$$

The derivative of potential energy with respect to position is just force. So in this case, the Euler-Lagrange equations are the familiar equation  $F = ma$ .

**Aside:** You might recall the *Lagrange multipliers* method for solving differentiable constrained optimization problems. The idea is that unconstrained optimization of a differentiable function is generally easy. The extrema of a differentiable function must be stationary points (points where the derivative is zero), so you can just find all of the stationary points, and pick the one that extremizes your function.

At first, it seems like constrained optimization should be much harder, since we can't apply the derivative trick anymore. However, Lagrange multipliers give us a simple, formulaic way of converting certain constrained optimization problems into higher-dimensional unconstrained problems. Suppose we have an objective function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  and the constraint that some function  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  must be 0. This is commonly written

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & g(x) = 0 \end{aligned} \quad (2.14)$$

Luckily for us, instead of solving this optimization problem, we can define a Lagrangian

$$\Lambda(x, \lambda) := f(x) - \lambda \cdot g(x) \quad (2.15)$$

and look for its stationary points. It turns out that if  $x$  is a minimizer of  $f$  subject to our equality constraint, then there exists a vector  $\lambda$  such that  $(x, \lambda)$  is a stationary point of  $\Lambda$ [16]. In this way, we can turn our constrained optimization problem into the problem of finding critical points of a function, which was our original strategy for optimization anyway.

The idea of Lagrangian mechanics is that we can frame classical mechanics as an optimization problem on the space of possible trajectories of our system. The action gives us an *objective function* on the space of trajectories, and the principle of

least action tells us that stationary points are the trajectories that physical systems actually take. Viewing Lagrangian mechanics in the context of optimization, it seems natural to hope that one could apply the method of Lagrange multipliers to constrain physical systems. And in fact, this works! Consult Appendix A to see the case of a 2D pendulum worked out using Lagrange multipliers.

### 2.1.3 Lagrangian Mechanics and Differential Geometry

In the last section, we introduced Lagrangian mechanics and worked through some basic examples. But there was a lot of symbol-pushing, and the definitions of what exactly a Lagrangian or variation is were not entirely clear. In this section, we'll give more rigorous definitions of the concepts involved, and explore the geometric structure behind Lagrangian mechanics. We follow the geometric presentation of mechanics given in [16].

We have seen before that a Lagrangian is a function that takes in both positions and velocities of particles. This hints that we should think about tangent bundles, because an element of a tangent bundle specifies a point on some base manifold as well as a tangent vector at that point.

Let  $Q$  be a manifold whose points represent possible configurations of a physical system. We call  $Q$  a *configuration manifold*. For example, the configuration manifold for a 2D pendulum is the circle  $S^1$ , since the configuration of a pendulum is described by its angle. The configuration manifold for a particle moving in space is  $\mathbb{R}^3$ , since the particle's position is described by a point in  $\mathbb{R}^3$ .

We associate a *state space* to our configuration manifold, which is simply the tangent bundle  $TQ$ . We denote by  $\pi_Q$  the canonical projection  $TQ \rightarrow Q$ . The state space describes possible configurations and velocities of the physical system. For example, the state space of a 2D pendulum is the cylinder  $S^1 \times \mathbb{R}$ , which makes sense because the pendulum's angle is described by a point on the circle, but its angular velocity can be any real number. The state space of a particle moving in space is  $\mathbb{R}^6$  with three dimensions corresponding to its position and three dimensions corresponding to its velocity.

**Definition 2.1.1.** A *Lagrangian* is a function  $\mathcal{L} : TQ \rightarrow \mathbb{R}$  from the tangent bundle of our configuration manifold to the reals.

**Definition 2.1.2.** The *action* is a map  $S : C(Q) \rightarrow \mathbb{R}$  induced by the Lagrangian. It is given by

$$S[q] := \int_{t_0}^{t_1} \mathcal{L}(q(t), \dot{q}(t)) dt \quad (2.16)$$

Since the Euler-Lagrange equations take a time derivative of (a derivative of) our Lagrangian, we should expect the general theory to depend on second derivatives of our curves. Thus, we define the *second-order submanifold* of  $T(TQ)$  to be

**Definition 2.1.3.**

$$\ddot{Q} := \{w \in T(TQ) \mid (\pi_Q)_*(w) = \pi_{TQ}(w)\} \quad (2.17)$$

where  $(\pi_Q)_* : T(TQ) \rightarrow TQ$  is the pushforward of the canonical projection  $TQ \rightarrow Q$ , and  $\pi_{TQ}$  is the canonical projection  $\pi_{TQ} : T(TQ) \rightarrow TQ$ .

Using the definition of tangent vectors as equivalence classes of curves which agree to first order, we can think of  $TQ$  as the space of first derivatives of curves. Similarly, we can think of  $\ddot{Q}$  as the space of second derivatives of curves in  $Q$ , since elements of  $\ddot{Q}$  have the form  $((q, \dot{q}), (\dot{q}, \ddot{q}))$ .

### 2.1.4 Euler Lagrange Equations

As before, we will use the Lagrangian to predict how our physical system will change over time. So we need to understand the space of trajectories our system can take. We define  $C(Q)$  to be the space of all smooth paths  $q : [t_0, t_1] \rightarrow Q$  on our configuration space. The path space is infinite-dimensional, so it is not a manifold according to our definitions. But we can use the calculus of variations to treat  $C(Q)$  like an infinite-dimensional manifold.

A variation is the analogue of a tangent vector on  $C(Q)$ . Just like a tangent vector at  $q \in Q$  can be specified by a curve  $\gamma : (-\epsilon, \epsilon) \rightarrow Q$  mapping 0 to  $q$ , we will specify variations using maps  $\eta : (-\epsilon, \epsilon) \times [t_0, t_1] \rightarrow Q$ . Often, we will denote  $\eta(\lambda, t)$  by  $\eta^\lambda(t)$ .

Let  $\dot{\eta}(\lambda, t) := \frac{\partial}{\partial t}\eta(\lambda, t)$ . We define the variation of  $S$  with respect to  $\eta$  to be

$$\frac{d}{d\lambda} \Big|_{\lambda=0} S(\eta^\lambda) = \frac{d}{d\lambda} \Big|_{\lambda=0} \int_{t_0}^{t_1} \mathcal{L}(\eta(\lambda, t), \dot{\eta}(\lambda, t)) dt \quad (2.18)$$

$$= \int_{t_0}^{t_1} \frac{\partial}{\partial \lambda} \Big|_{\lambda=0} \mathcal{L}(\eta(\lambda, t), \dot{\eta}(\lambda, t)) dt \quad (2.19)$$

$$= \int_{t_0}^{t_1} \frac{\partial \mathcal{L}}{\partial q} \cdot \frac{\partial \eta}{\partial \lambda} \Big|_{\lambda=0} + \frac{\partial \mathcal{L}}{\partial \dot{q}} \cdot \frac{\partial}{\partial \lambda} \frac{\partial \eta}{\partial t} \Big|_{\lambda=0} dt \quad (2.20)$$

$$= \int_{t_0}^{t_1} \frac{\partial \mathcal{L}}{\partial q} \cdot \frac{\partial \eta}{\partial \lambda} \Big|_{\lambda=0} + \frac{\partial \mathcal{L}}{\partial \dot{q}} \cdot \frac{\partial}{\partial t} \frac{\partial \eta}{\partial \lambda} \Big|_{\lambda=0} dt \quad (2.21)$$

$$= \int_{t_0}^{t_1} \left[ \frac{\partial \mathcal{L}}{\partial q} - \frac{\partial}{\partial t} \frac{\partial \mathcal{L}}{\partial \dot{q}} \right] \cdot \frac{\partial \eta}{\partial \lambda} \Big|_{\lambda=0} dt - \left[ \frac{\partial \mathcal{L}}{\partial \dot{q}} \cdot \frac{\partial \eta}{\partial \lambda} \Big|_{\lambda=0} \right]_{t_0}^{t_1} \quad (2.22)$$

$$(2.23)$$

Interestingly, we see that this only depends on  $\frac{\partial \eta}{\partial \lambda} \Big|_{\lambda=0}$ . Note that this defines a vector field along  $q$ . We denote this vector field  $\delta q$ . It turns out that given such a vector field  $\delta q$ , we can always construct a variation map  $\eta$  such that  $\delta q = \frac{\partial \eta}{\partial \lambda} \Big|_{\lambda=0}$ . So from now on, we will think of variations as being vector fields along curves. We write the variation as  $dS_q[\delta q]$ . In our new notation,

$$dS_q[\delta q] = \int_{t_0}^{t_1} \left[ \frac{\partial \mathcal{L}}{\partial q^i} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}^i} \right] \delta q^i dt + \left[ \frac{\partial \mathcal{L}}{\partial \dot{q}^i} \delta q^i \right]_{t_0}^{t_1} \quad (2.24)$$

Now, we need to understand what these expressions mean in our geometric picture.  $\delta q$  is a variation, so at each  $t$ ,  $\delta q(t) \in T_{q(t)}Q$  is a tangent vector. Inside the integral, we apply a linear function to  $\delta q(t)$  (since we multiply each component by a scalar and sum them together), so we can think of  $(\frac{\partial \mathcal{L}}{\partial q^i} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}^i})_i$  as a cotangent vector in  $T^*Q$ . Since  $\mathcal{L}$  is a function on  $TQ$ , and we take one more time derivative, this is a function on  $\ddot{Q}$ . So we can think of this expression as a map

$$D_{EL}\mathcal{L} : \ddot{Q} \rightarrow T^*Q \quad (2.25)$$

with coordinate expression

$$(D_{EL}\mathcal{L})_i := \frac{\partial L}{\partial q^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} \quad (2.26)$$

We call  $D_{EL}\mathcal{L}$  the *Euler-Lagrange map*

Similarly, we can think of the last term  $\frac{\partial \mathcal{L}}{\partial \dot{q}^i} \delta q^i$  as a 1-form on  $TQ$ <sup>1</sup>. We denote this 1-form  $\Theta_L$  and call it the *Lagrangian 1-form*. In coordinates, we have

$$\Theta_L = \frac{\partial \mathcal{L}}{\partial \dot{q}^i} dq^i \quad (2.27)$$

In terms of our new maps, we see that the variation is given by

$$dS_q[\delta q] = \int_{t_0}^{t_1} D_{EL}\mathcal{L}(q, \dot{q}, \ddot{q})(\delta q) dt + (\Theta_L)_{\dot{q}}(\hat{\delta}q)|_{t_0}^{t_1} \quad (2.28)$$

where  $\hat{\delta}q \in TTQ$  is given by

$$\hat{\delta}q(t) = \left( (q(t), \dot{q}(t)), \left( \delta q(t), \frac{\partial \delta t}{\partial t} \right) \right) \quad (2.29)$$

Recall that Hamilton's principle of least action states that the trajectory of our physical system will be a stationary point of the action. So we call a curve  $q : [t_0, t_1] \rightarrow Q$  a solution if the variation  $dS_q[\delta q]$  is 0 for all variations  $\delta q$  which vanish at  $t_0$  and  $t_1$ . In order for this to be true, we must have

$$D_{EL}\mathcal{L}(q, \dot{q}, \ddot{q}) = 0 \quad (2.30)$$

for all  $t \in [t_0, t_1]$ . In coordinates, this is just the familiar Euler-Lagrange equations

$$\frac{\partial \mathcal{L}}{\partial q^i} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}^i} = 0 \quad (2.31)$$

---

<sup>1</sup>At first, you might expect  $\Theta_L$  to be a 1-form on  $Q$ . But because  $\Theta_L$  depends on  $\frac{\partial \mathcal{L}}{\partial \dot{q}^i}$ , which itself depends on both  $q$  and  $\dot{q}$ , we need  $\Theta_L$  to be a 1-form on  $TQ$ . This means that  $\Theta_L$  takes in vectors in  $TTQ$  and returns scalars.

### 2.1.5 Lagrangian Flow

The Euler-Lagrange equations only depend on  $q$  and its first and second derivatives. This means that the Euler-Lagrange equations depend only on the section of  $\ddot{Q}$  induced by  $q$  (i.e. the map  $q(t) \mapsto (q(t), \dot{q}(t), \ddot{q}(t))$ ). Since  $\ddot{Q}$  is a subbundle of  $TTQ$ , we can instead say that the Euler-Lagrange equations depend on the induced section of  $TTQ$ .<sup>2</sup> We can interpret this condition by saying that solutions of the Euler-Lagrange equations are integral curves of a certain vector field on  $TQ$ .

We define the *Lagrangian vector field*  $X_L$  on  $TQ$  to be the section  $TQ \rightarrow T(TQ)$  such that<sup>3</sup>

$$D_{EL}\mathcal{L} \circ X_L = 0 \quad (2.32)$$

Associated to this vector field, we have a flow  $F_L^t : TQ \rightarrow TQ$  which we call the *Lagrangian flow*. The Lagrangian flow describes how our physical system evolves over time, since a flow produces integral curves of its underlying vector field.

### 2.1.6 Conservation of the Symplectic Form

In physical systems governed by nice Lagrangians, the initial conditions uniquely determine the trajectory of the system. This gives us a map  $f : TQ \rightarrow C(Q)$  which takes a tangent vector (i.e. initial condition) to the path the system takes starting at that initial condition. Explicitly,  $f(v)(t) = \pi(F_L^t(v))$ . We can pull back the action functional along this map to obtain a *restricted action map*  $\hat{S} : TQ \rightarrow \mathbb{R}$ . Taking the exterior derivative of this function corresponds to taking the first variation of the action functional.<sup>4</sup> Because our curve  $f(q)$  satisfies the Euler-Lagrange equations, the first term of equation 2.28 is zero for all variations. So we only need to look at the second term. We can also pull back the Lagrangian 1-form  $\Theta_L \in \Omega^1(TQ)$  to a *restricted Lagrangian 1-form*  $\hat{\Theta}_L \in \Omega^1(Q)$ . (Note: for this computation, we assume that our trajectory maps  $[0, T] \rightarrow Q$ )

$$d\hat{S}_q(v_q) := dS_q[\delta q] \quad (2.33)$$

$$= (\Theta_L)_{f(q)(T)}(\hat{\delta}q) - (\Theta_L)_{f(q)(0)}(\hat{\delta}q) \quad (2.34)$$

$$= (\hat{\Theta}_L)_{F_L^T(q)}((F_L^T)_*v_q) - (\hat{\Theta}_L)_q(v_q) \quad (2.35)$$

$$= (F_L^T)^*(\hat{\Theta}_L)_q(v_q) - (\hat{\Theta}_L)_q(v_q) \quad (2.36)$$

Since exterior derivatives commute with pullbacks, we can take the exterior derivative of this expression and use the fact that  $d^2\hat{S} = 0$  to obtain

$$(F_L^T)^*d\Theta_L = d\Theta_L \quad (2.37)$$

---

<sup>2</sup>This is like viewing the second-order Euler-Lagrange equations as a set of first-order differential equations in  $q$  and  $\dot{q}$ .

<sup>3</sup>For some Lagrangians, equation 2.32 might not define a unique vector field. But in nice cases, this is not a problem.

<sup>4</sup>Given a tangent vector  $v_q \in T_q Q$ , we can use our flow map to extend  $v_q$  to a vector field on the curve  $f(q)$ . This is just a variation of  $f(q)$ ! We simply define  $\delta q(t) := (F_L^t)_*v_q$ . One can check that  $d\hat{S}_q(v_q) = dS_q[\delta q]$ .

We call  $\Omega_L := d\Theta_L$  the *Lagrangian symplectic form*. Equation 2.37 tells us that time-evolution preserves  $\Omega_L$ .

### 2.1.7 Noether's Theorem

Noether's theorem gives us a simple algorithm for taking in symmetries of a physical system and finding associated conserved quantities. In order to understand symmetries of a physical system, we must first discuss actions of Lie groups on our configuration space. Let  $G$  be a Lie group with Lie algebra  $\mathfrak{g}$ , and let  $\Phi : G \times Q \rightarrow Q$  be a group action of  $G$  on  $Q$ . Often, we will denote  $\Phi(g, q)$  by  $\Phi_g(q)$ . We can lift this action to an action  $\Phi^{TQ}$  of  $G$  on  $TQ$  by defining

$$\Phi_g^{TQ}(v) := (\Phi_g)_*(v) \quad (2.38)$$

We can consider ‘infinitesimal’ versions of this action. Thinking of the Lie algebra as infinitesimal elements of the Lie group, we might expect to be able to use our group action to obtain infinitesimal transformations (vector fields) of  $Q$  from vectors in our Lie algebra. Explicitly, given  $\xi \in \mathfrak{g}$ , we define *infinitesimal generators*  $\xi_Q \in \Gamma(TQ)$  and  $\xi_{TQ} \in \Gamma(T(TQ))$  by

$$\xi_Q(q) := \left. \frac{d}{dg} \right|_{g=e} \Phi(g, q) \quad (2.39)$$

$$\xi_{TQ}(v_q) := \left. \frac{d}{dg} \right|_{g=e} \Phi^{TQ}(g, v_q) \quad (2.40)$$

where  $g$  denotes an element of a 1-parameter subgroup corresponding to  $\xi$  (i.e.  $\left. \frac{d}{dg} \right|_{g=e} g = \xi$ ).

Now, we define the *Lagrangian momentum map* to be a map  $J_L : TQ \rightarrow \mathfrak{g}^*$  defined by

$$J_L(v_q)(\xi) := \Theta_L(\xi_{TQ}(v_q)) \quad (2.41)$$

It turns out that Lagrangian momentum maps which come from symmetry transformations are conserved. For a proof, see [16].

## 2.2 Variational Field Theories

### 2.2.1 Introduction

In the previous section, we assumed that the configuration of our physical system could be described by a point in a finite-dimensional manifold. However, there are many physical systems for which this is not the case. For example, when we use a continuum approximation, the state of a fluid is characterized by a diffeomorphism from the fluid’s container to itself. Alternatively, we could characterize a fluid by its velocity field (this is the approach we will take later on). In either case, our characterization of the fluid has infinitely many degrees off

freedom. So we cannot represent our physical system with a finite-dimensional configuration space.

It would be very difficult to handle an arbitrary infinite-dimensional configuration space. But luckily, many physical systems with infinite-dimensional configuration spaces have highly-structured configurations spaces. We can take advantage of this structure to understand these physical systems.

Both possible configuration spaces we considered for fluids are spaces of maps between finite-dimensional manifolds. In particular, both configuration spaces are spaces of sections of fiber bundles. We can view a diffeomorphism from a manifold  $M$  to itself as a section of the trivial vector bundle  $M \times M$ . And a fluid's velocity field is simply a vector field, which is a section of the tangent bundle  $TM$ . It turns out that there are many physical systems whose configurations can be understood as sections of vector bundles. We call these section *fields* and we call the study of such systems *field theory*. In the following sections, we will present the beginning a geometric viewpoint on classical field theory, following [15], [14], and [17].

### 2.2.2 Geometric Setup

Studying field theory variationally is tricky. In order to formalize all of the variations required to understand the systems, we need to parameterize the sections of our fiber bundle. This leads to a complicated geometric setup with many similar-seeming maps.

The basic object of a field theory is a vector bundle

$$\begin{array}{ccc} S & \hookrightarrow & Y \\ & & \downarrow \pi_{YX} \\ & & X \end{array}$$

Where the base space  $X = \mathbb{R} \times \mathbb{R}^n$  is called *spacetime*, the fiber (vector space attached to each point) is  $S = \mathbb{R}^m$  is called *ambient space*, and the total space  $Y$  is called the *configuration bundle*.

We pick a subset of space  $M \subseteq \mathbb{R}^n$ . We also define a *parameter space*  $\mathcal{U} = [t_0, t_1] \times M$ . Now, we define our configuration space to be

$$\mathcal{C} := \{\varphi : \mathcal{U} \rightarrow Y \mid \pi_{YX} \circ \varphi : \mathcal{U} \rightarrow X \text{ is an embedding}\} \quad (2.42)$$

**Aside:** If you want to be very formal, you can define the configuration space to be an infinite-dimensional manifold by taking the completion of  $\mathcal{C}$  with respect to some norm. But this won't be necessary for us.

Let us denote  $\pi_{YX} \circ \varphi : \mathcal{U} \rightarrow X$  by  $\varphi_X$  and  $\varphi_X(\mathcal{U})$  by  $U_X$ . Now  $\varphi_X : \mathcal{U} \rightarrow U_X$  is a diffeomorphism. Note that a function  $\varphi \in \mathcal{C}$  gives us a section  $\varphi \circ \varphi_X^{-1} : U_X \rightarrow Y$ .

Our field-theoretic analogue of the state space is the first jet bundle of  $Y$ ,  $J^1Y$ . We will define a Lagrangian as a function on  $J^1Y$ . Because we want to integrate our Lagrangian, it will be convenient for the output of the Lagrangian to be a differential form which we can integrate<sup>5</sup>.

**Definition 2.2.1.** A *Lagrangian* is a functional  $\mathcal{L} : J^1Y \rightarrow \Lambda^{n+1}(X)$ .

Recall that  $X$  is  $(n+1)$ -dimensional since time is the zeroth dimension.

As before, we integrate our Lagrangian to obtain an action functional  $S : \mathcal{C} \rightarrow \mathbb{R}$ .

**Definition 2.2.2.** The *action functional* is the map  $S : \mathcal{C} \rightarrow \mathbb{R}$  given by

$$S(\varphi) := \int_{U_X} \mathcal{L}(j^1(\varphi \circ \varphi_X^{-1})) \quad (2.43)$$

Using this definition of the field-theoretic Lagrangian and action, one can find analogues of all of the theorems and identities we saw before. But the computations get much more complicated. For more details, see [15], [14], or [17].

## 2.3 Maxwell's Equations

### 2.3.1 Maxwell's Equations in Vector Calculus

In CGS units, Maxwell's equations are given by

1.  $\nabla \cdot E = 4\pi\rho$
2.  $\nabla \cdot B = 0$
3.  $\nabla \times E = -\frac{1}{c} \frac{\partial B}{\partial t}$
4.  $\nabla \times B = \frac{4\pi}{c} J + \frac{1}{c} \frac{\partial E}{\partial t}$

$E$  is the electric field,  $B$  is the magnetic field,  $J$  is the electric current density, and  $\rho$  is the electric charge density.  $E, B$  and  $J$  are vector fields and  $\rho$  is a scalar field.

If we want to write Maxwell's equations with differential forms, we need to decide what type of forms will represent  $E, B, J$  and  $\rho$ . Following [18] we will decide this based on how these fields are used in various equations.

First, we consider Faraday's law

$$\oint_C E \cdot d\ell = -\frac{d}{dt} \int_S B \cdot dA \quad (2.44)$$

---

<sup>5</sup>This is basically just taking the Hodge star of our previous notion of Lagrangian

$E$  is integrated over a curve, so  $E$  naturally corresponds to a 1-form (Since one-forms are objects that you can integrate along curves). We will write the 1-form associated to  $E$  as  $\eta = E^\flat$ . Meanwhile  $B$  is integrated over a surface, so  $B$  naturally corresponds to a 2-form. We will write this 2-form as  $\beta = \star B^\flat$  as the 2-form associated to  $B$ .  $\beta$  can be thought of as a function that measures the flux of  $B$  through oriented parallelograms.  $J$ , the current density, is integrated over surfaces to find the current passing through the surface, so  $J$  is naturally a 2-form. We will write this forms as  $\mathcal{J} = \star J^\flat$ . Finally,  $\rho$  is integrated over volumes to find the enclosed charge, so  $\rho$  is naturally a 3-form, and we will call the 3-form  $\rho$ .

We recall the following rules for translating from vector calculus to differential forms:

$$(\nabla \cdot v)^\flat = \star d \star v^\flat \quad (2.45)$$

$$(\nabla \times v)^\flat = \star d v^\flat \quad (2.46)$$

We can use these rules to write Maxwell's equations in terms of  $\eta, \beta, \mathcal{J}$ , and  $\rho$ .

We will start with the first equation. On the left hand side,  $\nabla \cdot E$  becomes  $\star d \star \eta$ , which is a 0-form. We want to set it equal to the 3-form  $4\pi\rho$ . So we use the Hodge star to turn  $\star d \star \eta$  into a 3 form and find that  $\star \star d \star \eta = 4\pi\rho$ . In 3D,  $\star \star = 1$ , so our equation is just

$$d \star \eta = 4\pi\rho \quad (2.47)$$

Now, we move on to the second equation.  $\nabla \cdot B$  becomes  $\star d \star (\star \beta)$ . Since  $\star \star = 1$ , this just becomes  $\star d \beta$ . So our equation is  $\star d \beta = 0$ . Applying  $\star$  to both sides gives  $d\beta = 0$ .

$$d\beta = 0 \quad (2.48)$$

For the third equation, our substitutions gives us  $\star d\eta = -\frac{1}{c} \frac{\partial \star \beta}{\partial t}$ . Applying  $\star$  to both sides yields  $d\eta = -\frac{1}{c} \star \frac{\partial \star \beta}{\partial t}$ . But we can pull the  $\star$  inside the derivative to get

$$d\eta = -\frac{1}{c} \frac{\partial \beta}{\partial t} \quad (2.49)$$

Finally, we translate the last equation. Our substitution rules give us

$$\star d \star \beta = \frac{4\pi}{c} \star \mathcal{J} + \frac{1}{c} \frac{\partial \eta}{\partial t} \quad (2.50)$$

Putting all of the equations together, we have

$$d \star \eta = 4\pi\rho \quad (2.51)$$

$$d\beta = 0 \quad (2.52)$$

$$d\eta = -\frac{1}{c} \frac{\partial \beta}{\partial t} \quad (2.53)$$

$$\star d \star \beta = \frac{4\pi}{c} \star \mathcal{J} + \frac{1}{c} \frac{\partial \eta}{\partial t} \quad (2.54)$$

Now, we have written the equations using differential forms. But the equations still don't look very relativistic yet - we still have a big distinction between space and time

derivatives. For our next step, we will stop thinking about forms on space that change over time, and instead think about forms on 3 + 1-dimensional spacetime (i.e. spacetime with 3 spatial dimensions and 1 time dimension). In spacetime, we have both the spatial exterior derivative, the spatial Hodge star, the spacetime exterior derivative, and the spacetime Hodge star. We will denote the spatial operators  $d_s, \star_s$  respectively, and we will denote the spacetime operators  $d$  and  $\star$ .

### 2.3.2 Working in Spacetime<sup>6</sup>

Before we can write Maxwell's equations in spacetime, we have to learn about how our spatial operators are related to our spacetime operators. We will use the convention that coordinates in spacetime are written like

$$(x^0, x^1, x^2, x^3) = (ct, x, y, z)$$

#### The Exterior Derivative

Now, we will look at the relationship between  $d$  and  $d_s$ . Let  $\omega = \sum_I \omega_I dx^I$  be a spatial differential form (i.e. no component of  $\omega$  involves a  $dx^0$ ). We can compute  $d\omega$  as follows

$$d\omega = \sum_I d\omega_I \wedge dx^I \tag{2.55}$$

$$= \sum_I \left[ \left( \sum_{i=0}^3 \partial_i \omega_I dx^i \right) \wedge dx^I \right] \tag{2.56}$$

$$= \sum_I \left( \partial_0 \omega_I dx^0 \wedge dx^I + \sum_{i=1}^3 \partial_i \omega_I dx^i \wedge dx^I \right) \tag{2.57}$$

$$= dx^0 \wedge \partial_0 \omega + d_s \omega \tag{2.58}$$

$$= (dx^0 \wedge \partial_0 + d_s) \omega \tag{2.59}$$

So we see that  $d = dx^0 \wedge \partial_0 + d_s$ . The spacetime exterior derivative is just the spatial exterior derivative with an extra term related to the time derivative.

#### The Hodge Star of Spatial Forms

Now, we will relate  $\star$  and  $\star_s$ . Let  $\omega$  be a spatial  $k$ -form. The spacetime Hodge star is defined by the property that

$$\omega \wedge \star \omega = \langle \omega, \omega \rangle \mu \tag{2.60}$$

Here  $\mu = dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3$  is the spacetime volume form. Let  $\mu_s = dx^1 \wedge dx^2 \wedge dx^3$  be the spatial volume form. Clearly  $\mu = dx^0 \wedge \mu_s$ . Furthermore, we know that  $\omega \wedge \star_s \omega = \langle \omega, \omega \rangle \mu_s$ .

---

<sup>6</sup>We won't need to worry about Maxwell's equations in spacetime for our later discussion of MHD. But it provides an interesting and natural example of how differential forms can simplify our representation of physical systems.

Therefore,

$$\omega \wedge \star \omega = \langle \omega, \omega \rangle \mu \quad (2.61)$$

$$= dx^0 \wedge \langle \omega, \omega \rangle \mu_s \quad (2.62)$$

$$= dx^0 \wedge \omega \wedge \star_s \omega \quad (2.63)$$

$$= \omega \wedge (-1)^k (dx^0 \wedge \star_s \omega) \quad (2.64)$$

So  $\star \omega = (-1)^k dx^0 \wedge \star_s \omega$  when  $\omega$  is purely a spatial  $k$ -form.

### The Hodge Star of Forms with a Time Component

Now, suppose that  $\omega = dx^0 \wedge \omega_s$  where  $\omega_s$  is a spatial form. Then  $\langle \omega, \omega \rangle = -\langle \omega_s, \omega_s \rangle$ , so we need  $\omega \wedge \star \omega = -\langle \omega_s, \omega_s \rangle \mu$ . We know that  $\omega_s \wedge \star_s \omega_s = \langle \omega_s, \omega_s \rangle \mu_s$ . Thus,

$$\omega \wedge \star_s \omega_s = dx^0 \wedge \omega_s \wedge \star_s \omega_s = dx^0 \wedge \mu_s = \mu \quad (2.65)$$

So,  $\star \omega = -\star_s \omega$  when  $\omega$  is the wedge product of  $dx^0$  and a spatial form.

Finally, we note for completeness that  $\star dx^0 = -\mu_s$ .

### 2.3.3 Covariant formulation of Maxwell's equations

Finally, we've developed all of the tools we need to write Maxwell's equations in spacetime. We will begin with the homogeneous equations (equations two and three). Maxwell's second equation tells us that  $d\beta = 0$  and Maxwell's third equation tells us that  $d_s \eta + \frac{1}{c} \frac{\partial \beta}{\partial t} = 0$ . Because, we write coordinates in spacetime as

$$(x^0, x^1, x^2, x^3) = (ct, x, y, z) \quad (2.66)$$

it turns out that  $\frac{1}{c} \frac{\partial \beta}{\partial t} = \frac{\partial \beta}{\partial x^0} =: \partial_0 \beta$ . So we can write Maxwell's third equation as  $d_s \eta + \partial_0 \beta = 0$

The second equation is an equation of 3-forms, and the third equation is an equation of 2-forms. We can make them both equations of 3-forms by wedging the third equation with  $dx^0$ . Then we have  $d_s \beta = 0$  and  $dx^0 \wedge (d_s \eta + \partial_0 \beta) = 0$ . We can add these together to get  $d_s \beta + dx^0 \wedge \partial_0 \beta + dx^0 \wedge d_s \eta = 0$ . We note that because we are adding together forms of different types, their sum is 0 if and only if the individual terms in the sum are 0. So this equation expresses both Maxwell's second law and his third law. Inspecting the sum, we see that the first two terms are our expression for  $d\beta$ ! Furthermore, the last term is  $d(\eta \wedge dx^0)$ .

$$d(\eta \wedge dx^0) = d\eta \wedge dx^0 \quad (2.67)$$

$$= (d_s \eta + dx^0 \wedge \partial_0 c \eta) \wedge dx^0 \quad (2.68)$$

$$= d_s \eta \wedge dx^0 \quad (2.69)$$

$$= dx^0 \wedge d_s \eta \quad (2.70)$$

Therefore, we can write Maxwell's second and third equations as  $d\beta + d(\eta \wedge dx^0) = 0$ , or  $d(\beta + \eta \wedge dx^0) = 0$ . To simplify even more, we call  $F := \beta + \eta \wedge dx^0$  the *Faraday tensor*, and simply write  $dF = 0$ .

Now, we consider  $d \star F$ . We'll start by computing  $\star F$ .

$$\star F = \star(\beta + \eta \wedge dx^0) = \star\beta + \star(\eta \wedge dx^0) \quad (2.71)$$

Since  $\beta$  is a spatial 2-form,  $\star\beta = dx^0 \wedge \star_s \beta$ . Since  $\eta$  is a spatial 1-form,

$$\star(\eta \wedge dx^0) = -\star(dx^0 \wedge \eta) = -(-\star_s \eta) = \star_s \eta \quad (2.72)$$

Putting this together shows us that  $\star F = dx^0 \wedge \star_s \beta + \star_s \eta$ . Now, we can take the exterior derivative.

$$d \star F = d(dx^0 \wedge \star_s \beta + \star_s \eta) \quad (2.73)$$

$$= (d_s + dx^0 \wedge \partial_0)(dx^0 \wedge \star_s \beta + \star_s \eta) \quad (2.74)$$

$$= -dx^0 \wedge d_s \star_s \beta + d_s \star_s \eta + dx^0 \wedge \partial_0 \star_s \eta \quad (2.75)$$

$$= d_s \star_s \eta + dx^0 \wedge (\partial_0 \star_s \eta - d_s \star_s d_s \beta) \quad (2.76)$$

Maxwell's first equation tells us that  $d_s \star_s \eta = 4\pi\rho$ . Maxwell's fourth equation tells us that  $\partial_0 \eta - \star_s d_s \star_s \beta = -\frac{4\pi}{c} \star_s \mathcal{J}$ . Therefore,

$$d \star F = 4\pi \star \rho dx^0 - \frac{4\pi}{c} dx^0 \wedge \mathcal{J} \quad (2.77)$$

$$(2.78)$$

We define  $c\rho - dx^0 \wedge \mathcal{J} = \mathfrak{J}$  to be the four-current. Now, our equation reads  $d \star F = \frac{4\pi}{c} \mathfrak{J}$ . This lets us finally express Maxwell's equations (in cgs units) as

$$dF = 0 \quad \text{and} \quad d \star F = \frac{4\pi}{c} \mathfrak{J} \quad (2.79)$$

### 2.3.4 Ideal Fluids

A simple approximation of a fluid is to imagine many particles traveling around, each with their own velocity. If we assume smoothness, we could represent the state of a fluid in a container  $M$  as a vector field  $v \in \Gamma(TM)$  representing the fluid's velocity at each point. However, it turns out to be quite helpful to work with the associated 1-form  $\eta := v^\flat$  instead, and we will primarily work with  $\eta$ . Since the particles move along the velocity field, and the particle's velocities add together to form the velocity field, these velocity vectors advect themselves along their own velocity field. Mathematically, we can write this claim as

$$\dot{\eta} = -\mathfrak{L}_{\eta^\sharp} \eta \quad (2.80)$$

But this is not quite the equations of motion for an ideal fluid. This allows the fluid to bunch up or spread out however it would like. In reality, fluids cannot compress infinitely or spread infinitely thin. They are, in fact, quite difficult to compress. A common assumption when modeling fluids is to assume that the fluids are incompressible. This means that the divergence of our velocity field must be 0. So we require that  $\delta\eta = 0$ . Of course, the equation of motion we proposed earlier might not preserve the property that  $\delta\eta = 0$ . To fix this, we note that if  $\delta\eta$ , initially, and if  $\delta\dot{\eta} = 0$  for all time, then  $\delta\eta$  will be zero for all time. So it is sufficient to ensure that  $\dot{\eta} = 0$ . Furthermore, given any 1-form, we can add an exact 1-form to shift its codifferential by any constant. So we can take the following pair of equations as our equations of motion

$$\dot{\eta} + \mathcal{L}_{\eta^\sharp}\eta + dp = 0 \quad (2.81)$$

$$\delta\eta = 0 \quad (2.82)$$

$p$  is called the *pressure*, and allows us to enforce the incompressibility constraint on  $\eta$ . These equations are called the *Euler equations* for an ideal fluid.

This model of fluids is very idealized. In particular, we assume that the fluid is incompressible, and also that it has no viscosity.

## 2.4 Magnetohydrodynamics

Magnetohydrodynamics(MHD) is the physics of current-carrying fluids. To study such fluids, we combine Maxwell's equations with the Euler equations for an ideal fluid, and make some additional simplifying assumptions. The material we present in this section can be found in [10], [9], or [19].

### 2.4.1 Equations of Motion

We will consider the case of a fluid which is not in the presence of external electric fields.

We assume that our fluid behaves like a current-carrying wire in a magnetic field  $\beta \in \Omega^2(M)$ . The force felt by a wire with current  $J$  in a magnetic field  $\beta$  is given by<sup>7</sup>

$$J \times (\star\beta)^\sharp = (-\iota_J\beta)^\sharp \quad (2.83)$$

If we assume that the electric field is roughly constant, then Maxwell's fourth equation (equation 2.54<sup>8</sup>) tells us that our current 1-form is given by  $\delta\beta$  (up to some scaling factors). So the electromagnetic force felt by the fluid is  $\iota_{(\delta\beta)^\sharp}\beta$ . This gives us the velocity equation

$$\dot{\eta} + \mathcal{L}_{\eta^\sharp}\eta + \iota_{(\delta\beta)^\sharp}\beta + dp = 0 \quad (2.84)$$

---

<sup>7</sup>See proposition 1 in Appendix B

<sup>8</sup>Beware: in the section on Maxwell's equations, we used  $\eta$  for the electric field, while now it represents the velocity 1-form.

There is a convenient trick we can apply to this equation. Using Cartan's magic formula, we see that

$$\dot{\eta} + d\iota_{\eta^\sharp}\eta + \iota_{\eta^\sharp}d\eta + \iota_{(\delta\beta)^\sharp}\beta + dp = 0 \quad (2.85)$$

Now,

$$d\iota_{\eta^\sharp}\eta = d|\eta|^2 \quad (2.86)$$

Since this is the differential of a scalar function, we can just redefine our pressure to be the old pressure plus  $|\eta|^2$  to obtain a simpler equation

$$\dot{\eta} + \iota_{\eta^\sharp}d\eta + \iota_{(\delta\beta)^\sharp}\beta + dp = 0 \quad (2.87)$$

The magnetic field is generated by the moving particles, so it is simply advected by the fluid. In terms of the Lie derivative, we write

$$\dot{\beta} + \mathfrak{L}_{\eta^\sharp}\beta = 0 \quad (2.88)$$

As one final simplification, we note that the magnetic field is divergence-free. Thus,

$$0 = \delta(\star\beta) = \star^{-1}d\beta \quad (2.89)$$

We conclude that  $d\beta = 0$ . This fact, along with Cartan's magic formula, shows us that

$$\mathfrak{L}_{\eta^\sharp}\beta = \iota_{\eta^\sharp}d\beta + d\iota_{\eta^\sharp}\beta = d\iota_{\eta^\sharp}\beta \quad (2.90)$$

This gives us the ideal MHD equations

$$\dot{\eta} + \iota_{\eta^\sharp}d\eta + \iota_{(\delta\beta)^\sharp}\beta + dp = 0 \quad (2.91)$$

$$\dot{\beta} + d\iota_{\eta^\sharp}\beta = 0 \quad (2.92)$$

$$\delta\eta = 0 \quad (2.93)$$

## 2.4.2 Conservation Laws

For more about conservation laws in fluids, see [19] or [20]

### Energy

The energy of a magnetic fluid in a domain  $M$  has two components: there is kinetic energy, and there is potential energy stored in the magnetic field. The kinetic energy has the form

$$K(\eta, \beta) = \frac{1}{2} \|\eta\|^2 = \frac{1}{2} \int_M \eta \wedge \star\eta \quad (2.94)$$

The potential energy has the form

$$U(\eta, \beta) = \frac{1}{2} \|\beta\|^2 = \frac{1}{2} \int_M \beta \wedge \star\beta \quad (2.95)$$

Thus, our energy is given by

$$E(\eta, \beta) := \frac{1}{2} \|\eta\|^2 + \frac{1}{2} \|\beta\|^2 = \frac{1}{2} \int_M \eta \wedge \star \eta + \beta \wedge \star \beta \quad (2.96)$$

We can prove conservation of energy with a simple computation.

$$\frac{d}{dt} E = \frac{1}{2} \frac{d}{dt} \|\eta\|^2 + \frac{1}{2} \frac{d}{dt} \|\beta\|^2 \quad (2.97)$$

$$= \langle\langle \eta, \dot{\eta} \rangle\rangle + \langle\langle \beta, \dot{\beta} \rangle\rangle \quad (2.98)$$

$$= \langle\langle \eta, -\iota_{\eta^\sharp} d\eta - \iota_{(\delta\beta)^\sharp} \beta - dp \rangle\rangle + \langle\langle \beta, -d\iota_{\eta^\sharp} \beta \rangle\rangle \quad (2.99)$$

$$= \langle\langle \eta \wedge \eta, d\eta \rangle\rangle - \langle\langle \delta\beta \wedge \eta, \beta \rangle\rangle - \langle\langle \delta\eta, p \rangle\rangle \quad (2.100)$$

$$- \langle\langle \eta \wedge \delta\beta, \beta \rangle\rangle \quad (2.101)$$

$$= \langle\langle \delta\beta \wedge \eta + \eta \wedge \delta\beta, \beta \rangle\rangle \quad (2.102)$$

$$= 0 \quad (2.103)$$

## Cross Helicity

The cross helicity is defined to be

$$H_\chi := \int_M \eta \wedge \beta \quad (2.104)$$

Again, we can compute its time derivative to show that it is conserved. Note that since  $\beta$  is a 2-form, we have  $\eta \wedge \beta = \beta \wedge \eta$ .

$$\frac{d}{dt} H_\chi = \frac{d}{dt} \int_M \eta \wedge \beta \quad (2.105)$$

$$= \langle\langle \dot{\eta}, \star \beta \rangle\rangle + \langle\langle \dot{\beta}, \star \eta \rangle\rangle \quad (2.106)$$

$$= \langle\langle -\iota_{\eta^\sharp} d\eta - \iota_{(\delta\beta)^\sharp} \beta - dp, \star \beta \rangle\rangle + \langle\langle -d\iota_{\eta^\sharp} \beta, \star \eta \rangle\rangle \quad (2.107)$$

$$= -\langle\langle d\eta, \eta \wedge \star \beta \rangle\rangle - \langle\langle \beta, \delta\beta \wedge \star \beta \rangle\rangle - \langle\langle p, \star d\beta \rangle\rangle - \langle\langle \beta, \eta \wedge \star d\eta \rangle\rangle \quad (2.108)$$

$$= - \int_M \eta \wedge \star \beta \wedge \star d\eta - \int_M \delta\beta \wedge \star \beta \wedge \star \beta - 0 - \int_M \eta \wedge \star d\eta \wedge \star \beta \quad (2.109)$$

$$= - \int_M \eta \wedge \star \beta \wedge \star d\eta + \int_M \eta \wedge \star \beta \wedge \star d\eta \quad (2.110)$$

$$= 0 \quad (2.111)$$

### 2.4.3 Magnetic Helicity

Suppose that our domain  $M$  satisfies  $H^1(M) = 0$ . Then, we get another conserved quantity called the *magnetic helicity*.

Since  $H^1(M) = 0$ , every closed form is exact. So we can write the magnetic field as the exterior derivative of a vector potential  $\beta = d\alpha$ . Now, we define the magnetic helicity to be

$$H_M := \int_M \alpha \wedge \beta \quad (2.112)$$

Like with cross helicity and energy, we can show that this is conserved via direct computation. This takes a little more work, as we first have to write the evolution equation for  $\alpha$ . Using the evolution equation for  $\beta$ , we see that

$$d\dot{\alpha} = -d\iota_{\eta^\#} d\alpha \quad (2.113)$$

Therefore, the form  $\dot{\alpha} + \iota_{\eta^\#} d\alpha$  is closed. Since  $H^1(M) = 0$ , it must also be exact. So we have

$$\dot{\alpha} = -\iota_{\eta^\#} d\alpha + dq \quad (2.114)$$

for some 0-form  $q$ .

Now, we can show that magnetic helicity is conserved.

$$\frac{d}{dt} H_M = \frac{d}{dt} \int_M \alpha \wedge \beta \quad (2.115)$$

$$= \langle\langle \dot{\alpha}, \star \beta \rangle\rangle + \langle\langle \dot{\beta}, \star \alpha \rangle\rangle \quad (2.116)$$

$$= \langle\langle -\iota_{\eta^\#} d\alpha + dq, \star \beta \rangle\rangle + \langle\langle -d\iota_{\eta^\#} \beta, \star \alpha \rangle\rangle \quad (2.117)$$

$$= -\langle\langle d\alpha, \eta \wedge \star \beta \rangle\rangle + \langle\langle q, \star d\beta \rangle\rangle - \langle\langle \beta, \eta \wedge \star d\alpha \rangle\rangle \quad (2.118)$$

$$= -\langle\langle d\alpha, \eta \wedge \star \beta \rangle\rangle - \langle\langle \beta, \eta \wedge \star d\alpha \rangle\rangle \quad (2.119)$$

$$= - \int_M \eta \wedge \star \beta \wedge \star d\alpha - \int_M \eta \wedge \star d\alpha \wedge \star \beta \quad (2.120)$$

$$= 0 \quad (2.121)$$

#### 2.4.4 2D MHD

We consider the case of a two-dimensional conducting fluid in the  $xy$ -plane. The equations we derive here are also given in [10]. Since the component of the magnetic 2-form which lives in the plane cannot affect the fluid's velocity, we assume that the magnetic 2-form does not lie in the plane.

We will denote the Hodge star in 3D space by  $\star_{3D}$ , and we will denote the Hodge star in the plane by  $\star_{2D}$ . Note if  $\eta$  is a 1-form in the  $xy$ -plane, then<sup>9</sup>

$$\star_{3D} (\star_{2D} \eta \wedge dz) = \eta \quad (2.122)$$

Let  $b = \star_{3D} \beta$ . Note that  $\beta = \star_{2D} b \wedge dz$ .

To find the equation of motion for  $\eta$ , we need to understand the term  $\iota_{(\delta\beta)^\#} \beta$ . We start with  $\delta\beta$ .

$$\delta\beta = \star_{3D} d_{3D} \star_{3D} \beta = \star_{3D} d_{3D} b \quad (2.123)$$

Note that because  $b$  has no  $z$ -dependence,  $d_{3D} b = d_{2D} b$ . So we have

$$\delta\beta = \star_{3D} d_{3D} b = \star_{3D} d_{2D} b = (\star_{2D} d_{2D} b) \wedge dz = (\star_{2D} d_{2D} b) dz \quad (2.124)$$

---

<sup>9</sup>See proposition 2 in Appendix B for proof

where the second-to-last equality used proposition 3 from Appendix B.

Therefore,

$$\iota_{(\delta_{3D}\beta)^{\sharp}}\beta = \iota_{(\star_{2D}d_{2D}b)dz}(\star_{2D}b \wedge dz) \quad (2.125)$$

$$= (\star_{2D}d_{2D}b)\iota_{dz}(\star_{2D}b \wedge dz) \quad (2.126)$$

$$= -(\star_{2D}d_{2D}b)(\star_{2D}b) \quad (2.127)$$

It is convenient to write this expression in a different form. It turns out that we have<sup>10</sup>

$$-(\star_{2D}d_{2D}b)(\star_{2D}b) = -\iota_{b^{\sharp}}d_{2D}b \quad (2.128)$$

This gives us our equation of motion for  $\eta$ :

$$\dot{\eta} + \iota_{\eta^{\sharp}}d\eta - \iota_{b^{\sharp}}db + dp = 0 \quad (2.129)$$

(where the exterior derivatives are taken to be  $d_{2D}$ )

We can follow the same procedure to find the equation of motion for  $b$ . We know that  $\beta = \star_{2D}b \wedge dz$ . Substituting this into the equation of motion for  $\beta$ , we find that

$$\star_{2D}\dot{b} \wedge dz + d_{3D}\iota_{\eta^{\sharp}}(\star_{2D}b \wedge dz) = 0 \quad (2.130)$$

Note that since  $\eta$  lies in the plane,  $dz(\eta^{\sharp}) = 0$ . Thus,

$$d_{3D}\iota_{\eta^{\sharp}}(\star_{2D}b \wedge dz) = d_{3D}[(\iota_{\eta^{\sharp}}\star_{2D}b) \wedge dz] \quad (2.131)$$

Next, we note that

$$\iota_{\eta^{\sharp}}\star_{2D}b = \langle \eta, \star_{2D}b \rangle \quad (2.132)$$

$$= \star_{2D}(\eta \wedge \star_{2D}^2 b) \quad (2.133)$$

$$= -\star_{2D}(\eta \wedge b) \quad (2.134)$$

$$= \star_{2D}(b \wedge \eta) \quad (2.135)$$

This clearly has no  $z$  dependence. So

$$d_{3D}[\star_{2D}(b \wedge \eta) \wedge dz] = [d_{2D}\star_{2D}(b \wedge \eta)] \wedge dz \quad (2.136)$$

Thus, our equation of motion has the form

$$[\star_{2D}\dot{b} + d_{2D}\star_{2D}(b \wedge \eta)] \wedge dz = 0 \quad (2.137)$$

Therefore, we see that

$$\star_{2D}\dot{b} + d_{2D}\star_{2D}(b \wedge \eta) = 0 \quad (2.138)$$

Now, we can multiply through by  $\star_{2D}^{-1}$  and use the fact that  $\delta_{2D} := \star_{2D}^{-1}d_{2D}\star_{2D}$  to obtain the equation of motion

$$\dot{b} + \delta_{2D}(b \wedge \eta) = 0 \quad (2.139)$$

---

<sup>10</sup>See Appendix B proposition 4

Finally, we will determine the initial condition on  $b$ . Recall that we require  $d_{3D}\beta = 0$  at  $t = 0$ . This implies that

$$0 = d_{3D}(\star_{2D} b \wedge dz) \quad (2.140)$$

$$= (d_{2D} \star_{2D} b) \wedge dz \quad (2.141)$$

So our initial condition on  $b$  is that  $\delta_{2D} b = 0$ .

Thus, we can express the 2D MHD equations entirely in terms of the exterior calculus operators in the plane, and have the form

$$\dot{\eta} + \iota_{\eta^*} d\eta - \iota_{b^*} db + dp = 0 \quad (2.142)$$

$$\dot{b} + \delta(b \wedge \eta) = 0 \quad (2.143)$$

$$\delta\eta = 0 \quad (2.144)$$

with initial condition  $\delta b = 0$ .

## 2.5 Formal Lagrangians

In this section, we will mostly follow the treatment in [21]. For more information about formal Lagrangians, see [21], [22], or [23] (the first has a fairly geometric viewpoint while the latter two approach the subject from a differential equations perspective).

Lagrangian mechanics gives us some powerful tools for understanding physical systems, but there is a significant drawback. In order to apply this framework, we need to have a Lagrangian which yields our system's equations of motion. Although many simple systems have Lagrangians which are easy to guess, this is not always the case. In fact, not all systems of differential equations admit a variational formulation [21].

However, there is a way of working around this limitation. Earlier<sup>11</sup>, we saw that one can add Lagrange multipliers to a Lagrangian to enforce constraints. So we can just enforce arbitrary equations of motion using Lagrange multipliers! Suppose we have a system of first-order differential equations

$$\mathcal{F}(u, \dot{u}) = 0 \quad (2.145)$$

We define a *formal Lagrangian*

$$\mathcal{L}(u, \dot{u}, \lambda, \dot{\lambda}) := \lambda \cdot \mathcal{F}(u, \dot{u}) \quad (2.146)$$

The Euler-Lagrange equations for the  $\lambda$  variables tell us that

$$\mathcal{F}(u, \dot{u}) = \frac{\partial \mathcal{L}}{\partial \lambda} = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\lambda}} = 0 \quad (2.147)$$

which means that our system follows the desired equations of motion!

---

<sup>11</sup>as an aside in section 2.1.2

Of course, this formal Lagrangian does not only solve our differential equations. We have doubled the number of variables, and the  $\lambda$ 's also evolve over time. The Euler-Lagrange equations for the  $u$  variables give us the evolution equations for  $\lambda$ .

$$\frac{\partial \mathcal{L}}{\partial u} = \lambda \frac{\partial \mathcal{F}}{\partial u} \quad (2.148)$$

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{u}} = \frac{d}{dt} \left( \lambda \frac{\partial \mathcal{F}}{\partial \dot{u}} \right) \quad (2.149)$$

$$= \dot{\lambda} \frac{\partial \mathcal{F}}{\partial \dot{u}} + \lambda \frac{\partial^2 \mathcal{F}}{\partial u \partial \dot{u}} \quad (2.150)$$

So the Euler-Lagrange equations tell us that

$$\dot{\lambda} \frac{\partial \mathcal{F}}{\partial \dot{u}} + \lambda \left( \frac{\partial \mathcal{F}}{\partial \dot{u}} - \frac{\partial^2 \mathcal{F}}{\partial u \partial \dot{u}} \right) = 0 \quad (2.151)$$

Now, suppose that given a solution  $u$  of  $\mathcal{F}$ , we can apply some transformation  $\Phi : u \mapsto (u, \lambda)$  to obtain a solution of our extended equations of motion.<sup>12</sup> In this case, solutions of our original system give us solutions of the larger system. So if we have a conserved quantity in the larger system, it must also be conserved by the original, smaller system. Furthermore, it turns out that symmetries of our original system yield symmetries of the larger system. Thus, we can use formal Lagrangians to find conserved quantities of systems of differential equations which do not naturally have a Lagrangian formulation.

---

<sup>12</sup>This seems like a strong assumption, but it is often the case. For an example, look at section C.1 in Appendix C

# Chapter 3

## Computational Preliminaries

### 3.1 Introduction

So far, we have built up a substantial amount of mathematical machinery and physical theories, but all of it lives in the setting of continuous mathematics. In order to apply these ideas to computational problems, we need to discretize everything. Discretization is a subtle procedure - frequently, equivalent concepts in the smooth setting lead naturally to different discretizations which each have their own positives and negatives.

### 3.2 Discrete Exterior Calculus

In this section, we will describe the theory of discrete exterior calculus (DEC), as presented in [24]. DEC provides an elegant and efficient way of applying the theory of manifolds and differential forms to computational problems. In the end, we will be able to write down the exterior calculus operators which we defined earlier as sparse matrices, which we can efficiently compute with. The fundamental idea of discrete exterior calculus is that we can approximate a smooth manifold by a simplicial complex (or cell complex). In the case of 2-dimensional manifolds, this means using a triangle (or polygonal) mesh.<sup>1</sup>

#### 3.2.1 Discrete Differential Forms

There is a very elegant discretization of differential forms. We cannot transfer the technical construction of the exterior algebra of the tangent bundle into the discrete setting, but we can bring the general intuition that a  $k$ -form is an object which is integrated over  $k$ -dimensional submanifolds. The natural notion of “discrete  $k$ -dimensional submanifold” is just a collection of  $k$ -simplices. So we could define a discrete differential  $k$ -form to be a real-valued function

---

<sup>1</sup>In the following sections, we restrict our attention to triangle meshes and simplicial complexes, but all of the ideas extend readily to meshes whose faces are polygons other than triangles. Indeed, in our MHD simulation, we will use a regular grid instead of a triangle mesh.

on  $k$ -simplices. With this definition, our discrete  $k$ -forms can be “integrated” over discrete submanifolds, and this integration satisfies the desirable property that

$$\int_{A \cup B} \omega = \int_A \omega + \int_B \omega \quad (3.1)$$

for disjoint discrete submanifolds  $A$  and  $B$ .

In the language of algebraic topology, a real-valued function on  $k$ -simplices is called a  $k$ -cochain. So we define

**Definition 3.2.1.** A discrete differential  $k$  form on a simplicial complex  $X$  is a  $k$ -cochain on  $X$ .

When thinking about our discrete differential forms as differential forms, we will often write the evaluation of a form  $\omega$  on a submanifold  $S$  as  $\int_S \omega$ . When thinking of discrete differential  $k$ -forms as real-valued functions on  $k$ -cells, we will sometimes write the evaluation of a form  $\omega$  on a submanifold  $S$  as  $\omega(S)$ .

Practically speaking, if our discrete mesh has  $n$  cells of dimension  $k$ , a discrete  $k$ -form is just a vector in  $\mathbb{R}^n$ , since we have to assign a real number to each  $k$ -cell.

### 3.2.2 Discrete Exterior Derivative

We can define a discrete version of the exterior derivative using Stokes’ Theorem. Recall that in the continuous setting, Stokes’ theorem tells us that

$$\int_{\partial S} \omega = \int_S d\omega \quad (3.2)$$

We already understand how to interpret the differential form  $\omega$  and the submanifold  $S$  in the discrete world. And taking the boundary of  $S$  is straightforward. So we can use Stokes’ theorem to *define*  $d\omega$  as “the differential form whose integral on  $S$  is  $\int_{\partial S} \omega$ ”.

**Definition 3.2.2.** Given a discrete  $k$ -form  $\omega$ , the discrete exterior derivative of  $\omega$  is the discrete  $(k+1)$ -form  $d\omega$  whose value on a  $(k+1)$ -simplex  $S$  is given by  $\int_{\partial S} \omega$ .

When we define the exterior derivative in this way, we ensure that Stokes’ theorem is exactly true even in our discrete equations. This is useful for reasoning about discrete systems.

Observe that linearity of our discrete exterior derivative follows from linearity of integration. For any  $(k+1)$ -cell  $S$ , we have

$$d(\omega + \eta)(S) = \int_{\partial S} (\omega + \eta) = \int_{\partial S} \omega + \int_{\partial S} \eta = d\omega(S) + d\eta(S) \quad (3.3)$$

Since this is true for every  $S$ , we conclude that  $d(\omega + \eta) = d\omega + d\eta$ . A similar argument shows that for any scalar  $r \in \mathbb{R}$ , we have  $d(r\omega) = r d\omega$ .

Since the discrete exterior derivative is a linear operator, we can write it down as a matrix. Let  $\omega$  be a discrete  $k$  form. Suppose our discrete manifold has  $n$   $k$ -cells and  $m$   $(k+1)$ -cells. Let

$$\omega = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_n \end{pmatrix} \quad (3.4)$$

where  $\omega_i$  is the value of  $\omega$  on the  $i$ th  $k$ -cell. Since the discrete exterior derivative takes  $k$ -forms to  $(k+1)$ -forms, it will be an  $m \times n$  matrix.

The  $i$ th entry in  $d\omega$ , which we will denote  $(d\omega)_i$ , is the value of  $d\omega$  on the  $i$ th  $(k+1)$ -simplex. Let us denote this cell by  $S_i$ . By definition,

$$\int_{S_i} d\omega = \int_{\partial S_i} \omega = \sum_{T_j \in \partial S_i} \int_{T_j} \omega = \sum_{T_j \in \partial S_i} \pm \omega_j \quad (3.5)$$

where the sign of  $\omega_j$  in the final sum depends on the orientation of  $T_j$  as an element of  $S_i$ 's boundary. So the  $i$ th row of  $d$  is mostly zeros, with  $\pm 1$  in the entries corresponding to elements of  $\partial S_i$ . This matrix has a nice interpretation. It is just the transpose of the matrix that takes a cell  $S$  to its boundary  $\partial S$ . If we write this map as  $\partial$ , then we obtain the simple identity

$$d = \partial^T \quad (3.6)$$

For this reason,  $d$  is sometimes called the *codifferential*.

Note that this matrix is very sparse. In general, the boundary of a  $(k+1)$ -simplex will contain only  $k+1$  boundary cells, so each column of  $\partial_k$  contains only  $k+1$  entries. Thus,  $d_k$  contains only  $k+1$  nonzero entries in each row.

### 3.2.3 Discrete Hodge Star

The Hodge star is a trickier operation to discretize than the exterior derivative. The exterior derivative (in the continuous setting) is a purely topological operation, whereas the Hodge star depends on the metric with which we endow our manifold. The topological-metric distinction is less clear in the continuous setting, but we can understand the difference as follows: the discrete exterior derivative depends only on the combinatorial structure of our discrete mesh (i.e. which simplices form the boundary of which )

Initially, it is not obvious that it even makes sense to talk about a discrete Hodge star. The continuous Hodge star on an  $n$ -dimensional smooth manifold  $M$  defines an isomorphism  $\Omega^k(M) \xrightarrow{\sim} \Omega^{n-k}(M)$ . But the space of discrete  $k$ -forms does not have to be isomorphic to the space of discrete  $(n-k)$ -forms! It's easy to come up with simplicial complexes which violate this condition. For example, consider a triangle. It has 3 vertices, so a discrete 0-form is a vector in  $\mathbb{R}^3$ . But it only has one face, so a discrete 2-form is a scalar in  $\mathbb{R}$ . Clearly  $\mathbb{R}$  and  $\mathbb{R}^3$  are not linearly isomorphic!

So we can't define a discrete Hodge star that takes in discrete  $k$ -forms and returns discrete  $(n - k)$ -forms. The basic problem is that we don't have a natural correspondence between  $k$ -cells and  $(n - k)$ -cells in a general simplicial complex.

Luckily for us, this problem has a neat solution. Although there is no natural correspondence between  $k$ -cells and  $(n - k)$ -cells, there *is* a natural correspondence between  $k$ -cells in our mesh and  $(n - k)$ -cells in the *dual* mesh. So we have a natural map between *primal*  $k$ -forms and *dual*  $(n - k)$ -forms. All that we have to do now is fix a scaling factor. Let  $S_i^k$  denote a primal  $k$ -cell, and  $\hat{S}_i^{n-k}$  denote its corresponding dual  $(n - k)$ -cell. Let  $\omega$  be a discrete  $k$ -form with

$$\int_{S_i^k} \omega = \omega_i \quad (3.7)$$

We need to determine

$$\int_{\hat{S}_i^{n-k}} \star \omega \quad (3.8)$$

We want  $\star \omega$  and  $\omega$  to have the same density, so we set

$$\int_{\hat{S}_i^{n-k}} \star \omega = \pm \frac{|\hat{S}_i^{n-k}|}{|S_i^k|} \int_{S_i^k} \omega \quad (3.9)$$

where  $|S_i^k|$  is the volume of simplex  $|S_i^k|$ . The sign depends on the relative orientations of  $S_i^k$  and its dual cell.

If we represent  $\omega$  as a vector, then  $\star \omega$  has the same number of components as  $\omega$  (since the primal and dual meshes have the same number of  $k$ -cells and  $(n - k)$ -cells respectively). So we can represent  $\star$  by a diagonal matrix.

### 3.2.4 Discrete Inner Product

Given two discrete  $k$ -forms  $\alpha$  and  $\beta$  (thought of as vectors), we define the discrete (Hodge) inner product to be

$$\langle\langle \alpha, \beta \rangle\rangle := \alpha^T \star_k \beta \quad (3.10)$$

Note that this is symmetric because our discrete Hodge star is diagonal (and thus  $\star_k^T = \star_k$ ).

This definition is analogous to the continuous identity

$$\langle\langle \alpha, \beta \rangle\rangle = \int_M \alpha \wedge \star \beta \quad (3.11)$$

One convenient property of this inner product is that the discrete exterior derivative and the discrete codifferential are adjoint

$$\langle\langle \alpha, d\beta \rangle\rangle = \alpha^T \star_k d\beta \quad (3.12)$$

$$= \alpha^T \star_k d \star_k^{-1} \star_k \beta \quad (3.13)$$

$$= (\star_k^{-1} d^T \star_k \alpha)^T \star_k \beta \quad (3.14)$$

$$= \langle\langle \delta \alpha, \beta \rangle\rangle \quad (3.15)$$

Where we have used the fact that  $\star_k^T = \star_k$ .

### 3.2.5 Discrete Codifferential

In the continuous setting, we defined the codifferential to be  $\delta := \star^{-1}d\star$ . We need to modify this, slightly, because our discrete Hodge star does not simply map discrete differential forms to discrete differential forms. Instead, it maps discrete primal forms to discrete dual forms (and vice versa). So when we define  $\delta$  on our primal mesh, we actually need to write it in terms of  $d$  on the dual mesh.

Luckily for us, the expression for  $d$  on the dual mesh is simple - it is simply the transpose of our primal  $d$ . So we define  $\delta := \star^{-1}d^T\star$ .

### 3.2.6 Discrete Wedge Product

It turns out that it is impossible to define a discrete wedge product which is compatible with our discrete exterior derivative and respects the (anti)-symmetry of the continuous wedge product. This is known as the *commutative cochain problem* [25], [26]. However, we can define a discrete wedge product of 1-forms provided we work on a regular grid.

### 3.2.7 Explicit Expressions on a Regular Grid

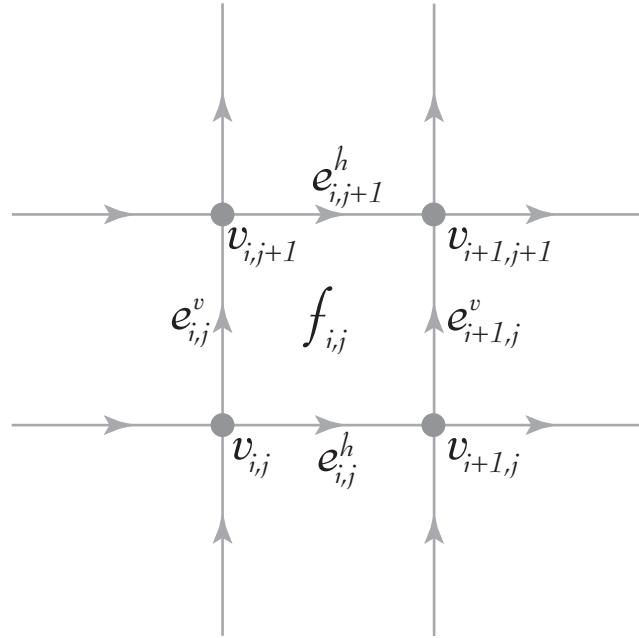


Figure 3.1: Reference for grid element naming convention

Now, we will give explicit expressions for these operators on a regular grid. Figure 3.1 shows our naming convention for the vertices, edges, and faces of the grid.

### Discrete Exterior Derivative

The discrete exterior derivative on 0-forms,  $d_0$ , takes a function  $f$  defined on vertices and returns a function  $df$  defined on edges. The value of  $df$  on an edge  $(v, w)$  from vertex  $v$  to vertex  $w$  is  $f(w) - f(v)$ . So we can express  $d_0$  as a matrix which takes

$$d_0 : v_{ij} \mapsto -e_{i,j}^h - e_{i,j}^v + e_{i-1,j}^h + e_{i,j-1}^v \quad (3.16)$$

Similarly, the discrete exterior derivative on 1-forms,  $d_1$ , takes a function  $f$  defined on edges and returns a function  $df$  defined on faces. The value of  $df$  on a face is the sum of the values of  $f$  around its boundary edges. So we can express  $d_1$  as a matrix which takes

$$d_1 : e_{i,j}^h \mapsto f_{i,j} - f_{i,j-1} \quad (3.17)$$

$$d_1 : e_{i,j}^v \mapsto -f_{i,j} + f_{i-1,j} \quad (3.18)$$

### Discrete Hodge Star

The discrete Hodge star is a diagonal matrix whose entries represent the ratio of the areas between primal and dual cells. Since the dual of a regular grid is just a translated copy of the grid, our Hodge star is simply the identity matrix.

### Discrete Hodge Inner Product

Since the discrete Hodge star is just the identity matrix, the discrete Hodge inner product is simply the standard inner product.

### Discrete Wedge Product

Let  $\alpha, \beta$  be 1-forms whose values on edge  $e_{i,j}^h$  are denoted by  $\alpha_{i,j}^h$  and  $\beta_{i,j}^h$  respectively (with the same convention for vertical edges). Then, we define the value of the discrete wedge product  $\alpha \wedge \beta$  on face  $f_{ij}$  to be

$$(\alpha \wedge \beta)_{i,j} := \alpha_{i,j+1/2}^h \beta_{i+1/2,j}^v - \alpha_{i+1/2,j}^v \beta_{i,j+1/2}^h \quad (3.19)$$

where  $\alpha_{i,j+1/2}^h := \frac{1}{2}(\alpha_{i,j} + \alpha_{i,j+1})$ , etc are the averaged values of the 1-forms on opposite edges of face  $f_{ij}$ .

## 3.3 Simulation

Now, we will consider the problem of representing and simulating physical systems on a computer. The material from this section largely comes from [16].

### 3.3.1 Euler Integration

One of the oldest and simplest methods of numerically solving differential equations is Euler integration. The idea behind Euler integration is simple. If we have a first-order differential equation  $\mathcal{F}(u, \dot{u}) = 0$ , then we can solve for  $\dot{u}$  to find  $\dot{u}$  for any given  $u$ . Taylor's theorem tells us that

$$u(t+h) \approx u(t) + h\dot{u}(t) \quad (3.20)$$

This gives us a straightforward algorithm for numerically solving this differential equation. We can just pick a fixed step size  $\Delta t$ , and set

$$u_{k+1} \leftarrow u_k + \Delta t \dot{u}_k \quad (3.21)$$

This is a simple algorithm, and in the limit as  $\Delta t \rightarrow 0$ , it converges to a solution of the equations of motion. However, for reasonable step sizes the algorithm can behave quite poorly.

As an example, consider the case of a 2D pendulum. We describe the state of the pendulum by its angle  $\theta$ , and the equation of motion is

$$\ddot{\theta} + \sin \theta = 0 \quad (3.22)$$

Since we defined Euler's method for first-order differential equations, we must convert this to a first order differential equation. So we add a velocity variable, and consider the equations of motion

$$\dot{\theta} - \omega = 0 \quad (3.23)$$

$$\dot{\omega} + \sin \theta = 0 \quad (3.24)$$

These equations give us the update rule

$$\theta_{k+1} \leftarrow \theta_k + \Delta t \omega_k \quad (3.25)$$

$$\omega_{k+1} \leftarrow \omega_k - \Delta t \sin \theta_k \quad (3.26)$$

However, if we use this update rule to simulate a pendulum, we do not obtain satisfactory results

One problem with this simulation is that our system does not conserve energy. In fact, it steadily gains energy over time. This is due to a systematic error in our Taylor expansions which we used to derive the update rule. Even though each approximation has a fairly small error, the errors accumulate over time.

### 3.3.2 Discrete Lagrangian Mechanics

In the previous section, we saw a simple example where Euler integration leads to a simulations which has poor long-term behavior due to accumulation of error. One problem we

observed with the simulation was that it did not conserve energy. So one might hope to produce more accurate simulations by designing simulations which explicitly conserve energy. Of course, just because a simulation conserves energy does not guarantee that it is accurate - a simulation which never moves will conserve energy perfectly, but will often give terrible results. However, energy conservation does ensure that our simulation will not blow up, and is often evidence that a simulation is accurate.

Lagrangian mechanics gives us a useful framework for understanding conserved quantities in continuous physics. Because we want to create simulations which respect conserved quantities, it is natural to turn to Lagrangian mechanics for aid.

We can follow the same steps we took when developing Lagrangian mechanics in Section 2.1.3 to develop a discrete-time version of Lagrangian mechanics. Suppose our system has a configuration manifold  $Q$ . We take  $Q \times Q$  as our discrete version of the tangent bundle. The idea is that on a computer, we represent velocities by finite differences, and finite differences are naturally represented by pairs of points.

**Definition 3.3.1.** A *discrete path* is an ordered set of points  $q = \{q_1, \dots, q_n\} \in Q^n$ . We denote the space of paths of length  $n$  by  $C_n^D(Q) := Q^n$ .

**Definition 3.3.2.** A *discrete Lagrangian* is a map  $\mathcal{L}^D : Q \times Q \rightarrow \mathbb{R}$ .

**Definition 3.3.3.** Given a discrete Lagrangian  $\mathcal{L}^D$ , we obtain a *discrete action functional*  $S^D : C_n^D(Q) \rightarrow \mathbb{R}$  defined by

$$S^D[q] := \sum_{i=1}^{n-1} \mathcal{L}^D(q_i, q_{i+1}) \quad (3.27)$$

Now, we obtain the trajectory of our system by applying Hamilton's principle of least action. This time, the variations are fairly simple to compute. Up to a boundary term, we have

$$\delta S^D[q] \cdot \delta q = \sum_{k=1}^{N-1} (D_2 \mathcal{L}^D(q_{k-1}, q_k) + D_1 \mathcal{L}^D(q_k, q_{k+1})) \delta q_k \quad (3.28)$$

where  $D_i$  denotes differentiation with respect to the  $i$ th argument.

By requiring that this variation vanish for all variations  $\delta q$ , we find our discrete Euler-Lagrange equations. They are simply

$$D_2 \mathcal{L}^D(q_{k-1}, q_k) + D_1 \mathcal{L}^D(q_k, q_{k+1}) = 0 \quad (3.29)$$

And a careful consideration of the boundary term in  $\delta S^D$  yields expression for discrete momentum maps which are conserved by a discrete version of Noether's theorem. See [16] for more details.

## Part II

## Results

# Chapter 4

## Discretization

### 4.1 Discrete Interior Product

Recall that in the continuous setting, we have the identity

$$\langle\!\langle \iota_{\alpha^\sharp} \beta, \gamma \rangle\!\rangle = \langle\!\langle \beta, \alpha \wedge \gamma \rangle\!\rangle \quad (4.1)$$

This identity was a key part of our derivation of the MHD conservation laws in section 2.4.2.

In the discrete setting, we can use this identity to *define* a discrete interior product. Since we have a discrete wedge product of 1-forms, and a discrete Hodge inner product of forms, we can define  $\iota_{\alpha^\sharp} \beta$  to be the unique form whose inner product with  $\gamma$  is  $\langle\!\langle \beta, \alpha \wedge \gamma \rangle\!\rangle$  for any  $\gamma$ . This uniquely defines the interior product of a discrete 1-form  $\alpha$  with a discrete 2-form  $\beta$ . Explicitly, this discrete interior product is given by the expressions<sup>1</sup>

$$(\iota_{\alpha^\sharp} \beta)_{i,j}^h = -\frac{1}{2} \left( \beta_{ij} \alpha_{i+1/2,j}^v + \beta_{i,j-1} \alpha_{i+1/2,j-1}^v \right) \quad (4.2)$$

$$(\iota_{\alpha^\sharp} \beta)_{i,j}^v = \frac{1}{2} \left( \beta_{i-1,j} \alpha_{i-1,j+1/2}^h + \beta_{i,j} \alpha_{i,j+1/2}^h \right) \quad (4.3)$$

This interior product was proposed in [27] (definition 8.2.2) as an ‘algebraic interior product’. It was used in [28] to discretize the incompressible Navier-Stokes equations. In the context of [27], this definition of the interior product is problematic, as the interior product of a general vector with a differential form should not depend on the metric, but this definition does. However, in our case, the definition is quite natural, as we already need the metric to define  $\alpha^\sharp$  anyway.

We can view the discretization of MHD presented by Kraus and Maj in [10] as a discretization using the standard DEC operators on a grid along with this choice of discretization of the interior product. Kraus and Maj take the inner product of the MHD equations with Lagrange multipliers to obtain a formal Lagrangian, and use the Hodge-adjointness of  $\wedge$  and  $\iota$  to obtain a formal Lagrangian written purely in terms of the standard DEC operators.

---

<sup>1</sup>For an explanation and derivation, see Appendix D

This simply defines the discrete interior product by the relation we used above. When Kraus and Maj differentiate the formal Lagrangian to obtain the Euler-Lagrange equations for the primal variables, they simply obtain a discretization of the MHD equations which uses our discrete interior product.

We that our integrator satisfies discrete conservation laws in Chapter 5.

## 4.2 Solving the Discrete Equations

Now, we have discrete versions of all of the operators involved in the ideal MHD equations. So our discrete MHD equations look exactly the same as our continuous 2D ideal MHD equations.

$$\dot{\eta} + \iota_{\eta^\sharp} d\eta - \iota_{b^\sharp} db + dp = 0 \quad (4.4)$$

$$\dot{b} + \delta(b \wedge \eta) = 0 \quad (4.5)$$

$$\delta\eta = 0 \quad (4.6)$$

This simple representation of the discrete equations is one of the advantages of discretizing with discrete exterior calculus.

Using the midpoint discretization at time step  $t + \frac{1}{2}$  with step size  $h$ , we let  $\dot{\eta} = \frac{1}{h}(\eta_{t+1} - \eta_t)$  and let  $\eta = \frac{1}{2}(\eta_{t+1} + \eta_t)$  (with analogous expressions for  $b$  and  $p$ ). Substituting these expressions into our MHD equations, we can solve for  $(\eta_{t+1}, b_{t+1}, p_{t+1})$  in terms of  $(\eta_t, b_t, p_t)$ .

This amounts to solving a system of nonlinear equations at each time step. To do so, we used Newton's method. Since all of the operators we use are either linear or bilinear, it is straightforward to find the Jacobian of our system which Newton's method requires.

$$\begin{matrix} \eta & \eta & b & p \\ b & \left[ \begin{array}{ccc} \frac{1}{h}I + \iota_{\bullet^\sharp} d\eta + \iota_{\eta^\sharp} d\bullet & \iota_{\bullet^\sharp} db + \iota_{b^\sharp} d\bullet & d \\ \delta(b \wedge \bullet) & \frac{1}{h}I - \delta(\eta \wedge \bullet) & 0 \\ \delta & 0 & 0 \end{array} \right] \\ p & & & \end{matrix} \quad (4.7)$$

A derivation of this expression is presented in Appendix E. We also provide an explicit formula for the Jacobian there.

Using this Jacobian, we can apply Newton's method to solve our nonlinear system. In our experiments, it usually converged within 10 steps. Because all of our DEC operators are sparse, this Jacobian is also sparse. However, it is not symmetric. So the linear solves were still fairly slow, and this prevented us from running large simulations or extending the method to 3D.

## 4.3 Grids with Boundary

We can define our DEC operators on a grid with boundary.  $d$  is still just  $\partial^T$ , the coboundary operator. The definition of the wedge product remains the same, because each of the squares

in the grid still has all of its boundary edges. The Hodge star still makes sense in this case, and this allows us to define the Hodge inner product for a grid with boundary. And using this Hodge inner product and wedge product, we can define the discrete interior product on grids with boundary. Now that we have defined all of the DEC operators on a grid with boundary, we can simulate MHD on this grid in exactly the same way that we performed simulations on our grid without boundary.

This implicitly enforces the constraint that the velocity field and magnetic field must be parallel to the boundary. Physically, this corresponds to a conducting fluid flowing around in a perfectly insulating container. Other boundary conditions would be more complicated to implement, but should be possible within the DEC framework.

# Chapter 5

## Conservation Behavior

Our discrete 2D MHD equations are

$$\dot{\eta} + \iota_{\eta^\sharp} d\eta - \iota_{b^\sharp} db + dp = 0 \quad (5.1)$$

$$\dot{b} + \delta(b \wedge \eta) = 0 \quad (5.2)$$

$$\delta\eta = 0 \quad (5.3)$$

We will show that these equations conserve energy and cross-helicity when we use the time discretization  $\dot{\eta} = \frac{1}{h}(\eta_{t+1} - \eta_t)$  and  $\eta = \frac{1}{2}(\eta_{t+1} + \eta_t)$ .

### 5.1 Energy

Recall that the energy at time  $t$  is given by

$$E_t := \langle\!\langle \eta_t, \eta_t \rangle\!\rangle + \langle\!\langle b_t, b_t \rangle\!\rangle \quad (5.4)$$

Let  $\dot{E} = \frac{1}{h}(E_{t+1} - E_t)$ . To show that energy is conserved, it suffices to show that  $\dot{E} = 0$ . Note that

$$\langle\!\langle \eta_{t+1}, \eta_{t+1} \rangle\!\rangle - \langle\!\langle \eta_t, \eta_t \rangle\!\rangle = \langle\!\langle \eta_{t+1}, \eta_{t+1} \rangle\!\rangle - \langle\!\langle \eta_{t+1}, \eta_t \rangle\!\rangle + \langle\!\langle \eta_{t+1}, \eta_t \rangle\!\rangle - \langle\!\langle \eta_t, \eta_t \rangle\!\rangle \quad (5.5)$$

$$= \langle\!\langle \eta_{t+1}, \eta_{t+1} - \eta_t \rangle\!\rangle + \langle\!\langle \eta_{t+1} - \eta_t, \eta_t \rangle\!\rangle \quad (5.6)$$

$$= 2h \langle\!\langle \dot{\eta}, \eta \rangle\!\rangle \quad (5.7)$$

Therefore, we see that

$$\frac{1}{2} \dot{E} = \langle\!\langle \dot{\eta}, \eta \rangle\!\rangle + \langle\!\langle \dot{b}, b \rangle\!\rangle \quad (5.8)$$

$$= \langle\!\langle -\iota_{\eta^\sharp} d\eta + \iota_{b^\sharp} db - dp, \eta \rangle\!\rangle + \langle\!\langle -\delta(b \wedge \eta), b \rangle\!\rangle \quad (5.9)$$

$$= -\langle\!\langle d\eta, \eta \wedge \eta \rangle\!\rangle + \langle\!\langle db, b \wedge \eta \rangle\!\rangle - \langle\!\langle p, \delta\eta \rangle\!\rangle - \langle\!\langle b \wedge \eta, db \rangle\!\rangle \quad (5.10)$$

$$= 0 \quad (5.11)$$

So energy is conserved

## 5.2 Cross-Helicity

We can perform a similar computation to show that cross-helicity is conserved. In 3D, our cross-helicity was given by  $\int \eta \wedge \beta$ . Since  $\beta = \star_{2D} b \wedge dz$ , the cross-helicity is  $\int \eta \wedge \star_{2D} b \wedge dz$ . So when we reduce to the 2D case, we get a cross-helicity function  $H = \int \eta \wedge \star b = \langle\!\langle \eta, b \rangle\!\rangle$ .

$$H_t := \langle\!\langle \eta_t, b_t \rangle\!\rangle = \langle\!\langle b_t, \eta_t \rangle\!\rangle \quad (5.12)$$

So

$$\dot{H} = \langle\!\langle \dot{\eta}, b \rangle\!\rangle + \langle\!\langle \eta, \dot{b} \rangle\!\rangle \quad (5.13)$$

$$= \langle\!\langle -\iota_{\eta^\sharp} d\eta + \iota_{b^\sharp} db - dp, b \rangle\!\rangle + \langle\!\langle \eta, -\delta(b \wedge \eta) \rangle\!\rangle \quad (5.14)$$

$$= -\langle\!\langle d\eta, \eta \wedge b \rangle\!\rangle + \langle\!\langle db, b \wedge b \rangle\!\rangle + \langle\!\langle p, \delta b \rangle\!\rangle - \langle\!\langle d\eta, b \wedge \eta \rangle\!\rangle \quad (5.15)$$

The first term cancels with the last term, the second term vanishes because  $b \wedge b = 0$ , and the third term vanishes because  $\delta b = 0$ . Thus, cross-helicity is conserved.

# Chapter 6

## Computational Examples

The proposed algorithm was implemented in Houdini 16, using numpy and scipy for sparse matrices and linear algebra. One step at  $20^2$  resolution takes approximately 3 seconds to run.

### 6.1 Alfvén Wave

Following Kraus and Maj [10], we consider an Alfvén wave in the  $x$  direction with initial conditions

$$\eta = (0, \sin(\pi x)), \quad b = (1, \sin(\pi x)) \quad (6.1)$$

We performed the simulation on a  $20 \times 20$  grid representing the region  $[-1, 1] \times [-1, 1]$ . We used a time step of 0.1 and ran the simulation for 100 steps. Our results are shown in Figure 6.2. Our implementation stored all data as floats, so a precision of  $10^{-8}$  is essentially machine precision.

### 6.2 MHD with boundary

We picked some arbitrary initial conditions  $\tilde{\eta} = (\tilde{\eta}_x, \tilde{\eta}_y)$  and  $\tilde{b} = (\tilde{b}_x, \tilde{b}_y)$ , and used pressure projection to obtain divergence-free initial conditions. We set

$$\tilde{\eta}_x = \begin{cases} 1 & y > \frac{1}{3} \\ -1 & -1/3 < y < 1/3 \\ 1 & y < -1/3 \end{cases} \quad \tilde{\eta}_y = \begin{cases} 1 & x > \frac{1}{3} \\ -1 & -1/3 < x < 1/3 \\ 1 & x < -1/3 \end{cases} \quad (6.2)$$

$$\tilde{b}_x = \begin{cases} 1 & y > \frac{1}{3} \\ -1 & -1/3 < y < 1/3 \\ 1 & y < -1/3 \end{cases} \quad \tilde{b}_y = 0 \quad (6.3)$$

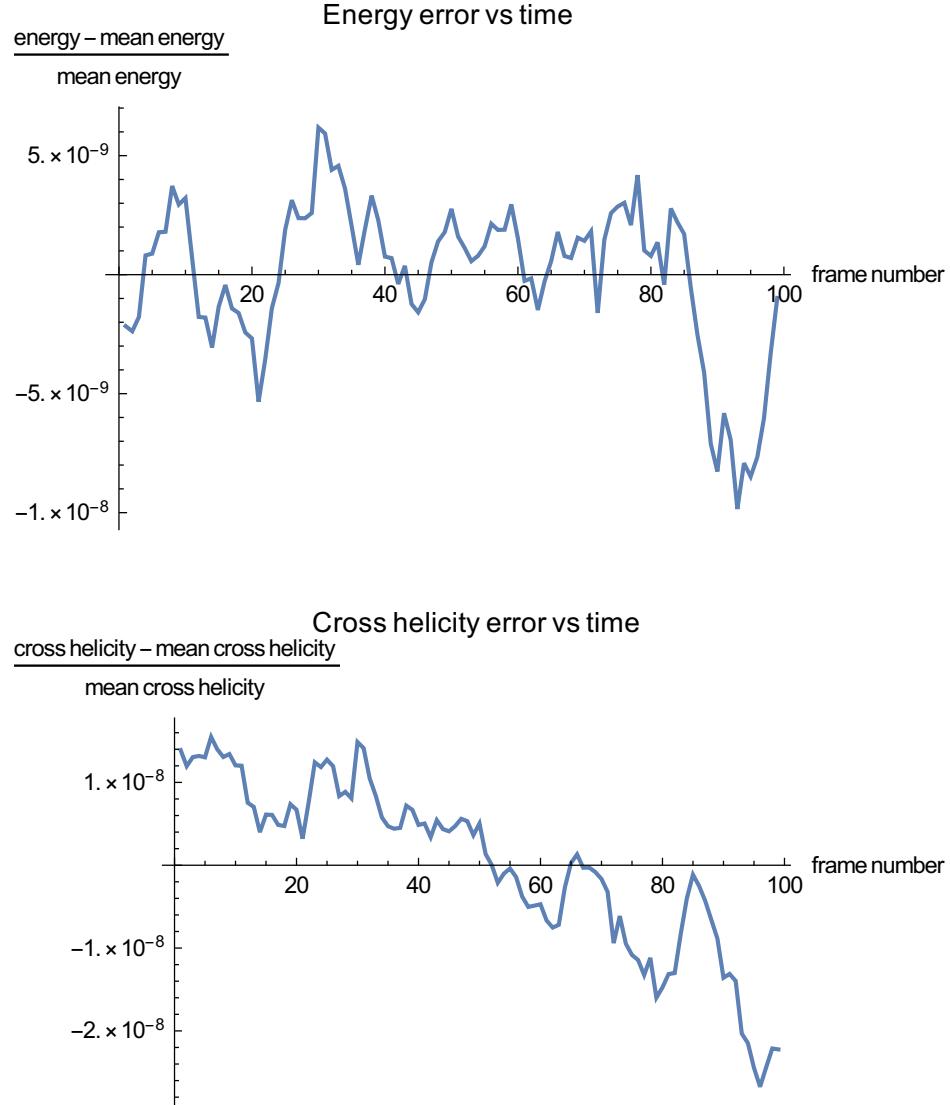


Figure 6.1: Energy and cross-helicity drift of the Alfvén wave simulation

We performed the simulation on a  $20 \times 20$  grid representing the region  $[-1, 1] \times [-1, 1]$ . We used a time step of 0.1 and ran the simulation for 100 steps. Our results are shown in Figure 6.3. Our energy drift is on the order of machine precision. The cross-helicity drift is slightly higher, but still very small.

### 6.2.1 Plume

We picked another set of initial conditions which cause a rising plume of fluid inside of an enclosed container. We used the following initial velocity vector field  $v$  and magnetic vector fields  $b$ , and then used the discrete flat operator to turn them into differential forms, and pressure-projected to remove the divergence.

$v_x$  and  $b_x$  are both zero everywhere.  $v_y$  and  $b_y$  are zero everywhere except in the rectangle

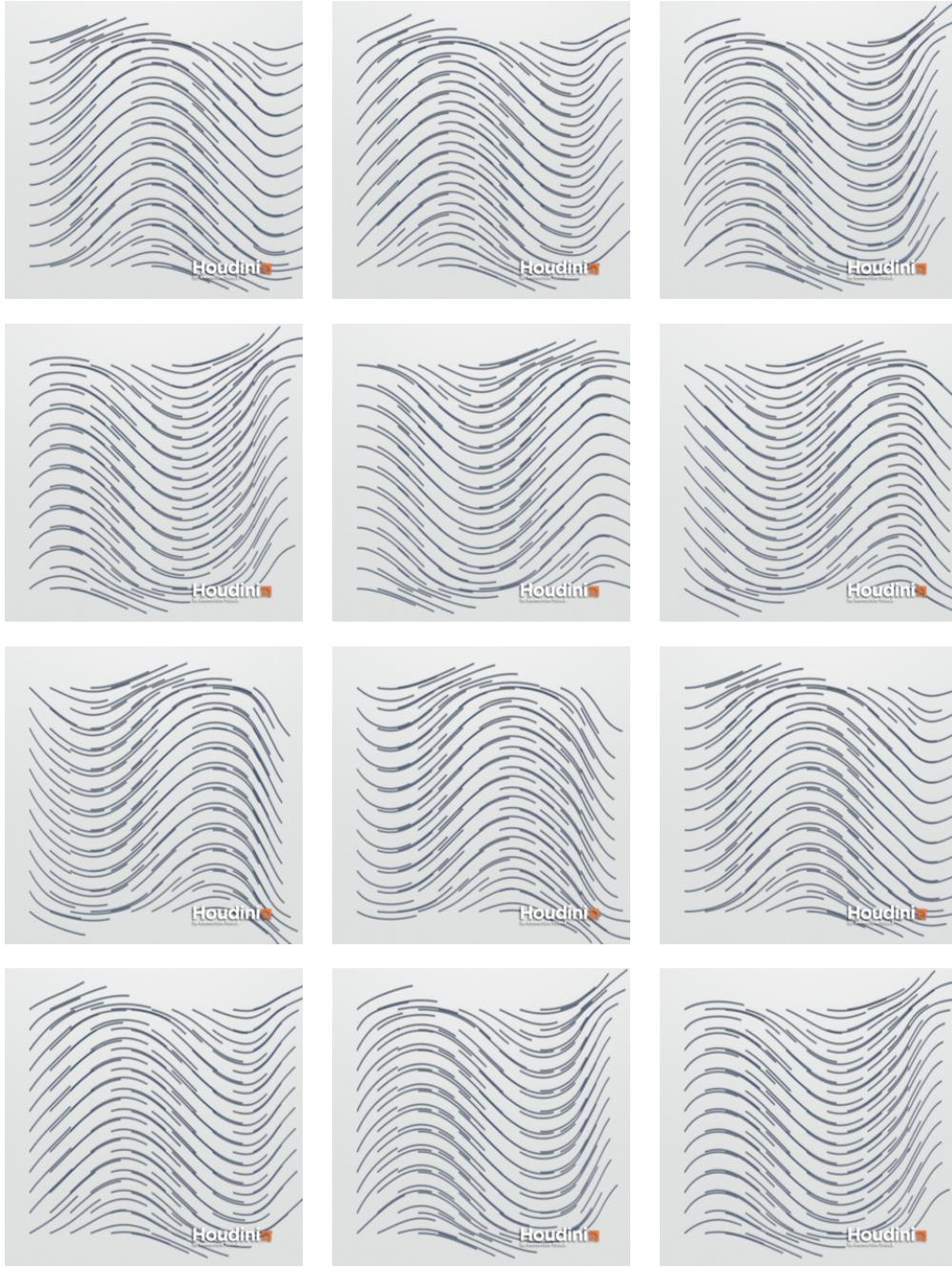


Figure 6.2: Magnetic field visualized for the first 12 frames of an Alfvén wave simulation

$[-0.1, 0.1] \times [-0.8, -0.2]$ . Inside this rectangle,  $b_y = 1$ , and  $v_y = x(1 - x)$ .

We then ran a simulation on a  $20 \times 20$  grid representing the region  $[-1, 1] \times [-1, 1]$  for 50 steps. Our results are shown in Figure 6.4. Again, the energy and cross-helicity are almost conserved at machine precision.

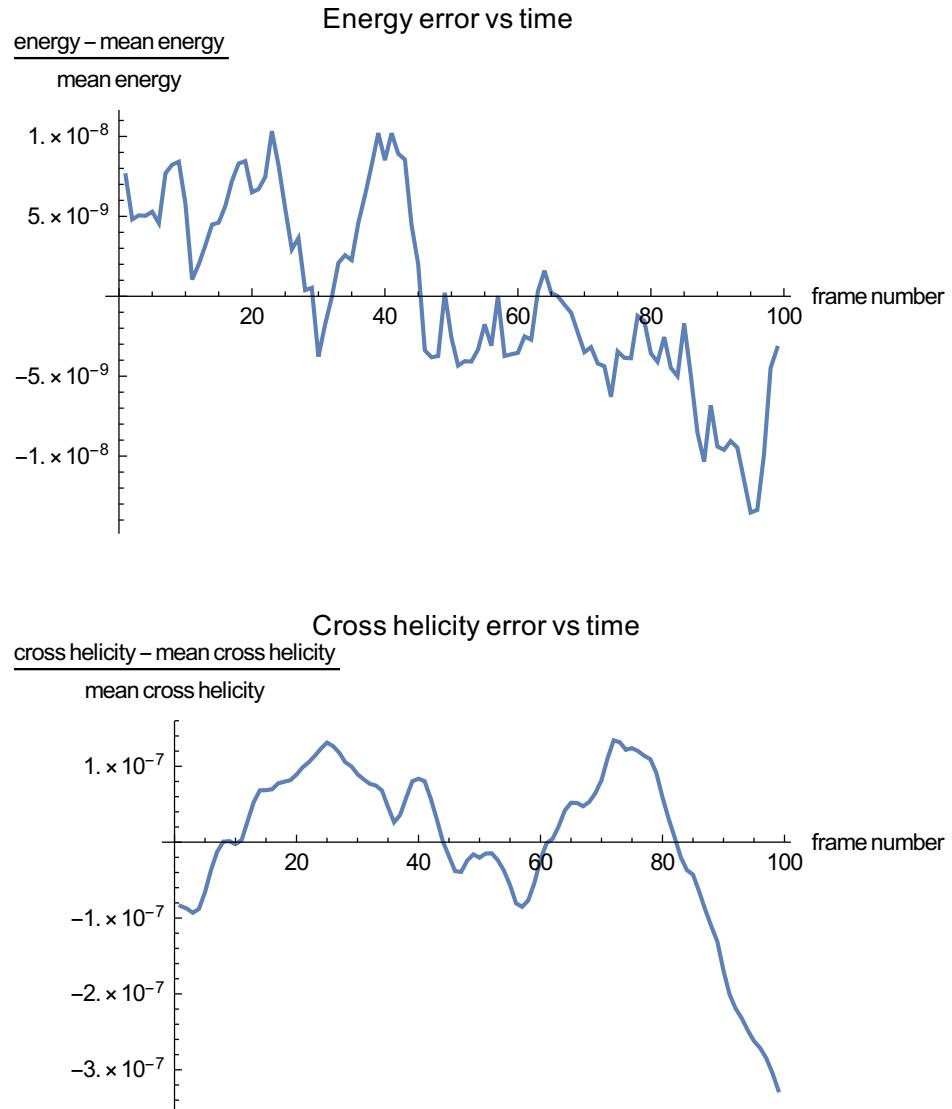


Figure 6.3: Energy and cross-helicity drift of the simulation of MHD in a bounded box

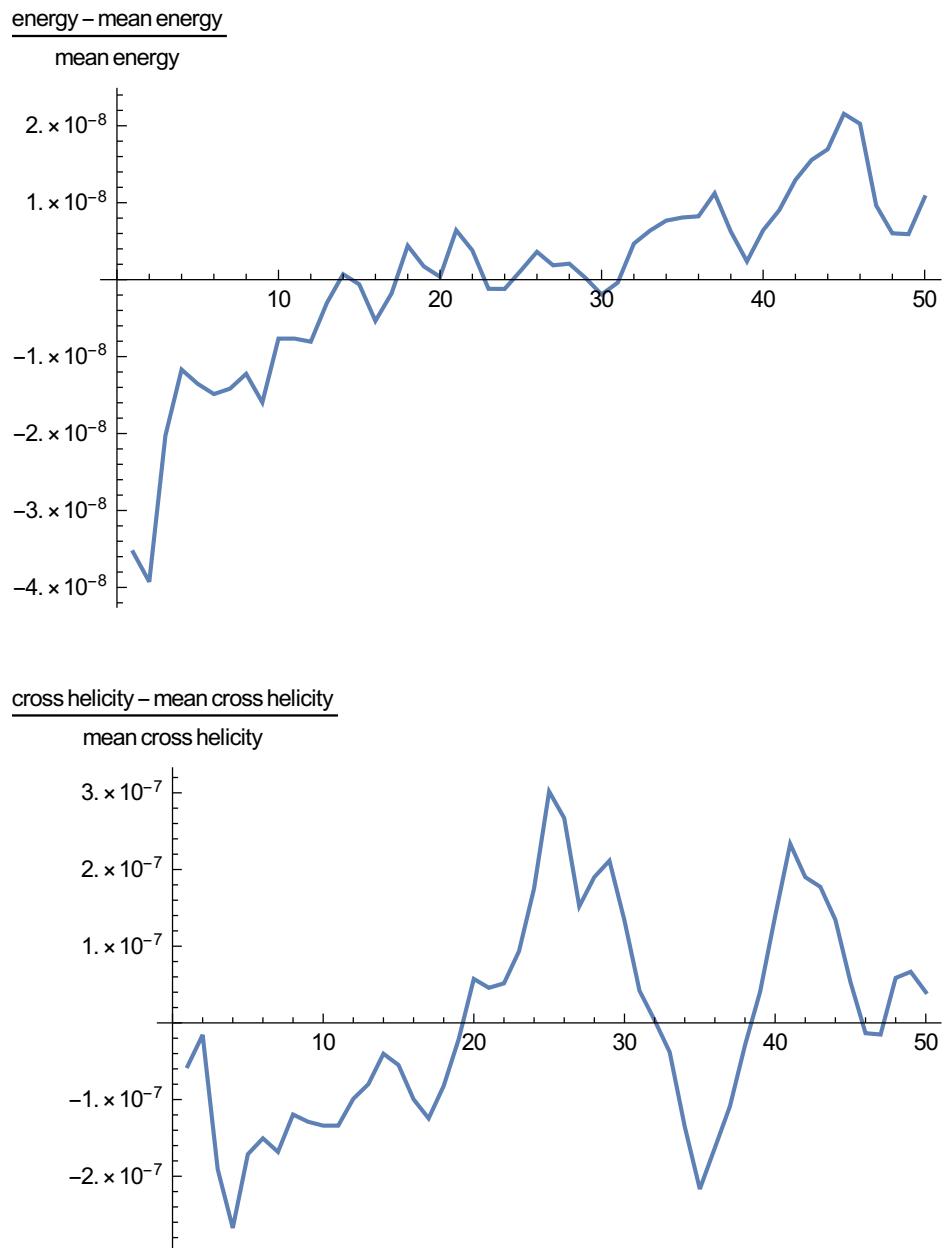


Figure 6.4: Energy and cross-helicity drift of the simulation of MHD in the plume in a bounded box

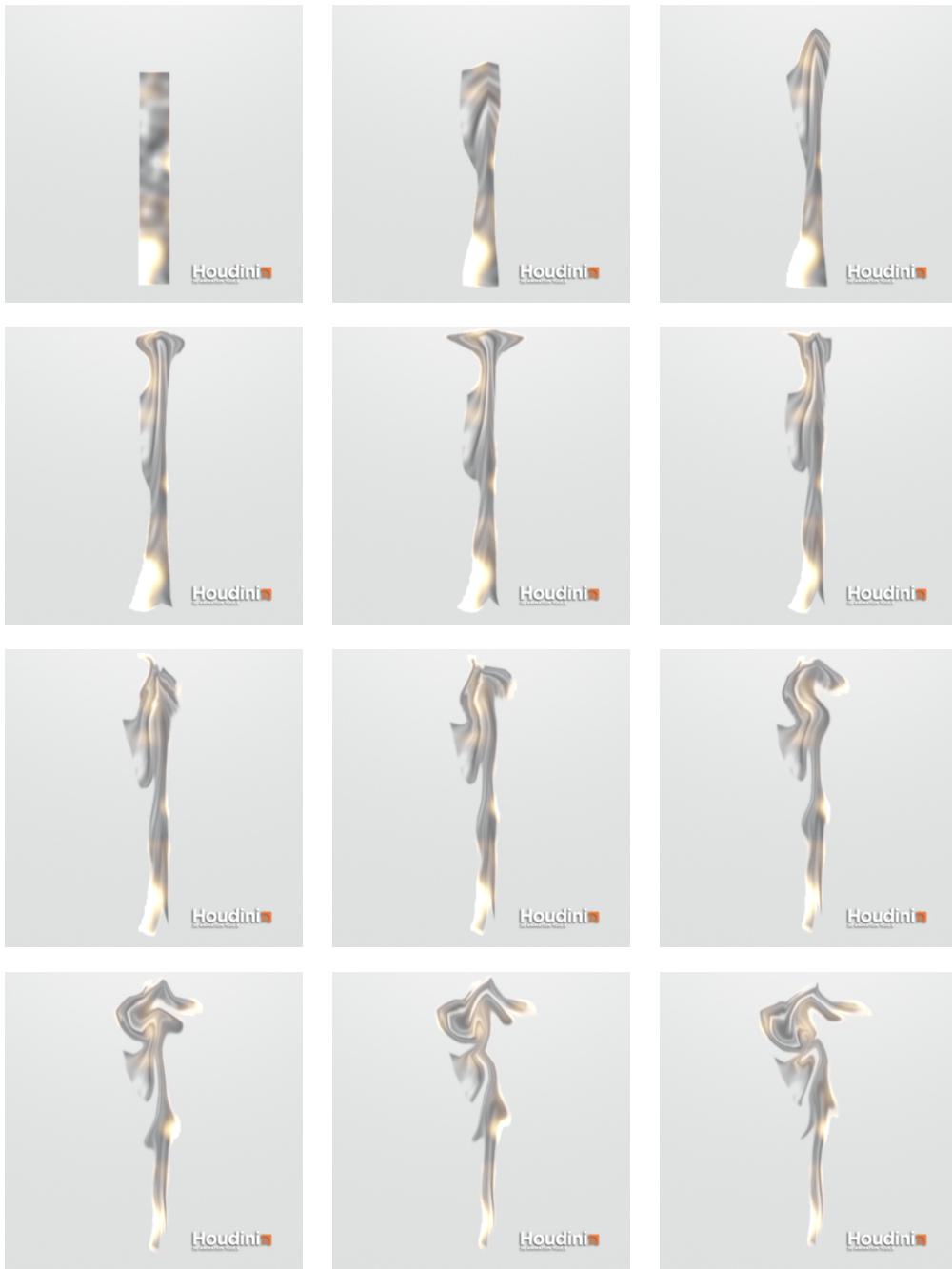


Figure 6.5: Plume visualized every frame for the first 12 frames. The visualizations are made by advecting a block of particles in the velocity field. The glowing is proportional to the square of the magnetic field strength.

# Index

- 0-form, 10
- 1-form, 10
- action, 18, 21
- Cartan's magic formula, 33
- category
  - of manifolds, 6
- closed, 12
- cochain, 40
- codifferential, 41
  - discrete, 43
- configuration manifold, 21
- configuration space
  - discrete, 46
- cotangent bundle, 10
- cross helicity, 34
- current, 32
- derivation, 4
- differential form, 8, 9
  - discrete, 40
- energy, 33
- Euler equations, 32
- Euler-Lagrange equation, 19
- Euler-Lagrange map, 23
- exterior derivative, 12
  - discrete, 40
- Faraday tensor, 31
- flat, 11
- Hodge inner product, 13
- Hodge star, 13
  - discrete, 41
- ideal fluid, 32
- incompressible, 32
- infinitesimal generator, 25
- inner product
  - discrete, 42
- iterated integral, 11
- jacobian, 49
- Lagrange multipliers, 20
- Lagrangian, 21
  - discrete, 46
  - field theory, 27
  - mechanics, 18
- Lagrangian symplectic form, 25
- Laplace-de Rham, 15
- Laplacian, 14
- Leibniz rule, 12
- magnetic helicity, 34
- manifold, 2
- MHD equations, 33
  - 2-dimensional, 37
- momentum map, 25
- monad, 6
- musical isomorphism, 11
- pressure, 32
- principle of least action, 23
- pullback
  - of differential forms, 9
  - of functions, 10
- pushforward
  - of tangent bundles, 6
  - of tangent vectors, 4
- Riemannian metric, 7
- second-order submanifold, 21
- sharp, 11
- state space, 21
- Stokes' theorem, 12, 40

tangent bundle, 5

tangent vector, 3

tensor

    alternating, 9

variation, 19, 22

vector field, 6

wedge product, 11

# Bibliography

- [1] Enrique Vazquez-Semadeni. Interstellar MHD Turbulence and Star Formation. *arXiv:1208.4132 [astro-ph]*, 407:401–444, 2015. arXiv: 1208.4132.
- [2] Interstellar Medium and Galactic Center | University of Iowa Astronomy and Astrophysics | Robert Mutel.
- [3] Kazunari Shibata and Tetsuya Magara. Solar Flares: Magnetohydrodynamic Processes. *Living Reviews in Solar Physics*, 8(1):6, December 2011.
- [4] Karen C Fox. Giant prominence erupts, 2012.
- [5] Earth’s Inconstant Magnetic Field | Science Mission Directorate.
- [6] E. J. Doyle et. al. Chapter 2: Plasma confinement and transport. *Nuclear Fusion*, 47(6):S18, 2007.
- [7] Jian-Guo Liu and Wei-Cheng Wang. An Energy-Preserving MACâŠYee Scheme for the Incompressible MHD Equation. *Journal of Computational Physics*, 174(1):12–37, November 2001.
- [8] Yao Zhou, Hong Qin, J. W. Burby, and A. Bhattacharjee. Variational integration for ideal magnetohydrodynamics with built-in advection equations. *Physics of Plasmas*, 21(10):102109, October 2014. arXiv: 1408.1346.
- [9] E. S. Gawlik, P. Mullen, D. Pavlov, J. E. Marsden, and M. Desbrun. Geometric, variational discretization of continuum theories. *Physica D: Nonlinear Phenomena*, 240(21):1724–1760, October 2011.
- [10] Michael Kraus and Omar Maj. Variational Integrators for Ideal Magnetohydrodynamics, July 2017.
- [11] Yao Zhou. Variational Integration for Ideal Magnetohydrodynamics and Formation of Current Singularities. *arXiv:1708.08523 [astro-ph, physics:physics]*, August 2017. arXiv: 1708.08523.
- [12] J. M. Lee. *Introduction to Smooth Manifolds*. Springer, 2003.
- [13] M. Spivak. *A Comprehensive Introduction to Differential Geometry*. Publish or Perish, Incorporated, 1975.

- [14] Jerrold E. Marsden, George W. Patrick, and Steve Shkoller. Multisymplectic geometry, variational integrators, and nonlinear PDEs. *Communications in Mathematical Physics*, 199(2):351–395, December 1998. arXiv: math/9807080.
- [15] Mark J. Gotay, James Isenberg, Jerrold E. Marsden, and Richard Montgomery. Momentum Maps and Classical Relativistic Fields. Part I: Covariant Field Theory. *arXiv:physics/9801019*, January 1998. arXiv: physics/9801019.
- [16] J. E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numerica*, 10:357–514, May 2001.
- [17] A. Lew, J. E. Marsden, M. Ortiz, and M. West. Asynchronous Variational Integrators. *Archive for Rational Mechanics and Analysis*, 167(2):85–146, April 2003.
- [18] Ari Stern, Yiying Tong, Mathieu Desbrun, and Jerrold E. Marsden. Geometric Computational Electrodynamics with Variational Integrators and Discrete Differential Forms. *arXiv:0707.4470 [physics]*, 73:437–475, 2015. arXiv: 0707.4470.
- [19] V. I. Arnold and Boris A. Khesin. *Topological methods in hydrodynamics*. Number 125 in Applied mathematical sciences. Springer, New York, corr. 2nd print edition, 1999.
- [20] Terence Tao. Conserved quantities for the Euler equations, February 2014.
- [21] Michael Kraus and Omar Maj. Variational Integrators for Nonvariational Partial Differential Equations. *Physica D: Nonlinear Phenomena*, 310:37–71, August 2015. arXiv: 1412.2011.
- [22] Nail H. Ibragimov. A new conservation theorem. *Journal of Mathematical Analysis and Applications*, 333(1):311–328, September 2007.
- [23] Nail H. Ibragimov. Integrating factors, adjoint equations and Lagrangians. *Journal of Mathematical Analysis and Applications*, 318(2):742–757, June 2006.
- [24] Mathieu Desbrun, Anil N. Hirani, Melvin Leok, and Jerrold E. Marsden. Discrete Exterior Calculus. *arXiv:math/0508341*, August 2005. arXiv: math/0508341.
- [25] Scott O. Wilson. On the Algebra and Geometry of a Manifold’s Chains and Cochains.
- [26] P. R. Kotiuga. Theoretical Limitations of Discrete Exterior Calculus in the Context of Computational Electromagnetics. *IEEE Transactions on Magnetics*, 44(6):1162–1165, June 2008.
- [27] Anil Nirmal Hirani. *Discrete exterior calculus*. phd, California Institute of Technology, 2003.
- [28] Mamdouh S. Mohamed, Anil N. Hirani, and Ravi Samtaney. Discrete exterior calculus discretization of incompressible Navier-Stokes equations over surface simplicial meshes. *Journal of Computational Physics*, 312:175–191, May 2016. arXiv: 1508.01166.

- [29] Dmitry Pavlov. *Structure-preserving discretization of incompressible fluids*. phd, California Institute of Technology, 2009.
- [30] Zensho Yoshida and Philip J. Morrison. Epi-two-dimensional flow and generalized enstrophy. *arXiv:1604.02339 [physics]*, April 2016. arXiv: 1604.02339.
- [31] Peter J. Olver. Jets and Differential Invariants.
- [32] P. Mullen, A. McKenzie, D. Pavlov, L. Durant, Y. Tong, E. Kanso, J. E. Marsden, and M. Desbrun. Discrete Lie Advection of Differential Forms. *Foundations of Computational Mathematics*, 11(2):131–149, April 2011.
- [33] D. Pavlov, P. Mullen, Y. Tong, E. Kanso, J. E. Marsden, and M. Desbrun. Structure-preserving discretization of incompressible fluids. *Physica D: Nonlinear Phenomena*, 240(6):443–458, March 2011.
- [34] Günter Schwarz. *Hodge decomposition: a method for solving boundary value problems*. Number 1607 in Lecture notes in mathematics. Springer-Verlag, Berlin ; New York, 1995.
- [35] Sharif Elcott, Yiying Tong, Eva Kanso, Peter Schröder, and Mathieu Desbrun. Stable, Circulation-preserving, Simplicial Fluids. *ACM Trans. Graph.*, 26(1), January 2007.
- [36] Máximo Bañados and Ignacio A. Reyes. A short review on Noether’s theorems, gauge symmetries and boundary terms. *International Journal of Modern Physics D*, 25(10):1630021, September 2016. arXiv: 1601.03616.
- [37] Peter J. Olver. *Applications of Lie groups to differential equations*. Number 107 in Graduate texts in mathematics. Springer-Verlag, New York, 1986.
- [38] Peter J. Olver. *Equivalence, invariants, and symmetry*. Cambridge University Press, Cambridge ; New York, 1995.
- [39] William L. Burke. *Applied differential geometry*. Cambridge University Press, Cambridge [Cambridgeshire] ; New York, 1985.
- [40] Theodore Frankel. *The geometry of physics: an introduction*. Cambridge University Press, New York, 2nd ed edition, 2004.
- [41] Scott O. Wilson. Geometric Structures on the Cochains of a Manifold. *arXiv:math/0505227*, May 2005. arXiv: math/0505227.
- [42] Roldao da Rocha and Waldyr A. Rodrigues Jr. Pair and Impar, Even and Odd Form Fields and Electromagnetism. *Annalen der Physik*, 19(1-2):6–34, February 2010. arXiv: 0811.1713.
- [43] J. M. Lee. *Riemannian Manifolds: An Introduction to Curvature*. Springer New York, 1997.
- [44] M. P. do Carmo. *Riemannian Geometry*. Birkhäuser, 1992.

- [45] Albert Ren-Haur Chern. *Fluid Dynamics with Incompressible Schrödinger Flow*. phd, California Institute of Technology, 2017.

# Appendices

# Appendix A

## Lagrange Multiplier Example: Pendulum

As a simple example of the power of Lagrange multipliers, consider the case of 2D pendulum. While it is simple to parameterize this system by the pendulum's angle, it is instructive to work through the system using Lagrange multipliers instead. For more complicated systems (e.g. three pendula in 3D), the Langrange multiplier method is much simpler than finding a suitable set of parameters for the system and using the Euler Lagrange equations to find equations of motion for those parameters.

Consider an idealized pendulum of length  $\ell$  swinging in a 2D space. The state of the pendulum is characterized by the position and velocity of the mass at the end, which we will denote  $x$  and  $\dot{x}$  respectively (note that  $x$  is a vector). If the mass were a free particle instead of a pendulum, its Lagrangian would simply be its kinetic energy minus its gravitational potential energy,  $\frac{1}{2}m\langle\dot{x}, \dot{x}\rangle - m\langle g, x\rangle$  (where  $g$  is the vector of gravitational acceleration). Since the mass is on the end of a pendulum, we add a constraint that its distance from the origin must be  $\ell$ . Equivalently, we can require that  $\langle x, x \rangle = \ell^2$ . So our Lagrangian for pendulum, including the Lagrange multiplier to enforce the constraint, is given by

$$\mathcal{L}(x, \lambda, \dot{x}, \dot{\lambda}) = \frac{1}{2}m\langle\dot{x}, \dot{x}\rangle - m\langle g, x\rangle - \lambda(\langle x, x \rangle - \ell^2) \quad (\text{A.1})$$

The Euler-Lagrange equations tell us that

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}} \quad (\text{A.2})$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\lambda}} \quad (\text{A.3})$$

Since our Lagrangian is independent of  $\dot{\lambda}$ , equation A.3 tells us that  $\frac{\partial \mathcal{L}}{\partial \lambda} = 0$ . Looking at our expression for  $\mathcal{L}$ , this precisely enforces our constraint that  $\|x\| = \ell$ .

Now, we turn our attention to equation A.2. From the definition of our Lagrangian, we see that  $\frac{\partial \mathcal{L}}{\partial x} = -mg - 2\lambda x$  and

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}} = \frac{d}{dt} m\dot{x} \quad (\text{A.4})$$

$$= m\ddot{x} \quad (\text{A.5})$$

Therefore, equation A.2 tells us that

$$m\ddot{x} + mg + 2\lambda x = 0 \quad (\text{A.6})$$

Now, we can solve for  $\lambda$  using our Euler-Lagrange equations. Dotting equation A.6 with  $x$ , we find that

$$m\langle \ddot{x}, x \rangle + m\langle g, x \rangle + 2\lambda \langle x, x \rangle = 0 \quad (\text{A.7})$$

Using the fact that  $\langle x, x \rangle = \ell^2$ , we see that

$$\lambda = -\frac{m}{2\ell^2} [\langle \ddot{x}, x \rangle + \langle g, x \rangle] \quad (\text{A.8})$$

Next, we note that because  $\langle x, x \rangle$  is constant, its second derivative must be zero. Taking this derivative, we find that

$$0 = \frac{d^2}{dt^2} \langle x, x \rangle \quad (\text{A.9})$$

$$= 2 \frac{d}{dt} \langle \dot{x}, x \rangle \quad (\text{A.10})$$

$$= 2 [\langle \ddot{x}, x \rangle + \langle \dot{x}, \dot{x} \rangle] \quad (\text{A.11})$$

Thus, we conclude that

$$\langle \ddot{x}, x \rangle = -\langle \dot{x}, \dot{x} \rangle \quad (\text{A.12})$$

Substituting this into equation A.8, we find that

$$\lambda = -\frac{m}{2\ell^2} [-\langle \dot{x}, \dot{x} \rangle + \langle g, x \rangle] \quad (\text{A.13})$$

Now, substituting this expression back into equation A.2, we obtain equations of motion

$$m\ddot{x} + mg + \frac{m}{\ell^2} \left( |\dot{x}|^2 - \langle g, x \rangle \right) x = 0 \quad (\text{A.14})$$

You can check that after switching to polar coordinates, you recover the standard equation for a pendulum.

**Aside:** Note that none of our computations actually used the fact that our pendulum was in 2D. So these equations of motion actually hold for a pendulum in arbitrary dimensions!

# Appendix B

## Exterior Calculus Identities

**Proposition 1.** Let  $M$  be a 3-dimensional manifold. Let  $\alpha \in \Omega^1(M)$  and  $\beta \in \Omega^2(M)$ . Furthermore, let  $a = \alpha^\sharp, b = (\star\beta)^\sharp$ . Then

$$(a \times b)^\flat = -\iota_\alpha \beta \quad (\text{B.1})$$

*Proof.* First, we note that

$$(a \times b)^\flat = \star(a^\flat \wedge b^\flat) \quad (\text{B.2})$$

$$= \star(\alpha \wedge \star\beta) \quad (\text{B.3})$$

$$= \star(\alpha \wedge \star\beta) \quad (\text{B.4})$$

So it is sufficient to show that  $-\iota_\alpha \beta = \star(\alpha \wedge \star\beta)$ . Furthermore, we recall that the fiberwise inner product on differential forms is nondegenerate. So it suffices to show that

$$\langle c, -\iota_\alpha \beta \rangle = \langle c, \star(\alpha \wedge \star\beta) \rangle \quad (\text{B.5})$$

for all  $c \in \Omega^1(M)$ . This is equivalent to showing that

$$\langle c, -\iota_\alpha \beta \rangle \mu = \langle c, \star(\alpha \wedge \star\beta) \rangle \mu \quad (\text{B.6})$$

where  $\mu$  is the volume form. We can show equation B.6 using the defining equation of the Hodge star operator.

$$\langle c, -\iota_\alpha \beta \rangle \mu = -\langle \alpha \wedge c, \beta \rangle \quad (\text{B.7})$$

$$= -\alpha \wedge c \wedge \star\beta \quad (\text{B.8})$$

$$= c \wedge \alpha \wedge \star\beta \quad (\text{B.9})$$

$$= c \wedge \star^2(\alpha \wedge \star\beta) \quad (\text{B.10})$$

$$= \langle c, \star(\alpha \wedge \star\beta) \rangle \mu \quad (\text{B.11})$$

□

## B.1 Conversion between forms on the plane and in $\mathbb{R}^3$

**Proposition 2.** Let  $\{dx, dy, dz\}$  be the standard basis for  $\Omega^1(\mathbb{R}^3)$ . Let  $\star_{3D}$  denote the Hodge star on  $\mathbb{R}^3$ , and let  $\star_{2D}$  denote the Hodge star on the  $xy$ -plane. Then for any 1-form  $\eta$  on the  $xy$ -plane, we have

$$\star_{3D} (\star_{2D}\eta \wedge dz) = \eta \quad (\text{B.12})$$

*Proof.* By the definition of the Hodge star, we have

$$dx \wedge dy \wedge dz = \star_{2D}\eta \wedge dz \wedge \star_{3D}(\star_{2D}\eta \wedge dz) \quad (\text{B.13})$$

$$= \star_{2D}\eta \wedge [-\star_{3D}(\star_{2D}\eta \wedge dz)] \wedge dz \quad (\text{B.14})$$

Since  $\star_{2D}\eta$  and  $\star_{3D}(\eta \wedge dz)$  must both be forms on the  $xy$ -plane, we conclude that

$$dx \wedge dy = \star_{2D}\eta \wedge [-\star_{3D}(\star_{2D}\eta \wedge dz)] \quad (\text{B.15})$$

This is the defining relation for  $\star_{2D}^2\eta$ . So we conclude that

$$\star_{3D}(\star_{2D}\eta \wedge dz) = -\star_{2D}^2\eta = \eta \quad (\text{B.16})$$

□

**Proposition 3.** Let  $\{dx, dy, dz\}$  be the standard basis for  $\Omega^1(\mathbb{R}^3)$ . Let  $\star_{3D}$  denote the Hodge star on  $\mathbb{R}^3$ , and let  $\star_{2D}$  denote the Hodge star on the  $xy$ -plane. Then for any 1-form  $\eta$  on the  $xy$ -plane, we have

$$\star_{3D}\eta = \star_{2D}\eta \wedge dz \quad (\text{B.17})$$

*Proof.* Again, we just use the definition of the Hodge star. We know that

$$\eta \wedge \star_{2D}\eta = dx \wedge dy \quad (\text{B.18})$$

Therefore,

$$\eta \wedge \star_{2D}\eta \wedge dz = dx \wedge dy \wedge dz \quad (\text{B.19})$$

Thus,  $\star_{2D}\eta \wedge dz$  satisfies the defining equation for  $\star_{3D}\eta$ . □

**Proposition 4.** Let  $\star$  denote the Hodge star on the  $xy$ -plane and let  $d$  denote the exterior derivative on the  $xy$ -plane. Then for any 1-forms  $\eta, b$  on the  $xy$ -plane, we have

$$\iota_{b^\#} d\eta = (\star d\eta) \star b \quad (\text{B.20})$$

*Proof.* Let  $c$  be any 1-form. Since the inner product is nondegenerate, it is sufficient to show that

$$\langle c, \iota_{b^\#} d\eta \rangle \mu = \langle c, (\star d\eta) \star b \rangle \mu \quad (\text{B.21})$$

Now, we show that the two sides are equal. We use the fact that  $(\star d\eta)$  is a 0-form, and we can thus treat wedge products with  $(\star d\eta)$  as scalar multiplication. Furthermore, we use that fact that in 2D, on 1-forms, we have  $\star^2 = -1$ .

$$\langle c, \iota_{b^\sharp} d\eta \rangle \mu = \langle b \wedge c, d\eta \rangle \mu \quad (\text{B.22})$$

$$= b \wedge c \wedge \star d\eta \quad (\text{B.23})$$

$$= (\star d\eta)(c \wedge \star^2 b) \quad (\text{B.24})$$

$$= \langle c, (\star d\eta) \star b \rangle \mu \quad (\text{B.25})$$

□

# Appendix C

## Formal Lagrangians

### C.1 Example of a mechanical system

Here, we will work through the formal Lagrangian construction in the case of a simple mechanical system. We will consider the harmonic oscillator  $\ddot{x} = -x$ . The formal Lagrangian construction we defined requires that our equations are first-order, so we will use the trick of introducing a new variable  $v = \dot{x}$ . This gives us two equations in two variables

$$\dot{x} = v \tag{C.1}$$

$$\dot{v} = -x \tag{C.2}$$

Now, we write down a Lagrangian which enforces these equations of motion

$$\mathcal{L}(x, \dot{x}, v, \dot{v}, \lambda, \dot{\lambda}, \mu, \dot{\mu}) := \lambda(\dot{x} - v) + \mu(\dot{v} + x) \tag{C.3}$$

Given this Lagrangian, we can use the Euler-Lagrange equations for  $x$  and  $v$  in order to find the equations of motion for  $\lambda$  and  $\mu$ .

$$\frac{\partial \mathcal{L}}{\partial x} = \mu \tag{C.4}$$

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}} = \frac{d}{dt} \lambda = \dot{\lambda} \tag{C.5}$$

$$\frac{\partial \mathcal{L}}{\partial v} = -\lambda \tag{C.6}$$

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{v}} = \frac{d}{dt} \mu = \dot{\mu} \tag{C.7}$$

The Euler-Lagrange equations tell us that

$$\dot{\lambda} = \mu \tag{C.8}$$

$$\dot{\mu} = -\lambda \tag{C.9}$$

These are just our original equations!

In the language we used before, we can define a map

$$\Phi(x, v) \mapsto (x, v, x, v) \quad (\text{C.10})$$

which takes solutions of our differential equations to solutions of our extended system.

# Appendix D

## Discrete Interior Product

We will compute an expression for the discrete interior product, using the identity

$$\langle\!\langle \iota_{\alpha} \beta, \gamma \rangle\!\rangle = \langle\!\langle \beta, \alpha \wedge \gamma \rangle\!\rangle \quad (\text{D.1})$$

Recall our standard grid, as shown in Figure D.1.

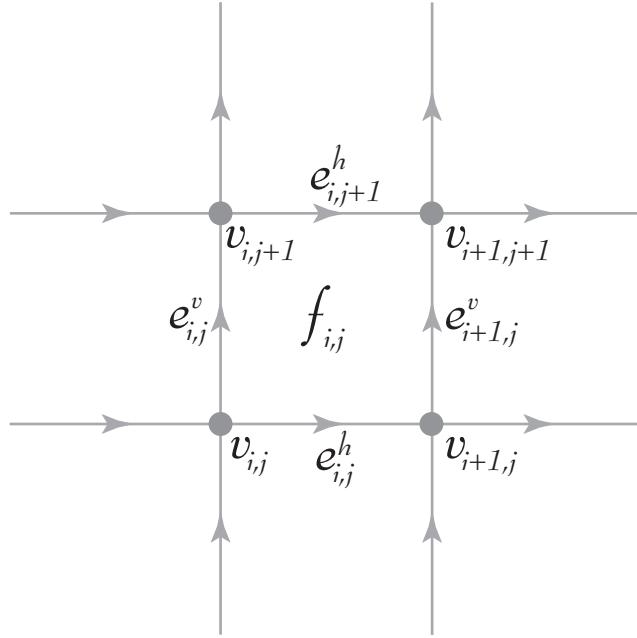


Figure D.1: Reference for grid element naming convention

The component of  $\iota_{\alpha} \beta$  on edge  $e_{i,j}^h$  can be computed as

$$\langle\!\langle \iota_{\alpha} \beta, e_{i,j}^h \rangle\!\rangle = \langle\!\langle \beta, \alpha \wedge e_{i,j}^h \rangle\!\rangle \quad (\text{D.2})$$

From the definition of the discrete wedge product, we see that  $\alpha \wedge e_{i,j}^h$  is only nonzero on faces  $f_{i,j}$  and  $f_{i,j-1}$ . So

$$\langle\!\langle \beta, \alpha \wedge e_{i,j}^h \rangle\!\rangle = b_{i,j}(\alpha \wedge e_{i,j}^h)_{i,j} + b_{i,j-1}(\alpha \wedge e_{i,j}^h)_{i,j-1} \quad (\text{D.3})$$

Using the formula for the discrete wedge product, we obtain

$$(\alpha \wedge e_{i,j}^h)_{i,j} = -\frac{1}{2}\alpha_{i+1/2,j}^v \quad (\alpha \wedge e_{i,j}^h)_{i,j-1} = -\frac{1}{2}\alpha_{i+1/2,j-1}^v \quad (\text{D.4})$$

where  $\alpha_{i+1/2,j}^v$  denotes the averaged quantity  $\frac{1}{2}(\alpha_{i,j}^v + \alpha_{i+1,j}^v)$ , etc.

Therefore, we conclude that

$$(\iota_{\alpha^\sharp}\beta)_{i,j}^h = -\frac{1}{2} \left( \beta_{ij}\alpha_{i+1/2,j}^v + \beta_{i,j-1}\alpha_{i+1/2,j-1}^v \right) \quad (\text{D.5})$$

By a similar argument, we can compute the component of  $\iota_{\alpha^\sharp}\beta$  on edge  $e_{i,j}^v$ . It is simply

$$(\iota_{\alpha^\sharp}\beta)_{i,j}^v = \langle\!\langle \iota_{\alpha^\sharp}\beta, e_{i,j}^v \rangle\!\rangle \quad (\text{D.6})$$

$$= \langle\!\langle \beta, \alpha \wedge e_{i,j}^v \rangle\!\rangle \quad (\text{D.7})$$

$$= \beta_{i-1,j}(\alpha \wedge e_{i,j}^v)_{i-1,j} + \beta_{i,j}(\alpha \wedge e_{i,j}^v)_{i,j} \quad (\text{D.8})$$

$$= \beta_{i-1,j} \cdot \frac{1}{2}\alpha_{i-1,j+1/2}^h + \beta_{i,j} \cdot \frac{1}{2}\alpha_{i,j+1/2}^h \quad (\text{D.9})$$

$$= \frac{1}{2} \left( \beta_{i-1,j}\alpha_{i-1,j+1/2}^h + \beta_{i,j}\alpha_{i,j+1/2}^h \right) \quad (\text{D.10})$$

# Appendix E

## Derivation of the Jacobian

Our discrete MHD equations are given by

$$\dot{\eta} + \iota_{\eta^\sharp} d\eta - \iota_{b^\sharp} db + dp = 0 \quad (\text{E.1})$$

$$\dot{b} + \delta(b \wedge \eta) = 0 \quad (\text{E.2})$$

$$\delta\eta = 0 \quad (\text{E.3})$$

Using the midpoint discretization at time step  $t + \frac{1}{2}$ , we let  $\dot{\eta} = \frac{1}{h}(\eta_{t+1} - \eta_t)$  and let  $\eta = \frac{1}{2}(\eta_{t+1} + \eta_t)$  (with analogous expressions for  $b$  and  $p$ ).

To solve the system for time stepping, we need to compute the Jacobian of these equations with respect to  $(\eta_{t+1}, b_{t+1}, p_{t+1})$ . We can do so by differentiating term-by-term with respect to each of the variables. This gives us the Jacobian as a block matrix.

### E.1 Velocity Equation

Differentiating the first term is simple.

$$\begin{aligned} \frac{\partial}{\partial \eta_{t+1}} \dot{\eta} &= \frac{\partial}{\partial \eta_{t+1}} \left[ \frac{1}{h}(\eta_{t+1} - \eta_t) \right] \\ &= \frac{1}{h} I \end{aligned}$$

where  $I$  is the identity matrix. Clearly the derivatives with respect to  $b$  and  $p$  are 0.

The second term is more complicated. First, we note that the expression  $\iota_\alpha^\sharp d\beta^\sharp$  is bilinear in  $\alpha$  and  $\beta$ . So

$$\begin{aligned} \iota_{\eta^\sharp} d\eta &= \frac{1}{4} \iota_{(\eta_{t+1} + \eta_t)^\sharp} d(\eta_{t+1} + \eta_t) \\ &= \frac{1}{4} \left[ \iota_{\eta_{t+1}^\sharp} d\eta_{t+1} + \iota_{\eta_{t+1}^\sharp} d\eta_t + \iota_{\eta_t^\sharp} d\eta_{t+1} + \iota_{\eta_t^\sharp} d\eta_t \right] \end{aligned}$$

The derivative of the last term with respect to  $\eta_{t+1}$  is clearly 0. The middle two terms are linear functions of  $\eta_{t+1}$ , so they are their own derivatives. The first term is the result of plugging  $\eta_{t+1}$  into a bilinear function, so we can differentiate it with the product rule. So we conclude that

$$\begin{aligned}\frac{\partial}{\partial \eta_{t+1}}(\iota_{\eta^{\sharp}} d\eta) &= \frac{1}{4} \left[ \iota_{\eta_{t+1}^{\sharp}} d\bullet + \iota_{\bullet^{\sharp}} d\eta_{t+1} + \iota_{\bullet^{\sharp}} d\eta_t + \iota_{\eta_t^{\sharp}} d\bullet \right] \\ &= \frac{1}{2} \left[ \iota_{\eta^{\sharp}} d\bullet + \iota_{\bullet^{\sharp}} d\eta \right]\end{aligned}$$

This also shows us that

$$\frac{\partial}{\partial b_{t+1}}(\iota_b^{\sharp} db) = \frac{1}{2} [\iota_b^{\sharp} d\bullet + \iota_{\bullet^{\sharp}} db]$$

And the fourth term is a linear function of  $p$ , so its derivative with respect to  $p$  is just  $d$ . Thus, we see that the first (block) row of our Jacobian is

$$\begin{bmatrix} \eta & b & p \\ \frac{1}{h} I + \iota_{\bullet^{\sharp}} d\eta + \iota_{\eta^{\sharp}} d\bullet & \iota_{\bullet^{\sharp}} db + \iota_b^{\sharp} d\bullet & d \end{bmatrix} \quad (\text{E.4})$$

Note that we've ignored some factors of two. This is okay, since we're looking for zeros, so we can just multiply the whole equation by 2. This adds on a factor of 2 to the pressure and the time step, but that is not a problem.

## E.2 Magnetic 1-Form Equation

Each of the terms in this equation is a linear function of  $\eta, b$  (taken individually). So each term is its own derivative. i.e.

$$\frac{\partial}{\partial \eta_{t+1}} [\delta(b \wedge \eta)] = \frac{\partial}{\partial \eta_{t+1}} \left[ \delta \left( b \wedge \frac{1}{2}(\eta_{t+1} - \eta_t) \right) \right] \quad (\text{E.5})$$

$$= \frac{1}{2} \delta(b \wedge \bullet) \quad (\text{E.6})$$

$$\frac{\partial}{\partial b_{t+1}} [\delta(b \wedge \eta)] = -\frac{\partial}{\partial b_{t+1}} [\delta(\eta \wedge b)] \quad (\text{E.7})$$

$$= -\frac{1}{2} \delta(\eta \wedge \bullet) \quad (\text{E.8})$$

$$\frac{\partial}{\partial b_{t+1}} b = \frac{1}{h} I \quad (\text{E.9})$$

Thus, we see that the second row of our Jacobian is given by

$$\begin{bmatrix} \eta & b & p \\ \delta(b \wedge \bullet) & \frac{1}{h} I - \delta(\eta \wedge \bullet) & 0 \end{bmatrix} \quad (\text{E.10})$$

where again, we have multiplied through by 2 to remove the  $\frac{1}{2}$  factors.

### E.3 Divergence-Free Constraint

Since  $\delta$  is linear, differentiating  $\delta\eta$  is simple. We just get that the third row of our Jacobian is

$$\begin{bmatrix} \eta & b & p \\ \delta & 0 & 0 \end{bmatrix} \quad (\text{E.11})$$

# Appendix F

## Assorted Vector and Matrix Derivatives

### F.1 Vector Derivatives

Throughout the section, let  $x \in \mathbb{R}^n$ .

$$\frac{\partial |x|}{\partial x} = \frac{x}{|x|} \quad (\text{F.1})$$

$$\frac{\partial}{\partial x} \frac{1}{|x|} = -\frac{x}{|x|^3} \quad (\text{F.2})$$

#### Proof of equation F.1

*Proof.* Write  $|x|$  as  $\sqrt{\langle x, x \rangle}$ . Now, we apply the chain rule

$$\frac{\partial \sqrt{\langle x, x \rangle}}{\partial x} = \frac{1}{2\sqrt{\langle x, x \rangle}} \frac{\partial \langle x, x \rangle}{\partial x} \quad (\text{F.3})$$

The desired result follows from the fact that  $\frac{\partial \langle x, x \rangle}{\partial x} = 2x$ .  $\square$

#### Proof of equation F.2

*Proof.* Applying the quotient rule, we see that

$$\frac{\partial}{\partial x} \frac{1}{|x|} = -\frac{1}{|x|^2} \frac{\partial}{\partial x} |x| \quad (\text{F.4})$$

The desired result follows from equation F.1  $\square$

### F.2 Matrix Derivatives

$$\frac{\partial x^T M x}{\partial x} = (M + M^T)x \quad (\text{F.5})$$

### Proof of equation F.5

*Proof.* The product rule tells us that

$$\frac{\partial x^T Mx}{\partial x} = \bullet^T Mx + x^T M\bullet \quad (\text{F.6})$$

(recall that the derivative of a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is another linear functional  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ). If we write this linear functional in the usual form (a column vector), we obtain

$$\frac{\partial x^T Mx}{\partial x} = Mx + M^T x \quad (\text{F.7})$$

□