

Companion Notes on Daniel Freed's *Quantum Theory from a Geometric Viewpoint*

Brennan Halcomb, Gary Hu, Rauan Kaldybayev, Theodore Mollano

Supervised by Ivo Terek

May 24, 2025

Abstract

This work aims to fill in the gaps and provide additional details in the first four chapters of Daniel Freed's notes [13], offering a more complete and accessible treatment of the material presented there.

We begin by introducing the Data of States and Observables (DOSO) framework, an axiomatic structure encoding states as convex sets and observables as complex topological vector spaces, linking measurement, dynamics, and probability in a common language applicable to both classical and quantum theories.

Building upon this, we develop the geometric foundations through a study of Riemannian and symplectic manifolds, highlighting their algebraic structures, global topological constraints, and canonical examples such as cotangent bundles. The theory of symmetries and moment maps is explored, emphasizing their role in conserved quantities and connections to geometric quantization.

The foundations of Hamiltonian and Lagrangian mechanics are then examined from a geometric and variational viewpoint. We rigorously formulate Hamiltonian flows on symplectic phase spaces and variational principles on path spaces, showcasing the connection between the Lagrangian and Hamiltonian formalisms via differential forms, conserved quantities, and symplectic structures on solution spaces.

Finally, we present an introduction to spectral theory, bridging finite-dimensional linear algebra and infinite-dimensional operator analysis. Key results include spectral decompositions via projection-valued measures, functional calculi, and the spectral theorem for self-adjoint operators. These concepts underpin the mathematical formulation of quantum observables and dynamics, linking to unitary group representations and C^* -algebraic structures.

Contents

1	States and Observables	4
1.1	Our Basic Axioms	5
1.2	DOSO for Classical Mechanics	17
1.3	DOSO for Quantum Mechanics	22
2	Riemannian and Symplectic Manifolds	27
2.1	Bilinear Forms in Linear Algebra	29
2.2	Euclidean Spaces and Affine Symplectic Spaces	37
2.3	Riemannian and Symplectic Manifolds	52
2.4	Symmetries and Moment Maps	62
3	Hamiltonian and Lagrangian Mechanics	68
3.1	Hamiltonian Systems	69
3.2	Variational Principles for Geodesics	73
3.3	Lagrangian Systems	76
3.4	Geometric Interpretation of γ	80
4	Spectral Theorems	82
4.1	Linear Operators in Finite Dimensions	83
4.2	Projection-Valued Measures and Functional Calculus for Diagonalizable Operators	86
4.3	Generalizations and Special Cases in Finite Dimensions	100
4.4	Spectral Theory on Hilbert Space	113

About

Introduction: This text grew out of a reading course at Williams College, MATH 497: Geometry and Quantum Theory, during the Spring 2025 term. The course was based on reading [13], and we met once a week for two hours to discuss the material in depth.

These notes were prepared collaboratively by four undergraduate students: Brennan Halcomb, Gary Hu, Rauan Kaldybayev, Theodore Mollano. We gratefully acknowledge the guidance and support of our advisor, Ivo Terek, whose expertise and encouragement greatly enriched this project.

Goal: Daniel Freed discusses quantum and classical physics from a geometric viewpoint. In our notes, we aim to fill in some of the gaps in [13], providing additional background, details, and motivation where helpful.

Background: We assume familiarity with functional analysis and differential geometry at the level of introductory graduate courses. A basic understanding of classical and quantum mechanics will also be helpful.

Remarks on Specific Chapters:

- **Chapter 1:** States and Observables

Freed begins by introducing DOSO (Data of States and Observables), an abstract framework that unifies classical, quantum, and statistical mechanics in a precise, formal way. It specifies some general properties we expect from a physical system. We recommend skimming Lecture 1 initially and returning to it as needed when specific concepts arise in later chapters. This approach lets you focus first on the more concrete and accessible material before revisiting the formal foundations.

- **Chapter 2:** Riemannian and Symplectic Manifolds

This chapter takes a geometric detour to contrast Riemannian and symplectic structures. Rather than developing either theory in depth, we focus on their foundational similarities and differences. The material sets the stage for the symplectic viewpoint on classical mechanics developed in the next chapter.

- **Chapter 3:** Hamiltonian and Lagrangian Mechanics

We approach classical mechanics from a geometric perspective, prioritizing geometric structure over physical intuition. Rather than modeling real-world systems, we use symplectic and Riemannian geometry to explore the underlying mathematics. Physical examples appear only to clarify abstract ideas; for a more physics-centered approach, traditional physics texts are recommended.

- **Chapter 4:** Spectral Theorems

Instead of looking at spectral theorems just from an analytic point of view, we also consider the algebraic and geometric parts to tell a more connected story. We also briefly explore generalizations using harmonic analysis, including ideas like Pontryagin duality.

1 States and Observables

Abstract

We begin by introducing the *Data of States and Observables* (DOSO) framework, a foundational axiomatic system for the mathematical description of physical systems. DOSO aims to unify core concepts of states, observables, measurement, and dynamics across different physical theories. It axiomatizes a state space \mathcal{S} as a real convex set (pure states being its extreme points) and a space of potential observables \mathcal{O} as a complex topological vector space with structures like a real structure, a Lie algebra on a dense subspace, and a functional calculus. A measurement pairing links observables and states to probability distributions, and a mechanism for dynamics describes transformations generated by observables.

We then illustrate how to apply the DOSO framework to classical mechanics. Here, pure states correspond to points in phase space, observables to real-valued functions, the Lie algebra to Poisson brackets, and dynamics to Hamiltonian flows. The framework naturally accommodates statistical classical mechanics through probability measures on phase space.

Finally, we discuss how to apply DOSO to quantum mechanics. States are represented by density operators on a Hilbert space (pure states by rank-one projectors), observables by self-adjoint operators, the Lie algebra by commutators, and dynamics by unitary evolution. The abstract DOSO structure provides a common language compare classical and quantum mechanics.

Contents

1.1	Our Basic Axioms	5
1.1.1	Data of States and Observables (DOSO)	5
1.1.2	Axiom 1: The Space of States (\mathcal{S})	6
1.1.3	Axiom 2: The Algebra of Observables (\mathcal{O})	7
1.1.4	Axiom 3: Dynamics and Symmetries	10
1.1.5	Summary of DOSO	12
1.1.6	Compatibilities Among DOSO Data	12
1.1.7	Expectation Values and Separability	14
1.1.8	Further Remarks on the Axiom System	14
1.1.9	Motion and Mechanical Systems	15
1.2	DOSO for Classical Mechanics	17
1.2.1	A Particle in Euclidean Space	17
1.2.2	DOSO Data for the Classical Particle	18
1.2.3	The Symplectic Form for the Particle Example	20
1.2.4	Dynamical Systems: Local versus Global Perspectives	21
1.3	DOSO for Quantum Mechanics	22
1.3.1	Hilbert Spaces and Planck's Constant	22
1.3.2	DOSO Data for Quantum Systems	23

References: For foundational texts on the mathematical foundations of quantum mechanics, see [7], [35], [31], [22]. For more modern treatments of quantum theory, see [16], [34], and [10]. For essential mathematical tools like functional analysis, see [5], [28], [30], [26], [6], and [25]. For classical mechanics, including Hamiltonian systems and symplectic geometry, see [1], [32], and [24]. For geometric tools related to Lie groups, see [36] and [15].

1.1 Our Basic Axioms

1.1.1 Data of States and Observables (DOSO)

Describing physical reality mathematically requires a precise language, which we develop here. Central to any physical theory are the notions of *state* (a complete description of a system at a given time), *observable* (a measurable property), and *dynamics* (how states or observables evolve in time). We adopt a general axiomatic framework encompassing classical mechanics, statistical mechanics, and, of particular interest, quantum mechanics and quantum field theory.

A physical system consists of its states, observables, and the pairing that assigns a value to each observable in a given state. For example, fix an integer $k > 0$ and consider k particles moving in Euclidean space. A state is the collection of the particles' positions and velocities at a fixed time. An observable is a real-valued function on trajectories: for instance, the distance between the 2nd and 13th particles at 4:45 AM on January 10, 2023.

Now that we have the intuition down, we describe our axiomatization of states and observables.

Definition 1.1 (Data of States and Observables (DOSO)). The data describing a physical system in terms of its states and observables comprises the following structures and their interactions:

1. (States.) A real convex space \mathcal{S} . Elements of \mathcal{S} are *states*, extreme points of \mathcal{S} are *pure states*, and non-extreme points are *mixed states*. Let $\mathcal{PS} \subset \mathcal{S}$ denote the subset of pure states.
2. (Observables.) A complex topological vector space \mathcal{O} called the *algebra of observables*¹, equipped with a *real structure*, i.e., an antilinear involution $A \mapsto A^*$. Elements of

$$\mathcal{O}_{\mathbb{R}} = \{A \in \mathcal{O} \mid A^* = A\},$$

the subspace of real (or self-adjoint) elements of \mathcal{O} , are called *observables*. There are the following additional structures.

- (a) A *measurement pairing*:

$$\begin{aligned} \mathcal{O}_{\mathbb{R}} \times \mathcal{S} &\longrightarrow \text{Prob}(\mathbb{R}) \\ (A, \sigma) &\mapsto \sigma_A \end{aligned}$$

where $\text{Prob}(\mathbb{R})$ is the space of probability measures on \mathbb{R} (equipped with its Borel σ -algebra).

- (b) A map, called *functional calculus* for observables:

$$\begin{aligned} \text{Borel}(\mathbb{R}; \mathbb{R}) \times \mathcal{O}_{\mathbb{R}} &\longrightarrow \mathcal{O}_{\mathbb{R}} \\ (f, A) &\mapsto f(A). \end{aligned}$$

3. (Dynamics.) A dense subspace $\mathcal{O}^{\infty} \subseteq \mathcal{O}$ equipped with a complex Lie algebra structure. Let $\mathcal{O}_{\mathbb{R}}^{\infty}$ be the space of real elements of \mathcal{O}^{∞} . For each $A \in \mathcal{O}_{\mathbb{R}}^{\infty}$, there exist

- (a) a one-parameter group of automorphisms of \mathcal{S} ;
- (b) a one-parameter group of automorphisms of \mathcal{O} .

Now, we briefly explain the purpose of each part of the axiomatization. Simultaneously, we introduce the full set of mathematical tools that this axiom provides us access to.

¹It is $\mathcal{O}_{\mathbb{R}}$ whose elements are observables. We consider the larger vector space \mathcal{O} for mathematical convenience.

1.1.2 Axiom 1: The Space of States (\mathcal{S})

Axiom 1: A real convex space \mathcal{S} . Elements of \mathcal{S} are states, extreme points of \mathcal{S} are pure states, and non-extreme points are mixed states. Let $\mathcal{PS} \subset \mathcal{S}$ denote the subset of pure states.

Remark 1.2. Typically, \mathcal{S} is a convex subset of a real topological vector space.

Remark 1.3. In addition to making possible a unified treatment of pure states and mixed states, the introduction of a convex structure opens the door to tools from convexity theory, such as the Krein-Milman Theorem² and Choquet's Theorem³.

Before we discuss why convex structure constitutes part of the “correct axiomatization,” let’s first define all of the relevant terms.

Definition 1.4.

- Let V be a real vector space. A subset $C \subseteq V$ is called a *convex set* if for any two elements $x, y \in C$ and for any real number $t \in [0, 1]$, the element $tx + (1-t)y$ is also in C . The expression $tx + (1-t)y$ is called a *convex combination* of x and y .
- Let C be a convex set. An element $e \in C$ is called an *extreme point* of C if it cannot be written as a non-trivial convex combination of two *distinct* points in C . More formally, $e \in C$ is an extreme point if for any $x, y \in C$ and any $t \in (0, 1)$ (note the open interval), the condition $e = tx + (1-t)y$ implies that $x = y = e$. A *non-extreme point* of a convex set \mathcal{S} is any element $\sigma \in \mathcal{S}$ that is not an extreme point.

Given two states σ_1 and σ_2 with probabilities p_1 and p_2 such that $p_1 + p_2 = 1$, there should be a state σ' representing the system being in σ_1 with probability p_1 and in σ_2 with probability p_2 . We call this a statistical mixture and write it as $p_1\sigma_1 + p_2\sigma_2$. Although it may not be obvious at first that statistical mixtures can be treated as linear combinations of states, this approach works well due to key properties of these mixtures. For instance, $p_1\sigma_1 + p_2\sigma_2 = p_2\sigma_2 + p_1\sigma_1$. Statistical mixtures play a fundamental role in statistical physics, and the framework of convex sets provides the proper mathematical setting for understanding and manipulating these combinations.

We choose the extreme points of the convex space to be our pure states because physically, pure states represent states of maximal knowledge. They are “indecomposable” in the sense that they cannot be viewed as statistical mixtures of other, different states. Identifying pure states is important as they often form the “building blocks” of the theory. Then, it naturally follows that we should define the mixed states as the non-extreme points.

Example 1.5 (Classical mechanics). In classical statistical mechanics, states are probability measures on the phase space, and the space of such measures is convex. A pure state is a Dirac measure

²A compact convex subset of a Hausdorff locally convex topological vector space is equal to the closed convex hull of its extreme points.

³Let C be a compact convex subset of a normed space V . Given $c \in C$, there exists a probability measure w supported on the set E of extreme points of C such that, for any affine function f on C ,

$$f(c) = \int f(e) dw(e).$$

at some phase space point, representing the system certainly being in that phase space configuration.

Example 1.6 (Quantum mechanics). In quantum mechanics, states are density operators ρ on a Hilbert space \mathcal{H} . These are positive trace-class operators⁴ with $\text{Tr}(\rho) = 1$. The set of such operators is convex. A pure state is equal to a projection $|\psi\rangle\langle\psi|$ for some unit vector $|\psi\rangle \in \mathcal{H}$. Such a state cannot be written as $\rho = p\rho_1 + (1-p)\rho_2$ for $0 < p < 1$ and $\rho_1, \rho_2 \neq \rho$ being other density operators.

Let's move onto the next axiom.

1.1.3 Axiom 2: The Algebra of Observables (\mathcal{O})

Axiom 2: A complex topological vector space \mathcal{O} called the algebra of observables⁵, equipped with a real structure, i.e., an antilinear involution $A \mapsto A^$. Elements of*

$$\mathcal{O}_{\mathbb{R}} = \{A \in \mathcal{O} \mid A^* = A\},$$

the subspace of real (or self-adjoint) elements of \mathcal{O} , are called observables. There are the following additional structures.

1. A measurement pairing:

$$\begin{aligned} \mathcal{O}_{\mathbb{R}} \times \mathcal{S} &\longrightarrow \text{Prob}(\mathbb{R}) \\ (A, \sigma) &\mapsto \sigma_A \end{aligned}$$

where $\text{Prob}(\mathbb{R})$ is the space of probability measures on \mathbb{R} (equipped with its Borel σ -algebra).

2. A map, called functional calculus for observables:

$$\begin{aligned} \text{Borel}(\mathbb{R}; \mathbb{R}) \times \mathcal{O}_{\mathbb{R}} &\longrightarrow \mathcal{O}_{\mathbb{R}} \\ (f, A) &\mapsto f(A). \end{aligned}$$

The space \mathcal{O} and its associated structures provide the mathematical representation of physical quantities that can be measured. There is quite a few properties to discuss, and we survey them one by one.

Definition 1.7 (Complex Topological Vector Space). A *complex topological vector space* is a vector space over the field of complex numbers \mathbb{C} that is endowed with a topology such that vector addition $(u, v) \mapsto u + v$ and scalar multiplication $(\lambda, v) \mapsto \lambda v$ are continuous functions.

Topological vector space structure. If A_1 and A_2 are observables and λ a real number, then, of course, we can define observables $A_1 + A_2$ and λA_1 – therefore, the mathematical object containing all observables should be a real vector space. Thus, we put observables into the real vector space \mathcal{O} .

Approximation is an invaluable technique in physics. Therefore, we require \mathcal{O} to be a *topological* vector space, meaning that there is a reasonable notion of continuity and limits. Any finite-dimensional vector space has an obvious topology, namely, the Euclidean one – yet one has to be careful in the

⁴Let \mathcal{H} be a separable Hilbert space. A (not necessarily positive) bounded operator $T : \mathcal{H} \rightarrow \mathcal{H}$ is called *trace-class* if and only if $\text{Tr}(|T|) < \infty$ where $|T| := \sqrt{T^*T}$.

⁵It is $\mathcal{O}_{\mathbb{R}}$ whose elements are observables. We consider the larger vector space \mathcal{O} for mathematical convenience.

infinite-dimensional case. Many observables, like position and momentum, are unbounded operators, requiring operator topologies (norm, strong, weak) for rigorous treatment. This topological structure enables the use of functional analysis, the most useful tool for infinite-dimensional systems in quantum mechanics and field theory.

Complex vector space structure. We go a step further and say that \mathcal{O} is a *complex* vector space. Making \mathcal{O} a *complex* vector space is motivated partly by quantum mechanics, a theory that is naturally expressed in complex vector spaces. Considering complex numbers is also mathematically convenient. Firstly, they offer a richer algebraic structure essential for tools like spectral theory and Fourier transforms. Secondly, complexifying function spaces, such as $C(X, \mathbb{C})$, simplifies analysis.

Example 1.8. In classical mechanics, \mathcal{O} might be a space of functions on phase space, like $C(M)$ (continuous functions on the phase space manifold M) or $L^p(M, d\mu)$ (functions whose p -th power is integrable with respect to a measure μ), which are often Banach spaces and thus topological vector spaces. In quantum mechanics, \mathcal{O} is typically related to the algebra $B(\mathcal{H})$ of bounded linear operators on a Hilbert space \mathcal{H} , or more generally, a C^* -algebra or a von Neumann algebra.

Definition 1.9 (Real Structure on a Complex Vector Space). A *real structure* on a complex vector space V is an antilinear map $J : V \rightarrow V$ (often denoted by $A \mapsto A^*$) such that $J^2 = \text{id}$ (i.e., $J(J(A)) = (A^*)^* = A$ for all $A \in V$). An antilinear map J satisfies $J(\lambda A + \mu B) = \bar{\lambda}J(A) + \bar{\mu}J(B)$ for all $A, B \in V$ and $\lambda, \mu \in \mathbb{C}$. The set of fixed points of J , $V_{\mathbb{R}} = \{A \in V \mid J(A) = A\}$, forms a real vector space.

Real structure. Complexifying \mathcal{O} comes with the obvious drawback that not every $A \in \mathcal{O}$ is an observable. Intuitively, the complex structure of \mathcal{O} would conflict with the fact that outcomes of physical measurements are always real numbers. (We will discuss measurement in later.) Complex numbers are nice to work with, yet the outcomes of physical measurements are always real numbers. The usefulness of a real structure lies in its ability to bridge the potentially complex algebraic framework of \mathcal{O} with the real-valued nature of physical measurements.

Example 1.10. In classical mechanics, if \mathcal{O} is taken as the space of complex-valued functions on phase space M , $C(M, \mathbb{C})$, then the star operation A^* corresponds to pointwise complex conjugation $f \mapsto \bar{f}$, and observables $\mathcal{O}_{\mathbb{R}}$ are the real-valued functions $C(M, \mathbb{R})$. In quantum mechanics, if \mathcal{O} is an algebra of operators on a Hilbert space \mathcal{H} (like $B(\mathcal{H})$), then A^* is the standard operator adjoint, and $\mathcal{O}_{\mathbb{R}}$ comprises the self-adjoint (Hermitian) operators.

This axiom formalizes the important connection between the abstract mathematical objects of the theory (states and observables) and the empirical results obtained from physical measurements. It essentially defines how to extract predictions for experimental outcomes from the theory.

Definition 1.11 (Probability Measure on \mathbb{R}). A *probability measure on \mathbb{R}* (more precisely, on the Borel σ -algebra of \mathbb{R} , denoted $\text{Borel}(\mathbb{R})$) is a function $P : \text{Borel}(\mathbb{R}) \rightarrow [0, 1]$ such that:

1. $P(E) \geq 0$ for all $E \in \text{Borel}(\mathbb{R})$ (non-negativity).
2. $P(\mathbb{R}) = 1$ (normalization).
3. For any countable sequence of pairwise disjoint Borel sets E_1, E_2, \dots , $P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$ (countable additivity or σ -additivity).

The space of all such probability measures on \mathbb{R} is denoted by $\text{Prob}(\mathbb{R})$.

Measurement pairing. This bullet point should come rather naturally. Observables are physical quantities that can be measured. If some observable is measured in some state, the outcome is not always deterministic – in quantum mechanics, for example, measurement includes an intrinsic element of randomness. For this reason, we say that measuring an observable A in a state σ produces results distributed according to some distribution σ_A .

Example 1.12. In classical mechanics, if the state σ is a pure state, corresponding to a specific point x_0 in phase space, and A is an observable represented by a function $A(x)$ on phase space, the measurement is deterministic. The resulting probability measure σ_A is then the Dirac delta measure $\delta_{A(x_0)}$, which assigns probability 1 to the single value $A(x_0)$ and 0 to any set not containing $A(x_0)$. If σ is a classical mixed state, represented by a probability measure μ on phase space (e.g., a Liouville distribution), then σ_A is the pushforward measure $A_*\mu$, defined by $(A_*\mu)(E) = \mu(A^{-1}(E))$ for any Borel set $E \subset \mathbb{R}$. This gives the probability distribution for the values of $A(x)$ when x is sampled according to μ . In quantum mechanics, the probabilistic nature is even more apparent. If σ is a pure state represented by a normalized vector $|\psi\rangle$ in a Hilbert space \mathcal{H} , and A is a self-adjoint operator, the spectral theorem associates A with a unique projection-valued measure (PVM) Π_A . The probability that a measurement of A yields a value in a Borel set $E \subset \mathbb{R}$ is then given by $\langle \psi, \Pi_A(E)\psi \rangle$. This expression defines the measure $\sigma_A(E)$. For a quantum mixed state described by a density operator ρ , this generalizes to $\sigma_A(E) = \text{Tr}(\rho\Pi_A(E))$. This formulation, often referred to as the Born rule (or its generalization), is very important in quantum theory.

Definition 1.13 (Borel Measurable Function). Let (X, \mathcal{T}) be a topological space, and let $\mathcal{B}(X)$ be the Borel σ -algebra on X (the smallest σ -algebra containing all open sets in \mathcal{T}). Let (Y, \mathcal{F}) be a measurable space (a set Y with a σ -algebra \mathcal{F} of subsets of Y). A function $h : X \rightarrow Y$ is *Borel measurable* (or simply Borel) if for every set $S \in \mathcal{F}$, the preimage $h^{-1}(S) = \{x \in X \mid h(x) \in S\}$ is in $\mathcal{B}(X)$. For the functional calculus, $f : \mathbb{R} \rightarrow \mathbb{R}$, both the domain and codomain are equipped with their standard Borel σ -algebras.

Functional calculus. Intuitively, if A is an observable and f is a function, it is natural that we can define an observable $f(A)$ as follows: for a physical measurement in which A would have been measured to have value $a \in \mathbb{R}$, the observable $f(A)$ is $f(a)$. The definition of a functional calculus for observables only makes this intuition precise.

The functional calculus, a map $(f, A) \mapsto f(A)$ from $\text{Borel}(\mathbb{R}; \mathbb{R}) \times \mathcal{O}_{\mathbb{R}}$ to $\mathcal{O}_{\mathbb{R}}$, allows for the construction of new real observables from existing ones by applying real-valued Borel measurable functions. If A is an observable representing a physical quantity, and $f : \mathbb{R} \rightarrow \mathbb{R}$ is such a function, then $f(A)$ is also considered a well-defined observable. The choice of Borel functions is significant because this class is sufficiently broad to include most functions of practical interest in physics, such as polynomials (e.g., forming A^2), exponentials (e.g., e^{cA}), and characteristic functions of intervals or other Borel sets (e.g., $\chi_E(A)$, which acts as a projector onto measurement outcomes of A lying in the set E). Furthermore, the set of Borel functions is well-behaved with respect to pointwise limits of sequences of functions.

Example 1.14. In classical mechanics, if an observable A is represented by a real-valued function $A(q, p)$ on phase space, then $f(A)$ is simply the composite function $(f \circ A)(q, p)$. In quantum

mechanics, for a self-adjoint operator A , the functional calculus $f(A)$ is rigorously defined via the spectral theorem. If A has a spectral decomposition $A = \int_{\sigma(A)} \lambda dP_\lambda$, where P_λ is its spectral measure and $\sigma(A)$ its spectrum, then $f(A)$ is defined as $f(A) = \int_{\sigma(A)} f(\lambda) dP_\lambda$.

1.1.4 Axiom 3: Dynamics and Symmetries

Axiom 3: A dense subspace $\mathcal{O}^\infty \subseteq \mathcal{O}$ equipped with a complex Lie algebra structure. Let $\mathcal{O}_{\mathbb{R}}^\infty$ be the space of real elements of \mathcal{O}^∞ . For each $A \in \mathcal{O}_{\mathbb{R}}^\infty$, there exist

1. *a one-parameter group of automorphisms of \mathcal{S} ;*
2. *a one-parameter group of automorphisms of \mathcal{O} .*

This axiom introduces the notion of evolution or transformation within the system, establishing how changes, such as time evolution or symmetry operations, are generated by certain well-behaved real observables. It provides the framework for dynamics and continuous symmetries.

Definition 1.15 (Complex Lie Algebra). A *complex Lie algebra* is a vector space L over the complex numbers \mathbb{C} , equipped with a binary operation $[\cdot, \cdot] : L \times L \rightarrow L$, called the Lie bracket, satisfying the following properties for all $A, B, C \in L$ and $\alpha, \beta \in \mathbb{C}$:

1. Bilinearity: $[\alpha A + \beta B, C] = \alpha[A, C] + \beta[B, C]$ and $[A, \alpha B + \beta C] = \alpha[A, B] + \beta[A, C]$.
2. Antisymmetry: $[A, B] = -[B, A]$.
3. Jacobi Identity: $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$.

The existence of a complex Lie algebra structure is a natural axiom motivated by the evolution of observables in the Heisenberg picture, which we talk about later. In classical mechanics, the Poisson bracket $\{f, g\}$ plays a similar role in Hamiltonian dynamics and canonical transformations. The “complex” nature of the Lie algebra makes the bracket \mathbb{C} -bilinear. Compatibility conditions like $[A_1, A_2]^* = [A_1^*, A_2^*]$ ensure the real structure interacts coherently with the Lie bracket, implying that the real subspace $\mathcal{O}_{\mathbb{R}}^\infty$ forms a real Lie subalgebra. This Lie algebra structure is essential for understanding the non-commutative nature of observables, central to both classical and quantum Hamiltonian mechanics, continuous symmetries (via Noether’s theorem), and quantization procedures that relate classical Poisson algebras to quantum Lie algebras.

Definition 1.16 (Dense Subspace). Let X be a topological space and S be a subspace of X (i.e., $S \subseteq X$ with the induced topology). S is a *dense subspace* of X if the closure of S in X is X itself ($\bar{S} = X$). Equivalently, for every point $x \in X$ and every open neighborhood U of x , $U \cap S \neq \emptyset$.

Putting the Lie algebra structure a dense subspace $\mathcal{O}^\infty \subset \mathcal{O}$ addresses the common situation in physics where certain algebraic operations (like forming Lie brackets) or analytical procedures (such as differentiation with respect to parameters) are well-defined or exhibit better behavior only on a smaller, more “regular” or “smoother” subset of observables. This \mathcal{O}^∞ represents such a well-behaved subset. The property of density is critical because it ensures that properties defined on \mathcal{O}^∞ , or elements within it, can often be extended by continuity or approximation to the entirety of \mathcal{O} . Thus, this dense subspace provides a well-behaved domain for structures like the Lie algebra, which might not be definable or manageable on all of \mathcal{O} .

Example 1.17. In classical mechanics, if \mathcal{O} is a space like $L^2(M)$ or $C_0(M)$ (continuous functions vanishing at infinity), \mathcal{O}^∞ might be $C^\infty(M)$, the space of smooth functions on phase space. Poisson brackets, important to Hamiltonian dynamics, are naturally and rigorously defined on $C^\infty(M)$, which is dense in many relevant larger function spaces. In quantum mechanics, if dealing with unbounded operators, \mathcal{O}^∞ could be a common, dense domain of important self-adjointness for a set of operators, or it might represent operators that are “smooth” with respect to a given dynamical evolution. For instance, the commutator of two unbounded self-adjoint operators might not be densely defined or lead to another self-adjoint operator without careful consideration of domains, necessitating the selection of a suitable \mathcal{O}^∞ .

Definition 1.18 (One-Parameter Group of Automorphisms). A *one-parameter group of automorphisms* on a space X (either \mathcal{S} or \mathcal{O}) is a family of automorphisms $\{\phi_t\}_{t \in \mathbb{R}}$ indexed by a real parameter t (often representing time or a continuous symmetry parameter) such that:

1. $\phi_0 = \text{id}$ (the identity map on X).
2. $\phi_t \circ \phi_s = \phi_{t+s}$ for all $t, s \in \mathbb{R}$ (group homomorphism property).
3. The map $t \mapsto \phi_t(x)$ is continuous for each $x \in X$ (continuity property, with respect to the topology on X and \mathbb{R}). The type of continuity (e.g., strong, weak) depends on the context and topology of X .

Physical transformations, such as time evolution, spatial translations, or rotations, are typically understood to be generated by specific real physical quantities: the Hamiltonian generates time evolution, momentum generates spatial translations, angular momentum generates rotations, and so on. The restriction of generators to $\mathcal{O}_{\mathbb{R}}^\infty$ ensures that these are indeed real observables and that they are “sufficiently well-behaved” (due to being in \mathcal{O}^∞) for the generated transformations to be well-defined, for example, leading to differentiable paths for states or observables, or forming strongly continuous one-parameter groups as per Stone’s theorem, which we discuss in Lecture 4.

By Axiom 3, every $A \in \mathcal{O}^\infty$ has associated with it a one-parameter group of automorphisms of \mathcal{S} (evolution of states) and another such group for \mathcal{O} (evolution of observables). The first map, where a generator $A \in \mathcal{O}_{\mathbb{R}}^\infty$ gives rise to a one-parameter group of automorphisms $\{\alpha_t^A\}_{t \in \mathbb{R}}$ on \mathcal{S} , corresponds to the *Schrödinger picture*. In this picture, the observables are considered fixed (or only explicitly time-dependent), while the state of the system σ evolves according to $\sigma(t) = \alpha_t^A(\sigma(0))$. The second map, where A generates a group $\{\beta_t^A\}_{t \in \mathbb{R}}$ on \mathcal{O} , corresponds to the *Heisenberg picture*. Here, the state σ is considered fixed, while the observables A evolve according to $A(t) = \beta_t^A(A(0))$. For the framework to be consistent, these two pictures must yield equivalent physical predictions for the measurement outcomes $\sigma_A(t)$. This implies a compatibility condition between how states and observables transform, ensuring that the probability measure $(\sigma(t))_A$ (Schrödinger) is the same as $\sigma_{(A(t))}$ (Heisenberg).

Example 1.19. In quantum mechanics, time evolution is governed by the Hamiltonian operator $H \in \mathcal{O}_{\mathbb{R}}^\infty$. In the Schrödinger picture, state vectors evolve as $|\psi(t)\rangle = U_t |\psi(0)\rangle$, where $U_t = e^{-itH/\hbar}$ is a one-parameter unitary group. This induces automorphisms on the space of density operators (states \mathcal{S}) via $\rho(t) = U_t \rho(0) U_t^*$. Observables (operators) do not evolve in time unless explicitly specified. In the Heisenberg picture, observables evolve as $A(t) = U_t^* A(0) U_t$, while states do not evolve in time. Both $U_t(\cdot) U_t^*$ and $U_t^*(\cdot) U_t$ are one-parameter groups of automorphisms on the appropriate spaces.

Example 1.20. In classical mechanics, a Hamiltonian function $H \in C^\infty(M) \subset \mathcal{O}_\mathbb{R}^\infty$ generates a Hamiltonian flow ϕ_t on phase space M , which is a one-parameter group of symplectomorphisms, which we discuss in Lecture 2. This flow induces an evolution of states (probability measures on M) via the pushforward, $\mu_t = (\phi_t)_*\mu_0$, and an evolution of observables (functions on M) via the pullback, $f_t = f \circ \phi_{-t}$.

The usefulness of Axiom 3 lies in how it allows the theory to make predictions about how physical systems change over time. Beyond time evolution, it provides the mathematical machinery to describe continuous symmetries of the system. If a transformation leaves the system's properties invariant, its generator (an element of $\mathcal{O}_\mathbb{R}^\infty$) is often a conserved quantity associated with that symmetry. This is the concept behind Noether's theorem, linking symmetries to conservation laws, an important principle in physics.

1.1.5 Summary of DOSO

The following table summarizes the relationship between DOSO and three physical frameworks: classical mechanics, statistical mechanics, and quantum mechanics. The three physical theories are, of course, not mutually exclusive: statistical physics studies classical as well as quantum systems. We show how DOSO translates into the terms and conventions commonly used in classical mechanics, statistical mechanics, and quantum mechanics.

<i>DOSO Component</i>	<i>Classical Mechanics (CM)</i>	<i>Statistical Mechanics (SM)</i>	<i>Quantum Mechanics (QM)</i>
1. <i>Space of States (S)</i>	<i>Pure:</i> Phase space point. <i>Mixed:</i> Probability measure on phase space.	<i>State:</i> Probability measure on phase space. <i>Pure (SM):</i> Dirac delta measure (microstate). <i>Mixed:</i> Ensemble.	<i>State:</i> Density operator. <i>Pure:</i> Rank-one projector. <i>Mixed:</i> Non-projector density operator.
2. <i>Space of Operators (O) & Structures</i>	<i>Observables:</i> Real functions on phase space. <i>Smooth Subspace:</i> Smooth functions. <i>Lie Algebra:</i> Poisson bracket. <i>Functional Calculus:</i> Function composition.	<i>Observables:</i> Real functions on phase space. <i>Smooth Subspace:</i> Smooth functions. <i>Lie Algebra:</i> Poisson bracket (for Liouville eq.). <i>Functional Calculus:</i> Function composition (for averages).	<i>Observables:</i> Self-adjoint operators. <i>Smooth Subspace:</i> E.g., Bounded operators or specific domains. <i>Lie Algebra:</i> Commutator. <i>Functional Calculus:</i> Defined via the spectral theorem.
3. <i>Measurement Pairing</i>	Deterministic value for pure states.	Pushforward measure for ensembles.	Born Rule (probabilistic outcomes).
4. <i>Dynamics & Symmetries</i>	Hamilton's equations for phase space points. Poisson bracket with the Hamiltonian for observables.	Liouville's equation for probability density.	Schrödinger equation for states (density operators/wavefunctions). Heisenberg equation of motion for operators.

1.1.6 Compatibilities Among DOSO Data

The structures outlined in Axiom System 1.1 are not independent but must satisfy certain important compatibility conditions for the framework to be coherent and physically meaningful. The following are some examples of such requirements.

For $f \in \text{Borel}(\mathbb{R}; \mathbb{R})$, $A, A_1, A_2 \in \mathcal{O}$, $\sigma, \sigma_i \in \mathcal{S}$, and positive real numbers x^i (more accurately $p_i \in [0, 1]$) with $\sum p_i = 1$ (for a convex combination $p_1\sigma_1 + p_2\sigma_2 + \dots$), we require:

1. Linearity of measurement with respect to state mixing: If $\sigma = \sum_i p_i \sigma_i$ is a mixed state, then the probability distribution of A in state σ is the corresponding mixture of probability distributions:

$$\left(\sum_i p_i \sigma_i \right)_A = \sum_i p_i (\sigma_i)_A$$

This condition is important to the interpretation of mixed states as statistical ensembles. The probability distribution obtained from a mixed state is precisely the weighted average of the

probability distributions obtained from its underlying states. This applies to the probability measures themselves, implying linearity for expectation values as well.

2. Behavior of measurement for sums/functions of independent observables:

$$\sigma_{A_1+A_2} = \sigma_{A_1} * \sigma_{A_2} = \int_{-\infty}^{\infty} \sigma_{A_1}(y) \cdot \sigma_{A_2}(x-y) dy$$

This states that the probability distribution for the sum of two independent observables A_1, A_2 is the convolution of their individual probability distributions. This is a standard probability assumption that is true in classical, statistical, and quantum mechanics.

3. Compatibility of the Lie bracket with the real structure:

$$[A_1, A_2]^* = [A_1^*, A_2^*] \quad \text{for } A_1, A_2 \in \mathcal{O}^\infty$$

This condition ensures that the real structure (the $*$ -involution) is compatible with the Lie algebra structure on \mathcal{O}^∞ . It means that $*$ is an antilinear Lie algebra automorphism. An important consequence is that if A_1, A_2 are real observables (i.e., $A_1^* = A_1, A_2^* = A_2$), then $[A_1, A_2]^* = [A_1, A_2]$. This implies that $[A_1, A_2]$ is also a real observable, provided the Lie bracket itself maps pairs of real observables to real observables.

In classical mechanics, the compatibility of the Poisson bracket with the real structure ensures that the evolution of observables in phase space preserves their real values, while maintaining the consistency of symmetries and conserved quantities (e.g., energy, momentum) under canonical transformations. In statistical mechanics, the real structure guarantees that macroscopic thermodynamic quantities, such as the partition function and free energy, remain real and well-defined, while ensuring that the statistical evolution respects the symmetries of the system (e.g., time-reversal or gauge invariance). In quantum mechanics, this ensures that commutation relations between Hermitian operators preserve the Hermitian property, thus guaranteeing real eigenvalues for physical observables.

For $H \in \mathcal{O}_{\mathbb{R}}^\infty$ (a real, well-behaved observable, often the Hamiltonian or a symmetry generator), let

$$\begin{aligned} \sigma &\mapsto \sigma^{(t)} \\ A &\mapsto A^{(t)} \end{aligned}$$

denote the generated one-parameter groups of automorphisms of \mathcal{S} and \mathcal{O} , respectively. Then:

1. The Heisenberg equation of motion:

$$\frac{d}{dt} A^{(t)} + [H, A^{(t)}] = 0, \quad \text{for } A \in \mathcal{O}^\infty, t \in \mathbb{R}$$

This is the abstract form of the Heisenberg equation of motion. It describes how an observable A (from the “smooth” set \mathcal{O}^∞) evolves in time, $A \rightarrow A^{(t)}$, under the influence of the generator H . The Lie bracket $[H, A^{(t)}]$ dictates the rate and manner of this change.

The existence and uniqueness of solutions to this differential equation are very important, but we will not discuss it here.

2. Equivalence of Schrödinger and Heisenberg pictures for measurement outcomes:

$$(\sigma^{(t)})_A = \sigma_{A^{(t)}}, \quad \text{for } \sigma \in \mathcal{S}, A \in \mathcal{O}, t \in \mathbb{R}$$

This is an important consistency condition. It states that the probability distribution for measuring a fixed observable A in an evolved state $\sigma^{(t)}$ (Schrödinger picture) is identical to the probability distribution for measuring the evolved observable $A^{(t)}$ in the initial state σ (Heisenberg picture). This ensures that the physical predictions are independent of the “picture” we artificially adopted, establishing a duality between the two pictures.

1.1.7 Expectation Values and Separability

From the measurement pairing, we can define the *expectation value* (or average measured value) of an observable.

Definition 1.21 (Expectation Value). Given a state $\sigma \in \mathcal{S}$ and a real observable $A \in \mathcal{O}_{\mathbb{R}}$, the probability measure $\sigma_A \in \text{Prob}(\mathbb{R})$ describes the distribution of outcomes if A is measured in state σ . The *expectation value* of A in state σ , denoted $\langle A \rangle_{\sigma}$, is defined as the mean of this distribution:

$$\langle A \rangle_{\sigma} := \int_{\mathbb{R}} \lambda d\sigma_A(\lambda)$$

provided the integral exists. This defines a real-valued pairing $\mathcal{S} \times \mathcal{O}_{\mathbb{R}} \rightarrow \mathbb{R}$.

The expectation value $\langle A \rangle_{\sigma}$ represents the average result that would be obtained from a large number of independent measurements of the observable A on identical copies of the system, each prepared in the state σ . It is one of the most important quantities extracted from the probability distribution σ_A .

An important property of this pairing is that it “separates points”:

1. If $\sigma_1, \sigma_2 \in \mathcal{S}$ satisfy $\langle A \rangle_{\sigma_1} = \langle A \rangle_{\sigma_2}$ for all $A \in \mathcal{O}_{\mathbb{R}}$ for which the expectation values are defined, then $\sigma_1 = \sigma_2$.

This means that distinct states must be distinguishable by at least one observable’s expectation value. If two states yield the same average outcome for every possible measurement, they are the same state. This ensures that the space of states \mathcal{S} does not contain redundant information.

2. If $A_1, A_2 \in \mathcal{O}_{\mathbb{R}}$ satisfy $\langle A_1 \rangle_{\sigma} = \langle A_2 \rangle_{\sigma}$ for all $\sigma \in \mathcal{S}$ for which the expectation values are defined, then $A_1 = A_2$ (in $\mathcal{O}_{\mathbb{R}}$).

This is the same thing, but for observables. Distinct observables must be distinguishable by their expectation value in at least one state. If two observables yield the same average outcome in every possible state of the system, they are the same observable. This ensures that $\mathcal{O}_{\mathbb{R}}$ does not contain operationally redundant elements.

This separation property implies that the pairing $\langle \cdot, \cdot \rangle : \mathcal{S} \times \mathcal{O}_{\mathbb{R}} \rightarrow \mathbb{R}$ is non-degenerate.

1.1.8 Further Remarks on the Axiom System

We make several remarks about Axiom System 1.1.

Remark 1.22 (Interpretations and Implications).

1. *Schrödinger vs. Heisenberg Pictures:* The equation $\sigma^{(t)}_A = \sigma_{A(t)}$ is the mathematical expression of the physical equivalence between the Schrödinger picture (where states evolve in time and observables are fixed) and the Heisenberg picture (where states are fixed and observables evolve). The laws of physics should not depend on which calculational framework one chooses to describe the evolution of measurement outcome probabilities.
2. *Probabilistic Nature of Measurement:* Measurement is inherently probabilistic in quantum mechanics. In classical mechanics, if a system is in a pure state σ (e.g., a specific point in phase space detailing all positions and momenta), then the outcome σ_A of measuring an observable A (a function on phase space) is a point measure (Dirac delta measure at the value $A(\sigma)$). Thus, for pure classical states, measurement is deterministic. Probability enters classical mechanics only through mixed states (statistical ensembles, e.g., in statistical mechanics), where σ_A will generally be a measure with positive variance, reflecting our ignorance of the precise pure state.

In quantum mechanics, by contrast, even for a pure state (e.g., a state vector $|\psi\rangle$), the result of measuring most observables A is probabilistic, and σ_A will have positive variance (unless $|\psi\rangle$ is an eigenstate of A). This probabilistic nature is inherent to quantum mechanics.
3. *Construction of Mixed States from Pure States:* There is a map

$$\text{Prob}(\mathcal{PS}_0) \longrightarrow \mathcal{S}$$

(where \mathcal{PS}_0 is the set of pure states) that averages pure states over a probability measure to produce a general (mixed) state. If μ is a probability measure on \mathcal{PS}_0 , then a state $\sigma_\mu \in \mathcal{S}$ can be defined by $\sigma_\mu = \int_{\mathcal{PS}_0} \pi d\mu(\pi)$. This expresses the idea that any state can be seen as a statistical mixture of pure states. This is formally captured by Choquet theory for compact convex sets, where states are represented by barycentric integrals over extreme points.

4. *Symmetries of DOSO:* A symmetry of the “Data of States and Observables” (DOSO) is a collection of automorphisms of \mathcal{S} and of \mathcal{O} that preserve all the defining structures and compatibilities: the convex structure of \mathcal{S} , the algebraic and topological structure of \mathcal{O} , the real structure, the Lie bracket on \mathcal{O}^∞ , the functional calculus, and the measurement pairing. The group of all such symmetries is typically very large. Physical theories often study specific subgroups of this full symmetry group, which correspond to physical invariances (e.g., Poincaré symmetry in relativistic theories). Wigner’s theorem in quantum mechanics provides a useful characterization of symmetries.

1.1.9 Motion and Mechanical Systems

The general DOSO framework provides the static structure. Dynamics is introduced by selecting a specific notion of “motion.”

Definition 1.23 (Motion and Mechanical System). Suppose a DOSO (Axiom System 1.1) is given.

- A *motion* is the data of compatible one-parameter groups of automorphisms of \mathcal{S} and of \mathcal{O} . These are denoted

$$\begin{aligned}\sigma &\mapsto \sigma^{(t)} \\ A &\mapsto A^{(t)}\end{aligned}$$

for $t \in \mathbb{R}$.

- The compatibility condition is that these flows are adjoint with respect to the measurement pairing:

$$\sigma^{(t)}_A = \sigma_{A^{(t)}}$$

- A *mechanical system* is a DOSO together with a fixed, specified motion.

A “motion” singles out a particular way the system evolves, typically interpreted as time evolution. A mechanical system is thus not just the set of possible states and measurements, but also the rule governing how these change over time. We have ensured that this rule is physically consistent regardless of whether we view states or observables as evolving.

Remark 1.24 (Properties of Motion).

1. *Generation by a Hamiltonian:* In many important cases (Hamiltonian systems), the flows defining the motion are generated by a specific real observable $H \in \mathcal{O}_{\mathbb{R}}^{\infty}$, called the Hamiltonian of the system, via the mechanism described in Axiom 4. The Hamiltonian represents the total energy of the system in many physical contexts. If the flow is generated by H , then H itself is conserved under this flow (i.e., $H^{(t)} = H$, or $[H, H] = 0$).
2. *Time Translation Group:* The flows typically represent an action of the additive group \mathbb{R} of real numbers, interpreted as time translations. This is the idea that the laws of physics are invariant under shifts in time.
3. *Structure of Time:* Time is modeled as an affine line $\mathbb{A}^1(\mathbb{R})$, whose group of translations is \mathbb{R} . For simplicity, one often uses the standard real line \mathbb{R} . More sophisticated models might involve relativistic notions where time is part of a spacetime manifold, but for non-relativistic mechanics, this simple model of a uniform, universal time is usually sufficient. The affine structure means there’s a notion of duration, but no preferred origin. If a Euclidean structure is added, one has a metric for time intervals.
4. *Symmetries of a Mechanical System:* The symmetry group of a mechanical system is a subgroup of the full symmetry group of its underlying DOSO. In addition to preserving the static structures, these symmetries must also preserve the specific motion (dynamics). That is, if α is a symmetry transformation, it must commute with the time evolution: $\alpha(\sigma^{(t)}) = (\alpha(\sigma))^{(t)}$ and $\alpha(A^{(t)}) = (\alpha(A))^{(t)}$ (or, if the symmetry itself is time-dependent, a more general condition applies). For Hamiltonian systems, this often means the symmetry transformation must commute with the Hamiltonian, or leave the Hamiltonian invariant.

The possibility of time-reversal symmetries requires special treatment (e.g., they are often realized by anti-unitary operators in quantum mechanics), which we will not do now.

Often, a particular state $\sigma \in \mathcal{S}$ (e.g., a vacuum state or ground state) might be fixed as part of the data of a mechanical system, and symmetries might be required to preserve this special state.

Definition 1.25 (Stationary State). Fix a mechanical system with its defined motion $\sigma \mapsto \sigma^{(t)}$. A state $\sigma \in \mathcal{S}$ is called a *stationary state* (or an equilibrium state, or a time-invariant state) if it does not change under this motion, i.e.,

$$\sigma^{(t)} = \sigma \quad \text{for all } t \in \mathbb{R}.$$

Stationary states are very important. In quantum mechanics, eigenstates of the Hamiltonian are stationary states (their corresponding density operators are invariant under unitary evolution generated by the Hamiltonian). In statistical mechanics, equilibrium ensembles (microcanonical, canonical, grand canonical) are described by stationary states. Identifying and characterizing stationary states is often a primary goal when studying a mechanical system.

Remark 1.26. *The DOSO framework, along with its compatibility conditions and the definition of a mechanical system, provides a sophisticated and abstract foundation for a vast range of physical theories. Its power lies in its generality, allowing for a unified discussion of concepts in classical, statistical, and quantum mechanics.*

1.2 DOSO for Classical Mechanics

We now illustrate how to apply the DOSO formulation to classical, non-relativistic mechanics. This serves not only to ground the axiomatic framework in familiar territory but also to highlight the unifying power of the DOSO formulation. The chosen example, a single particle in Euclidean space, while elementary, is sufficiently rich to exhibit the important features.

1.2.1 A Particle in Euclidean Space

For concreteness, consider a single particle of constant mass $m \in \mathbb{R}^{>0}$ moving in a standard d -dimensional Euclidean space \mathbb{E}^d , where $d \in \mathbb{Z}^{>0}$. The particle's motion is impacted by a potential energy function $V : \mathbb{E}^d \rightarrow \mathbb{R}$. For the purposes of defining smooth dynamics, we typically assume V to be sufficiently regular, e.g., $V \in C^k(\mathbb{E}^d, \mathbb{R})$ for $k \geq 2$, often $V \in C^\infty(\mathbb{E}^d, \mathbb{R})$. For convenience, we denote the configuration space $M := \mathbb{E}^d$.

A *motion* or *trajectory* of the particle is a sufficiently regular map $x : \mathbb{R} \rightarrow M$, $t \mapsto x(t)$, where t represents time. The space of all such conceivable motions (prior to imposing physical laws) can be taken as $\mathcal{F} := C^k(\mathbb{R}, M)$ for appropriate k ; if we desire maximal regularity, we consider $\mathcal{F} = C^\infty(\mathbb{R}, M)$. This \mathcal{F} can be endowed with the structure of an infinite-dimensional Fréchet manifold, but we will not say more about that here.

Physical, or *classical*, trajectories are those that satisfy Newton's second law:

$$m\ddot{x}(t) = -\nabla V(x(t))$$

where $\ddot{x}(t) = \frac{d^2x}{dt^2}(t)$ is the acceleration and $\nabla V(x(t))$ is the gradient of V evaluated at the particle's position $x(t)$. Newton's second law is a system of d second-order ordinary differential equations. An important assumption for simplifying the structure of the space of solutions is that for any given initial conditions $(x_0, v_0) \in M \times \mathbb{R}^d$ (initial position and initial velocity), there exists a unique solution $x(t)$ to Newton's second law for all $t \in \mathbb{R}$. This global existence and uniqueness is a strong condition, not always satisfied (e.g., for potentials that lead to finite-time singularities). However, when it holds, the set $N \subset \mathcal{F}$ of all physical trajectories forms a $2d$ -dimensional differentiable manifold.

This manifold N can be explicitly parameterized. For any chosen reference time $t_0 \in \mathbb{R}$, the map

$$\begin{aligned} \Psi_{t_0} : N &\longrightarrow M \times \mathbb{R}^d \cong \mathbb{E}^d \times \mathbb{R}^d \\ x(\cdot) &\longmapsto (x(t_0), \dot{x}(t_0)) \end{aligned}$$

is a diffeomorphism. This map, which we may call the *initial data map* at time t_0 , provides a global chart for N . It's important to note that while N itself is defined as the solution space, the specific parameterization explicitly breaks the inherent time-translation symmetry of Newton's equations (if V is time-independent). The space $\mathbb{E}^d \times \mathbb{R}^d$ can be identified with the tangent bundle TM of $M = \mathbb{E}^d$.

1.2.2 DOSO Data for the Classical Particle

We now proceed to write the components of DOSO for this classical system.

In this deterministic classical framework, a state of maximal information corresponds to knowing the entire trajectory of the particle. Thus, the space of pure states is precisely the manifold of classical trajectories:

$$\mathcal{PS} := N$$

Each point $x(\cdot) \in N$ represents a unique, complete history of the particle satisfying the laws of motion.

For classical systems, mixed states arise from statistical uncertainty about the pure state. The encompassing space of states \mathcal{S} is the space of all Borel probability measures on N :

$$\mathcal{S} := \text{Prob}(N)$$

This space \mathcal{S} is a convex set. A pure state $x_N \in N$ can be identified with the Dirac delta measure $\delta_{x_N} \in \mathcal{S}$. These are the extreme points of \mathcal{S} . A general mixed state $\sigma \in \mathcal{S} \setminus \mathcal{PS}$ describes a statistical ensemble where the particle follows one of the classical trajectories in N according to the probability distribution σ . This is the domain of classical statistical mechanics.

Recall that an observable is a physical quantity whose value can be determined if the state of the system is known. In the pure state picture, an observable is a function on N .

- *The Algebra \mathcal{O} and its "Smooth" Subspace \mathcal{O}^∞ :* We define the space of complex-valued observables as

$$\mathcal{O} := \text{Borel}(N; \mathbb{C})$$

the space of complex-valued Borel-measurable functions on N . This is a complex vector space. The subspace of "smooth" observables is

$$\mathcal{O}^\infty := C^\infty(N; \mathbb{C})$$

the space of complex-valued smooth functions on N . Given that N is a $2d$ -dimensional smooth manifold, $C^\infty(N; \mathbb{C})$ is well-defined and forms a dense subspace of \mathcal{O} under suitable topologies (e.g., L^p -like topologies if N carries a natural measure, or in the sense of approximating continuous functions). Endowing \mathcal{F} (the space of all paths) with an infinite-dimensional manifold structure is a more complicated topic that we will not do. For defining observables that depend only on the physical trajectory (elements of N), restricting to functions on N is sufficient and avoids these complexities.

- *Real Structure:* The real structure $A \mapsto A^*$ on \mathcal{O} is given by pointwise complex conjugation: if $A \in \mathcal{O}$, then $A^*(x_N) := \overline{A(x_N)}$ for all $x_N \in N$. Consequently, the space of real observables $\mathcal{O}_{\mathbb{R}}$ (the fixed points of $*$) is

$$\mathcal{O}_{\mathbb{R}} = \text{Borel}(N; \mathbb{R})$$

These are the functions that correspond to physically measurable quantities, yielding real numbers. Similarly, $\mathcal{O}_{\mathbb{R}}^\infty = C^\infty(N; \mathbb{R})$.

- *Lie Algebra Structure:* The space $\mathcal{O}_{\mathbb{R}}^{\infty} = C^{\infty}(N; \mathbb{R})$ is endowed with a real Lie algebra structure via the *Poisson bracket*, which requires the symplectic structure on N discussed below. For $A, B \in C^{\infty}(N; \mathbb{R})$, their Poisson bracket $\{A, B\} \in C^{\infty}(N; \mathbb{R})$. The complexification $\mathcal{O}^{\infty} = C^{\infty}(N; \mathbb{C})$ then becomes a complex Lie algebra. The compatibility $[A_1, A_2]^* = [A_1^*, A_2^*]$ is satisfied by the Poisson bracket since $\{A, B\}^* = \overline{\{A, B\}}$, and if $A^* = A, B^* = B$, then $\{A, B\}$ is real, so $\{\bar{A}, \bar{B}\} = \overline{\{A, B\}}$.

The measurement pairing connects observables to probability distributions of outcomes for a given state. For $\sigma \in \mathcal{S} = \text{Prob}(N)$ and $A \in \mathcal{O}_{\mathbb{R}} = \text{Borel}(N; \mathbb{R})$, the probability measure $\sigma_A \in \text{Prob}(\mathbb{R})$ is given by the *pushforward measure*:

$$\sigma_A := A_*\sigma$$

That is, for any Borel set $E \subset \mathbb{R}$, $(A_*\sigma)(E) := \sigma(A^{-1}(E))$.

- *Deterministic Measurement for Pure States:* If $\sigma = \delta_{x_N}$ is a pure state (corresponding to a specific trajectory $x_N \in N$), then $\sigma_A = (A_*)(\delta_{x_N}) = \delta_{A(x_N)}$. This is a point measure concentrated at the value $A(x_N)$. Thus, for pure states in classical mechanics, the measurement of any observable yields a definite, deterministic outcome. This is a key distinction from quantum mechanics.
- *Functional Calculus:* The functional calculus $f, A \mapsto f(A)$ is simply function composition: if $A \in \mathcal{O}_{\mathbb{R}}$ and $f \in \text{Borel}(\mathbb{R}; \mathbb{R})$, then $f(A)$ is the function $x_N \mapsto f(A(x_N))$, which is also in $\mathcal{O}_{\mathbb{R}}$. The compatibility $\sigma_{f(A)} = f_*\sigma_A$ is a standard property of pushforward measures: $(f \circ A)_*\sigma = f_*(A_*\sigma)$.

The motion described in DOSO is induced by a flow on the manifold of pure states N . This flow, in turn, is governed by Hamiltonian mechanics, which relies on a symplectic structure on N . The time evolution $x_N \mapsto (x_N)^{(t)}$ on N (where $(x_N)^{(t)}(s) = x_N(s + t)$ by time translation along the trajectory) defines a one-parameter group of automorphisms of N . This induces:

- Automorphisms on $\mathcal{S} = \text{Prob}(N)$: $\sigma \mapsto \sigma^{(t)}$, where $\sigma^{(t)}(S) = \sigma(\{x_N \mid (x_N)^{(t)} \in S\})$.
- Automorphisms on $\mathcal{O} = \text{Borel}(N; \mathbb{C})$: $A \mapsto A^{(t)}$, where $A^{(t)}(x_N) = A((x_N)^{(-t)})$.

This motion is generated by a Hamiltonian function $H \in C^{\infty}(N; \mathbb{R})$ associated with the system's energy. The Hamiltonian flow on N is a one-parameter group of symplectomorphisms. The symplectic structure itself is therefore important. We will discuss this in more detail in Lecture 2.

Without going into too much detail, a pair (N, H) , where N is a symplectic manifold (the space of pure states) and $H \in C^{\infty}(N; \mathbb{R})$ is a Hamiltonian function, determines a complete classical mechanical system in the DOSO framework. For our particle example, the symplectic structure on N can be derived from a more foundational *variational principle* (the principle of least action) applied to a Lagrangian function. This is the *Lagrangian formulation*. The transition to (N, H) is the *Hamiltonian formulation*. The derivation of the symplectic structure within the Lagrangian framework is detailed in Lecture 3. Here, we shall merely state the symplectic 2-form on N .

Remark 1.27 (General Symplectic Flows). *More generally, a classical mechanical system can be defined by a pair $(N, \{\varphi_t\}_{t \in \mathbb{R}})$, where N is a symplectic manifold and $\{\varphi_t\}$ is a one-parameter group of symplectomorphisms (diffeomorphisms preserving the symplectic form ω). Such a system always fits the DOSO framework. However, not every such flow $\{\varphi_t\}$ is Hamiltonian in the sense of being generated by a time-independent Hamiltonian function $H \in C^{\infty}(N; \mathbb{R})$ (i.e., φ_t being the flow of the Hamiltonian vector field X_H). For instance, systems with explicitly time-dependent Hamiltonians in*

the traditional sense give rise to symplectic flows that are not necessarily Hamiltonian on N (which is the space of solutions of an autonomous system). In Lecture 3 we will see a non-Hamiltonian symplectic flow.

1.2.3 The Symplectic Form for the Particle Example

The symplectic structure on N (the space of physical trajectories) is naturally inherited from a canonical structure on the cotangent bundle T^*M of the configuration space $M = \mathbb{E}^d$.

Let $\mathcal{F} = C^\infty(\mathbb{R}, M)$ be the ambient space of smooth paths. One can define a presymplectic 2-form⁶ on \mathcal{F} . Let δ denote the exterior derivative (variational derivative) on this infinite-dimensional manifold \mathcal{F} . The evaluation map is

$$\begin{aligned} e : \mathcal{F} \times \mathbb{R} &\longrightarrow M \\ (x(\cdot), t) &\longmapsto x(t) \end{aligned}$$

Consider variations δe (variation of position) and $\delta \dot{e}$ (variation of velocity). A 2-form on $\mathcal{F} \times \mathbb{R}$ (or, by fixing time, on \mathcal{F}) can be defined as

$$\Omega = m \int \langle \delta \dot{x}(t) \wedge \delta x(t) \rangle_{\mathbb{R}^d} dt$$

A more direct approach for the manifold of solutions N is to use the parameterization $\Psi_{t_0} : N \xrightarrow{\cong} TM \cong \mathbb{E}^d \times \mathbb{R}^d$ given by $(x(t_0), \dot{x}(t_0))$. Let (q^1, \dots, q^d) be standard coordinates on \mathbb{E}^d (representing $x(t_0)$) and (v^1, \dots, v^d) be coordinates for the velocities $\dot{x}(t_0)$ in \mathbb{R}^d . The pullback $(\Psi_{t_0})^* \omega_{TM}$ of the canonical symplectic form ω_{TM} on TM (derived from the standard Lagrangian $L(q, v) = \frac{1}{2}m\|v\|^2 - V(q)$) to N is given by:

$$\omega_N|_{t_0} = \sum_{i=1}^d m dv^i \wedge dq^i$$

This form is understood as follows: dq^i and dv^i are basis 1-forms on TM at (q, v) . The important insight is that this form on N is independent of the choice of t_0 . This is a consequence of Liouville's theorem for Hamiltonian flows (which implies that the Hamiltonian flow preserves the symplectic form). Thus, N carries a canonical symplectic structure.

Remark 1.28 (Symplectic Form on Phase Space T^*M). *It is more conventional in Hamiltonian mechanics to work with phase space T^*M , with canonical coordinates (q^i, p_i) , where $p_i = \frac{\partial L}{\partial \dot{q}^i}$ are the canonical momenta. For $L = \frac{1}{2}m\sum(\dot{q}^i)^2 - V(q)$, we have $p_i = m\dot{q}^i = mv^i$. In these (q, p) coordinates, the canonical symplectic 2-form on T^*M is*

$$\omega_{T^*M} = \sum_{i=1}^d dp_i \wedge dq^i$$

*The form from earlier becomes $m\sum dv^i \wedge dq^i$. If we set $p_i = mv^i$, then $dp_i = m dv^i$, so the forms are identical under this identification. Thus, the space of solutions N is naturally identified with T^*M (or TM), equipped with its canonical symplectic structure.*

⁶A presymplectic 2-form is a closed differential 2-form of constant rank on a manifold

1.2.4 Dynamical Systems: Local versus Global Perspectives

Remark 1.29 (Classical Dynamics: Short Times vs. Long Times, Particles vs. Fields). A symplectic manifold (N, ω) together with a smooth function $H : N \rightarrow \mathbb{R}$ (the Hamiltonian) constitutes a Hamiltonian dynamical system. The dynamics are governed by Hamilton's equations, which can be written as $\dot{z}(t) = X_H(z(t))$, where $z(t)$ is a curve in N and X_H is the Hamiltonian vector field associated with H , uniquely defined by the relation $i_{X_H}\omega = -dH$.

Local Existence and Uniqueness: The theory of ordinary differential equations (e.g., the Picard-Lindelöf theorem, applicable since X_H is smooth if H is C^2) guarantees the local existence and uniqueness of a flow $\varphi_t : N \rightarrow N$ that integrates X_H . This addresses the behavior of the system for small times t . Constructing this local flow is relatively straightforward from a mathematical standpoint, often relying on fixed-point theorems.

Global Dynamical Questions: The more interesting and challenging questions in dynamical systems theory concern the long-time behavior of these flows. These include:

- *Existence of Periodic Orbits:* Do trajectories exist that return to their initial state after a finite time? These are important to understanding stability and resonances.
- *Stability of Orbits:* How do trajectories behave under small perturbations of initial conditions (Lyapunov stability)?
- *Ergodicity and Mixing:* Does the system explore the available phase space uniformly over long times?
- *Chaotic Dynamics:* Does the system exhibit sensitive dependence on initial conditions, leading to complex, unpredictable long-term behavior despite its deterministic nature? Hamiltonian systems, despite preserving phase space volume (Liouville's theorem, a consequence of $\mathcal{L}_{X_H}\omega = 0$), can indeed be chaotic (e.g., KAM theory, Smale horseshoes).
- *Limit Cycles:* These are isolated periodic orbits towards which nearby trajectories converge. While prominent in dissipative systems, classical Hamiltonian systems (which are volume-preserving) do not possess attracting limit cycles in the usual sense. However, more complex recurrent structures exist.

Observe that the foundational local existence and uniqueness concerns operate at small time scales, whereas these deeper dynamical questions probe the system's behavior as $t \rightarrow \infty$.

Analogy with Classical Field Theory: Classical field theory (describing systems with infinitely many degrees of freedom, like electromagnetism or fluid dynamics) presents something similar, now involving both time and space.

- *Local/Short-Range Questions:* Here, "local" means small regions in spacetime (small time differences and small spatial distances). The dynamics are typically governed by partial differential equations (PDEs), such as wave equations, Maxwell's equations, or equations from fluid dynamics (e.g., Euler or Navier-Stokes). The interesting mathematical questions about local behavior are all similar: given suitable initial data, does a unique solution exist for a short time, and does it depend continuously on the initial data? This is already a highly non-trivial area of PDE theory.
- *Global/Long-Range Questions:* These pertain to the behavior of solutions for all times and over large spatial domains. Examples include: Does a solution exist globally in time, or do singularities (e.g., wave collapse, shock formation in fluids, gravitational singularities) develop

in finite time? What is the asymptotic behavior of fields as $t \rightarrow \infty$? How do spatially extended structures form and evolve?

Analogy with Quantum Field Theory (QFT): The questions of QFT can often be categorized into kinematics and dynamics.

- *Kinematics: This involves defining the important objects: quantum fields, state spaces, and observables. A key challenge at short distances (high energies/momenta) is the appearance of infinities (UV divergences) in perturbative calculations, necessitating the very complicated machinery of renormalization. The behavior of correlation functions at short distances (operator product expansions) is important.*
- *Dynamics: Much of this concerns the behavior of the theory at large distances (low energies/momenta) and long times. Key questions include: The structure of the vacuum state, the particle spectrum (existence of a mass gap), scattering theory (S-matrix), phase transitions and critical phenomena, and long-range forces and phenomena like confinement in quantum chromodynamics.*

In relativistic QFTs, due to Lorentz invariance, long-time and large-distance behaviors are related. The understanding of these long-range phenomena often leads to the most interesting physical insights and drives the development of powerful non-perturbative mathematical techniques.

The long-range questions in classical mechanics, classical field theory, and quantum field theory are often the most intriguing, with significant mathematical implications. These questions frequently lead to unexpected connections between physics and seemingly unrelated areas of mathematics.

1.3 DOSO for Quantum Mechanics

We now transition our focus to quantum mechanics, illustrating how its foundational structure can be also be encapsulated via DOSO.

A notable simplification adopted in this initial presentation is the restriction to *bounded observables*. This is a common starting point in many rigorous treatments as it allows the full power of C^* -algebra theory to be deployed. However, it is important to recognize that many of the most important physical observables in quantum mechanics, such as position, momentum, and typically the Hamiltonian itself (which generates time evolution), are represented by *unbounded* self-adjoint operators. The theory of unbounded operators is significantly difficult. Nevertheless, the framework for bounded observables provides a good foundation and its principles often extend to the unbounded case.

1.3.1 Hilbert Spaces and Planck's Constant

At the center of quantum theory lies a universal constant of nature, *Planck's constant* \hbar . This constant has the physical dimensions of *action*:

$$[\hbar] = \frac{\text{mass} \cdot \text{length}^2}{\text{time}}$$

Recalling that energy has units of $\text{mass} \cdot \text{length}^2 / \text{time}^2$, the units of \hbar are equivalently those of $\text{energy} \cdot \text{time}$. Planck's constant quantifies the important scale at which quantum effects become manifest, appearing often in commutation relations (e.g., $[x, p] = i\hbar$), uncertainty principles (e.g.,

$\Delta x \Delta p \geq \hbar/2$), and the quantization of physical quantities (e.g., energy levels of bound systems, angular momentum).

The entire DOSO structure of a quantum mechanical system is derived from a single, foundational mathematical object: The Hilbert space. To every quantum mechanical system is associated a *separable complex Hilbert space* \mathcal{H} , which may be finite or countably infinite dimensional. This space \mathcal{H} is the place where the states and observables of the system are defined.

Complex numbers are essential in quantum mechanics for describing superposition, interference, and time evolution via the Schrödinger equation. Probability amplitudes, central to the Born rule, rely on the complex structure of Hilbert space. The Hilbert space's inner product defines measurement probabilities, orthogonality, and self-adjoint operators (observables), while completeness ensures convergence and enables the spectral theorem. Foundational quantum systems typically use separable Hilbert spaces with countable orthonormal bases, though non-separable spaces appear in quantum field theory.

1.3.2 DOSO Data for Quantum Systems

With the Hilbert space \mathcal{H} established, we can now define the components of the DOSO framework for a quantum system.

In quantum mechanics, a pure state corresponds to a *ray* in the Hilbert space \mathcal{H} .

Definition 1.30. A *ray* is an equivalence class of non-zero vectors differing by a non-zero complex scalar multiple: $\psi = \{c\psi \mid c \in \mathbb{C}, c \neq 0\}$, where $\psi \in \mathcal{H} \setminus \{0\}$.

Conventionally, rays are represented by unit vectors $\|\psi\| = 1$, with the understanding that vectors differing only by a phase factor $e^{i\alpha}$ (an unobservable global phase) represent the same physical state. The space of all such rays is the *projective Hilbert space* $P(\mathcal{H})$.

$$\mathcal{PS} := P(\mathcal{H})$$

The general space of states \mathcal{S} , encompassing both pure and mixed states, is the set of *density operators* on \mathcal{H} .

Definition 1.31. A *density operator* ρ is a linear operator on \mathcal{H} satisfying:

1. ρ is self-adjoint: $\rho = \rho^*$.
2. ρ is non-negative (positive semi-definite): $\langle \phi, \rho \phi \rangle \geq 0$ for all $\phi \in \mathcal{H}$.
3. ρ is trace-class, and its trace is unity: $\text{Tr}(\rho) = 1$.

$$\mathcal{S} := \{\rho \in \mathcal{T}_1(\mathcal{H}) \mid \rho = \rho^*, \rho \geq 0, \text{Tr}(\rho) = 1\}$$

where $\mathcal{T}_1(\mathcal{H})$ is the space of trace-class operators on \mathcal{H} .

The set \mathcal{S} is a convex subset of the real Banach space of self-adjoint trace-class operators. A density operator ρ represents a pure state if and only if $\rho^2 = \rho$ (in addition to the other defining properties), or equivalently, if $\text{Tr}(\rho^2) = 1$. Otherwise, if $\text{Tr}(\rho^2) < 1$, ρ represents a mixed state, which can be expressed as a convex combination of pure states (e.g., $\rho = \sum_k p_k |\psi_k\rangle \langle \psi_k|$ with $p_k > 0, \sum p_k = 1$).

- *The Algebra \mathcal{O} and \mathcal{O}^∞ : Bounded Operators on \mathcal{H}*

In this simplified exposition focusing on bounded observables, we define:

$$\mathcal{O} = \mathcal{O}^\infty := \text{End}(\mathcal{H}) \equiv \mathcal{B}(\mathcal{H})$$

the complex Banach space (in fact, a C^* -algebra) of all bounded linear operators mapping \mathcal{H} to itself. The restriction to bounded operators ensures that they are defined on the entirety of \mathcal{H} and possess desirable analytic properties.

- *Real Structure: Operator Adjoint*

The real structure $A \mapsto A^*$ on $\mathcal{O} = \mathcal{B}(\mathcal{H})$ is given by the standard Hilbert space *adjoint operation*. An operator A^* is the adjoint of A if $\langle \phi, A\psi \rangle = \langle A^*\phi, \psi \rangle$ for all $\phi, \psi \in \mathcal{H}$. This map is antilinear and involutive.

- *Real Observables $\mathcal{O}_\mathbb{R}$: Self-Adjoint Operators*

Consequently, the space of real observables $\mathcal{O}_\mathbb{R}$ consists of all bounded *self-adjoint operators* on \mathcal{H} :

$$\mathcal{O}_\mathbb{R} = \{A \in \mathcal{B}(\mathcal{H}) \mid A = A^*\}$$

These operators have real spectrum and correspond to physically measurable quantities whose outcomes are real numbers.

- *Lie Algebra Structure: The Quantum Commutator*

The complex Lie algebra structure on $\mathcal{O} = \mathcal{B}(\mathcal{H})$ is defined via a scaled commutator:

$$[A_1, A_2]_\hbar := \frac{-i}{\hbar}(A_1 A_2 - A_2 A_1)$$

for $A_1, A_2 \in \mathcal{B}(\mathcal{H})$. Here, $A_1 A_2$ denotes operator composition.

If A_1 and A_2 are self-adjoint, their commutator $C = A_1 A_2 - A_2 A_1$ is skew-adjoint ($C^* = -C$). Multiplication by $-i/\hbar$ (or $1/(i\hbar)$) then makes $[A_1, A_2]_\hbar$ self-adjoint. Thus, this definition ensures that the Lie bracket of two real (self-adjoint) observables is itself a real (self-adjoint) observable, making $\mathcal{O}_\mathbb{R}$ a real Lie algebra under this bracket. This Lie algebra structure is the quantum analogue of the Poisson bracket algebra in classical mechanics and is important to the Heisenberg uncertainty principle and to the equations of motion. Thus, the compatibility condition $[X, Y]^* = [X^*, Y^*]$ is satisfied by this definition.

- *Functional Calculus for Observables*

Functions of observables are defined via the *spectral theorem* for self-adjoint operators, which we discuss in Lecture 4. If $A \in \mathcal{O}_\mathbb{R}$ is a bounded self-adjoint operator and $f : \mathbb{R} \rightarrow \mathbb{R}$ is a Borel function, then $f(A)$ is a well-defined bounded self-adjoint operator. This is important for constructing observables like A^2 or $\exp(A)$, and particularly for defining projection-valued measures $\Pi_A(E) := \chi_E(A)$ (where χ_E is the characteristic function of a Borel set $E \subset \mathbb{R}$), which are central to the theory of measurement.

The measurement pairing connects an observable $A \in \mathcal{O}_\mathbb{R}$ and a state (pure or mixed) to a probability distribution $\sigma_A \in \text{Prob}(\mathbb{R})$ for the outcomes of measuring A . Let Π_A denote the projection-valued measure associated with the self-adjoint operator A via the spectral theorem (so $\Pi_A(E)$ is the projection onto the subspace of states for which the value of A lies in the Borel set $E \subset \mathbb{R}$).

- *Measurement in a Pure State:* Suppose the system is in a pure state represented by a unit vector $\psi \in \mathcal{H}$ (equivalently, by the line $L = \text{span}\{\psi\} \in P(\mathcal{H})$ or the density operator $\rho_\psi = |\psi\rangle\langle\psi|$).

The probability measure $\sigma_A \equiv (\rho_\psi)_A$ on \mathbb{R} is given by the *Born rule*:

$$(\rho_\psi)_A(E) := \langle \psi, \Pi_A(E) \psi \rangle = \text{Tr}(\rho_\psi \Pi_A(E))$$

for any Borel subset $E \subset \mathbb{R}$. This value is the probability that a measurement of A yields a result in E . The independence of this probability from the choice of unit vector representing the ray L is ensured by the quadratic dependence on ψ (an overall phase $e^{i\alpha}$ cancels out: $\langle e^{i\alpha}\psi, \Pi_A(E) e^{i\alpha}\psi \rangle = \langle \psi, \Pi_A(E) \psi \rangle$).

- *Measurement in a Mixed State*: If the system is in a general (possibly mixed) state described by a density operator $S \in \mathcal{S}$, then we can generalize to:

$$S_A(E) := \text{Tr}(S \Pi_A(E))$$

for any Borel subset $E \subset \mathbb{R}$. This formula is linear in S , reflecting the convex nature of states, and provides the statistical predictions of quantum theory for any preparation S and any (bounded) observable A . If $S = \sum_k p_k |\psi_k\rangle \langle \psi_k|$, then $S_A(E) = \sum_k p_k \langle \psi_k, \Pi_A(E) \psi_k \rangle$.

Quantum Motion (Dynamics) The motion (time evolution) of a quantum system is generated by a specific self-adjoint observable $H \in \mathcal{O}_{\mathbb{R}}^\infty$ (here, $\mathcal{O}_{\mathbb{R}}^\infty = \mathcal{O}_{\mathbb{R}}$ are bounded self-adjoint operators), designated as the *Hamiltonian* of the system. In most realistic systems, H is unbounded, but the principles extend via Stone's theorem.

- *Unitary Time Evolution Group*: The Hamiltonian H generates a one-parameter group of unitary transformations $\{U_t\}_{t \in \mathbb{R}}$ on \mathcal{H} :

$$U_t := e^{-itH/\hbar}, \quad t \in \mathbb{R}$$

This is defined using the functional calculus for self-adjoint operators (applying $f_t(\lambda) = e^{-it\lambda/\hbar}$ to H). Since H is self-adjoint, U_t is unitary for all $t \in \mathbb{R}$ ($U_t^* = U_t^{-1} = U_{-t}$). These operators describe the evolution of the system over a time interval t . For bounded H , this group is norm-continuous.

- *Evolution of States (Schrödinger Picture)*: If a state at time $t = 0$ is $S \equiv S(0)$, its evolution to time t is given by:

$$S^{(t)} \equiv S(t) := U_t S(0) U_t^*$$

This preserves the properties of a density operator (self-adjointness, non-negativity, trace one). If $S(0) = |\psi_0\rangle \langle \psi_0|$ is a pure state, then $S(t) = |\psi(t)\rangle \langle \psi(t)|$ where $|\psi(t)\rangle = U_t |\psi_0\rangle$. The latter satisfies the *Schrödinger equation*: $i\hbar \frac{d}{dt} |\psi(t)\rangle = H |\psi(t)\rangle$. More generally, density operators evolve according to the *Liouville-von Neumann equation*: $i\hbar \frac{d}{dt} S(t) = [H, S(t)] \equiv HS(t) - S(t)H$.

- *Evolution of Observables (Heisenberg Picture)*: Alternatively, one can consider the states fixed and the observables evolving. An observable $A \equiv A(0)$ evolves to $A^{(t)} \equiv A(t)$ according to:

$$A^{(t)} \equiv A(t) := U_t^* A(0) U_t$$

These evolved observables $A(t)$ satisfy the *Heisenberg equation of motion*:

$$\frac{d}{dt} A(t) = \frac{i}{\hbar} [H, A(t)]$$

The compatibility condition $(S^{(t)})_A = S_{(A^{(t)})}$ is satisfied:

$$\mathrm{Tr}(S^{(t)}\Pi_A(E)) = \mathrm{Tr}(U_t S U_t^* \Pi_A(E)) = \mathrm{Tr}(S U_t^* \Pi_A(E) U_t).$$

The term $U_t^* \Pi_A(E) U_t$ is the PVM for the evolved observable $A^{(t)}$, evaluated on E . That is,

$$\Pi_{A^{(t)}}(E) = U_t^* \Pi_A(E) U_t.$$

Thus

$$\mathrm{Tr}(S \Pi_{A^{(t)}}(E)) = S_{(A^{(t)})}(E).$$

We have finished explaining how to axiomatize quantum mechanics using DOSO, and this is a good place to stop.

2 Riemannian and Symplectic Manifolds

Abstract

We develop the basics on Riemannian and symplectic structures. We begin with linear algebraic preliminaries, analyzing bilinear forms on finite-dimensional real vector spaces. Symmetric positive-definite forms yield Euclidean structures with notions of length and angle, while skew-symmetric non-degenerate forms define symplectic structures with oriented area. Their properties are treated rigorously. The spaces of such forms, $Met(\mathbb{R}^n)$ and $Symp(\mathbb{R}^{2m})$, are identified as homogeneous spaces $GL(n, \mathbb{R})/O(n)$ and $GL(2m, \mathbb{R})/Sp(2m, \mathbb{R})$, whose topological distinctions reflect differences.

We then transition to affine geometry, defining affine spaces as vector spaces without a distinguished origin. Affine maps, groups, and coordinates are introduced, leading to Euclidean affine spaces (with constant inner product structures) and affine symplectic spaces (with closed symplectic 2-forms). We compare their induced geometries, and define vector fields, flows, and gradients (Euclidean and Hamiltonian) in this flat setting.

Globalization to smooth manifolds follows, with Riemannian metrics and almost symplectic forms constructed as smooth tensor fields. A key distinction emerges: Riemannian metrics exist on all smooth manifolds, whereas symplectic structures face topological obstructions. Symplectic geometry is defined by the integrability condition $d\omega = 0$, leading to Darboux's theorem and the focus on global invariants such as de Rham cohomology. Cotangent bundles T^*M serve as canonical symplectic examples. Hamiltonian vector fields, their flows, the Poisson bracket on $C^\infty(M)$, and the role of $H_{dR}^1(M)$ in obstructing Hamiltonianity are examined.

The section concludes with the theory of symmetries and moment maps. We define symplectic and Hamiltonian vector fields, relate them to conserved quantities via Noether's theorem, and introduce moment maps $\mu : N \rightarrow \mathfrak{g}^*$ for Lie algebra actions. Their existence and equivariance are tied to cohomological obstructions ($H^1(N; \mathbb{R})$, $H^2(\mathfrak{g}; H^0(N; \mathbb{R}))$). The moment map equation $d\mu_\xi = -\iota_{\alpha(\xi)}\omega$ and G -equivariance are explored through examples. A geometric interpretation via prequantization bundles connects the moment map to geometric quantization, connecting it to classical and quantum mechanics.

Contents

2.1	Bilinear Forms in Linear Algebra	29
2.1.1	Bilinear Forms: Generalities	29
2.1.2	Symplectic Vector Spaces	30
2.1.3	Model Forms	32
2.1.4	The Spaces $Met(\mathbb{R}^n)$ and $Symp(\mathbb{R}^n)$	34
2.2	Euclidean Spaces and Affine Symplectic Spaces	37
2.2.1	Linear Algebra, But For Affine Spaces	37
2.2.2	Affine Maps and the Affine Group	40
2.2.3	Euclidean Affine Spaces	43
2.2.4	Affine Symplectic Spaces	45
2.2.5	Vector Fields, Flows, and Gradients on Affine Spaces	47
2.3	Riemannian and Symplectic Manifolds	52
2.3.1	Pointwise Structures, Fiber Bundles, and Global Tensor Fields	53
2.3.2	Existence of Structures	54
2.3.3	The Symplectic Condition: Integrability	56

2.3.4	Canonical Example: The Cotangent Bundle T^*M	58
2.3.5	Dynamics on Symplectic Manifolds: Hamiltonian Formalism	59
2.4	Symmetries and Moment Maps	62
2.4.1	Finite Symmetries: Symplectomorphism Groups	62
2.4.2	More on Symplectic Vector Fields	63
2.4.3	Lie Algebra Actions and the Moment Map Problem	64
2.4.4	Properties and Characterization of Moment Maps	66
2.4.5	Geometric Interpretation: Prequantization Bundles	67

References: We refer the reader for foundational and comprehensive treatments of Riemannian geometry, including metrics, curvature, and manifold structures, to [19] [8]; for the principles of affine geometry and its interplay with Euclidean and symplectic structures on vector and affine spaces, see [3] and relevant sections in [21]; for detailed expositions on symplectic geometry, Hamiltonian mechanics, Poisson brackets, and the structure of cotangent bundles, see [2] alongside comprehensive texts such as [21] [14]; the theory of symmetries, Lie group actions, and moment maps is extensively covered in [23] [14] [33] [27]; and for an introduction to prequantization bundles and geometric quantization, which provides context for concepts like the moment map, refer to [37] [4].

2.1 Bilinear Forms in Linear Algebra

We begin by establishing the algebraic foundations upon which symplectic geometry rests: the theory of bilinear forms on vector spaces, with a particular focus on the skew-symmetric variant that defines linear symplectic structures. Understanding this framework, and contrasting it with the more familiar Euclidean structure created from inner products, is essential for appreciating the unique characteristics of symplectic manifolds.

2.1.1 Bilinear Forms: Generalities

Let V be a finite-dimensional real vector space.

Definition 2.1. A **bilinear form** on V is a map $B : V \times V \rightarrow \mathbb{R}$ which is linear in each argument separately:

- $B(au_1 + bu_2, v) = aB(u_1, v) + bB(u_2, v)$ for all $u_1, u_2, v \in V$ and $a, b \in \mathbb{R}$.
- $B(u, av_1 + bv_2) = aB(u, v_1) + bB(u, v_2)$ for all $u, v_1, v_2 \in V$ and $a, b \in \mathbb{R}$.

The space of all bilinear forms on V can be naturally identified with the tensor product $V^* \otimes V^*$, where V^* denotes the dual space of V .

A bilinear form B induces two linear maps from V to its dual V^* , given by $v \mapsto B(v, \cdot)$ and $v \mapsto B(\cdot, v)$.

Definition 2.2. A bilinear form B is **non-degenerate** if any one of the following (equivalent) conditions hold:

- the induced linear maps $V \rightarrow V^*$ are isomorphisms
- for every non-zero $v \in V$, there exists some $w, w' \in V$ such that $B(v, w) \neq 0$ and $B(w', v) \neq 0$.
- if $B(v, w) = 0$ for all $w \in V$ or if $B(w, v) = 0$ for all $w \in V$, then v must be the zero vector

If we choose a basis $\{e_1, \dots, e_n\}$ for V , a bilinear form B is determined by the matrix A with entries $A_{ij} = B(e_i, e_j)$. The form B is non-degenerate if and only if this matrix A is invertible (i.e., $\det(A) \neq 0$).

Two important classes of bilinear forms are distinguished by their symmetry properties:

Definition 2.3.

- B is **symmetric** if $B(u, v) = B(v, u)$ for all $u, v \in V$. In terms of the matrix representation, this means $A = A^T$.
- B is **skew-symmetric** (or alternating) if $B(u, v) = -B(v, u)$ for all $u, v \in V$. This implies $B(v, v) = -B(v, v)$, so $2B(v, v) = 0$, which gives $B(v, v) = 0$ for all $v \in V$. The matrix representation satisfies $A = -A^T$.

The condition $B(v, v) = 0$ for skew-symmetric forms highlights an immediate contrast with familiar Euclidean structures.

2.1.2 Symplectic Vector Spaces

We now define the central algebraic object through the lens of linear algebra.

Definition 2.4. A **symplectic form** on a real vector space V is a bilinear form $\omega : V \times V \rightarrow \mathbb{R}$ that is both

1. Skew-symmetric: $\omega(u, v) = -\omega(v, u)$ for all $u, v \in V$.
2. Non-degenerate.

A vector space V equipped with a symplectic form ω is called a **symplectic vector space** (V, ω) .

Skew-symmetry immediately implies $\omega(v, v) = -\omega(v, v)$, so $\omega(v, v) = 0$ for all $v \in V$. Non-degeneracy means that the map $\omega^\flat : V \rightarrow V^*$ defined by $\omega^\flat(u) = \omega(u, \cdot)$ (where $\omega(u, \cdot)$ is the linear functional $v \mapsto \omega(u, v)$) is an isomorphism. This isomorphism is fundamental, providing a canonical way to identify vectors with covectors in a symplectic setting, analogous to the musical isomorphism provided by an inner product. Let's briefly discuss this:

Definition 2.5 (Musical Isomorphism). Let (V, g) be a Euclidean vector space.

1. The **flat map** $g^\flat : V \rightarrow V^*$ is defined by $v \mapsto v^\flat$, where $v^\flat \in V^*$ is the linear functional $v^\flat(w) = g(v, w)$ for all $w \in V$.
2. The **sharp map** $g^\sharp : V^* \rightarrow V$ is the inverse of g^\flat , defined by $\alpha \mapsto \alpha^\sharp$, where $\alpha^\sharp \in V$ is the unique vector such that $g(\alpha^\sharp, w) = \alpha(w)$ for all $w \in V$.

The terms “flat” ($^\flat$) and “sharp” ($^\sharp$) evoke the musical analogy of lowering and raising indices. If $v = \sum v^i e_i$ in an orthonormal basis $\{e_i\}$, then $v^\flat = \sum v_i e^i$ where the components $v_i = g(v, e_i) = \sum_j v^j g(e_j, e_i) = \sum_j v^j \delta_{ji} = v^i$. Thus, in an orthonormal basis, the components of a vector and its corresponding covector under the musical isomorphism are identical. The map g^\sharp takes the covector $\alpha = \sum \alpha_j e^j$ to the vector $\alpha^\sharp = \sum \alpha_j e_j$. The non-degeneracy of g ensures that g^\flat and g^\sharp are indeed isomorphisms.

Back to the symplectic form, this definition stands in contrast to that of an inner product.

Definition 2.6. An **inner product** on a real vector space V is a bilinear form $g : V \times V \rightarrow \mathbb{R}$ that is both

1. Symmetric: $g(u, v) = g(v, u)$ for all $u, v \in V$.
2. Positive-definite: $g(v, v) > 0$ for all non-zero $v \in V$.

Note that positive-definiteness implies non-degeneracy for a symmetric bilinear form. A vector space equipped with an inner product is a **Euclidean vector space**.

The fundamental difference between these two lie in symmetry and definiteness. An inner product g is symmetric and satisfies $g(v, v) > 0$ for $v \neq 0$, giving rise to notions of length ($\|v\| = \sqrt{g(v, v)}$) and angle (via $g(u, v) = \|u\|\|v\|\cos\theta$). A symplectic form ω , being skew-symmetric, necessarily satisfies $\omega(v, v) = 0$. It does not define lengths or angles but rather provides a notion of oriented area. For

vectors $u, v \in V$, the quantity $\omega(u, v)$ can be interpreted as the signed area of the parallelogram spanned by u and v , projected onto a “symplectic plane.”

Example 2.7.

- On $V = \mathbb{R}^n$, the **model inner product** is given by

$$g(x, y) = \sum_{i=1}^n x_i y_i$$

for $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$.

- The canonical example of a symplectic vector space is $V = \mathbb{R}^{2m}$ with standard coordinates $(q^1, \dots, q^m, p_1, \dots, p_m)$, equipped with the bilinear form

$$\omega_0(z, z') = \sum_{i=1}^m (q^i p'_i - q'^i p_i),$$

where $z = (q, p)$ and $z' = (q', p')$.

An important consequence of these properties relates to the dimension of the vector space.

Proposition 2.8. *Any finite-dimensional symplectic vector space (V, ω) must have even dimension.*

Proof. Let $\dim V = n$. The bilinear form ω can be represented by a matrix Ω with respect to any basis $\{e_i\}$, where $\Omega_{ij} = \omega(e_i, e_j)$. Skew-symmetry implies $\Omega^T = -\Omega$. Non-degeneracy implies $\det(\Omega) \neq 0$. From linear algebra, we know $\det(\Omega^T) = \det(\Omega)$ and $\det(-\Omega) = (-1)^n \det(\Omega)$. Therefore, $\det(\Omega) = (-1)^n \det(\Omega)$. Since $\det(\Omega) \neq 0$, we must have $(-1)^n = 1$, which requires n to be even. \square

We thus write $\dim V = 2m$ for some integer m .

Remark 2.9. *Contact geometry is the odd-dimensional analogue of symplectic geometry: whereas symplectic manifolds are even-dimensional with a closed, nondegenerate 2-form ω , contact manifolds are odd-dimensional and equipped with a 1-form α satisfying maximal non-integrability via $\alpha \wedge (d\alpha)^n \neq 0$.*

Non-degeneracy of Symplectic Forms

In this section, we explore the relationship between symplectic forms, their non-degeneracy, and the top exterior power. Our goal is to expand Remark 2.5 from [13].

Recall that every alternating bilinear form

$$\omega: V \times V \longrightarrow k, \quad \omega(v, w) = -\omega(w, v),$$

is naturally identified with a decomposable element of the second exterior power of the dual space,

$$\omega \longleftrightarrow \sum_{i < j} \omega(e_i, e_j) e_i^\vee \wedge e_j^\vee \in \bigwedge^2 V^*,$$

where $\{e_i\}_{i=1}^{2m}$ is any ordered basis of V and $\{e_i^\vee\}$ the dual basis.

Proposition 2.10. ω is non-degenerate \iff the top exterior power

$$\frac{\omega^{\wedge m}}{m!} \in \bigwedge^{2m} V^* = \det V^*$$

is nonzero. Equivalently, $\omega^{\wedge m} \neq 0$.

Proof. Choose an ordered basis $\mathcal{B} = \{e_1, \dots, e_{2m}\}$ such that the matrix of ω in \mathcal{B} is A . Under the canonical isomorphism $(\Lambda^2 V^*)^{\otimes m} \rightarrow \Lambda^{2m} V^*$, one has

$$\omega^{\wedge m} = \sum_{\sigma \in S_{2m}} \text{sgn}(\sigma) \omega(e_{\sigma(1)}, e_{\sigma(2)}) \cdots \omega(e_{\sigma(2m-1)}, e_{\sigma(2m)}) e_1^\vee \wedge \cdots \wedge e_{2m}^\vee.$$

Thus $\omega^{\wedge m} = \text{Pf}(A) e_1^\vee \wedge \cdots \wedge e_{2m}^\vee$, where $\text{Pf}(A)$ is the Pfaffian of the skew-symmetric matrix A . Because A is skew-symmetric, $\det A = \text{Pf}(A)^2$. Hence

$$\begin{aligned} \omega^{\wedge m} \neq 0 &\iff \text{Pf}(A) \neq 0 \\ &\iff \det A \neq 0 \\ &\iff A \text{ invertible} \\ &\iff \flat_\omega \text{ isomorphism.} \end{aligned}$$

Therefore ω is non-degenerate if and only if $\omega^{\wedge m}$ spans the one-dimensional determinant line $\det V^* \cong \Lambda^{2m} V^*$, completing the proof. \square

The Musical Isomorphism: Symplectic Case

Now, we present the musical isomorphism, but for the symplectic case. Similar to earlier, the non-degenerate skew-symmetric form $\omega : V \times V \rightarrow \mathbb{R}$ induces an isomorphism between V and V^* .

Definition 2.11. Let (V, ω) be a symplectic vector space.

1. The map $\omega^\flat : V \rightarrow V^*$ is defined by $v \mapsto \omega^\flat(v)$, where $\omega^\flat(v)(w) = \omega(v, w)$ for all $w \in V$.
2. The map $\omega^\sharp : V^* \rightarrow V$ is the inverse of ω^\flat .

The non-degeneracy of ω guarantees that ω^\flat is an isomorphism. In canonical coordinates $(q^1, \dots, q^m, p_1, \dots, p_m)$ with basis vectors $e_i = \frac{\partial}{\partial q^i}$ and $f_j = \frac{\partial}{\partial p_j}$, and dual basis dq^i, dp_j , we have:

- $\omega^\flat(e_i) = \omega(e_i, \cdot) = \sum_k \omega(e_i, e_k) dq^k + \sum_k \omega(e_i, f_k) dp_k = \sum_k \delta_{ik} dp_k = dp_i$.
- $\omega^\flat(f_j) = \omega(f_j, \cdot) = \sum_k \omega(f_j, e_k) dq^k + \sum_k \omega(f_j, f_k) dp_k = \sum_k (-\delta_{jk}) dq^k = -dq^j$.

Thus, ω^\flat maps $\frac{\partial}{\partial q^i} \mapsto dp_i$ and $\frac{\partial}{\partial p_j} \mapsto -dq^j$. Correspondingly, ω^\sharp maps $dq^i \mapsto -\frac{\partial}{\partial p_i}$ and $dp_j \mapsto \frac{\partial}{\partial q^j}$.

These isomorphisms are the essential tools for converting differentials of functions (df , which are 1-forms) into vector fields, leading to the concepts of gradient and Hamiltonian vector fields.

2.1.3 Model Forms

Earlier, we introduced the model inner product and the model symplectic form. One might wonder why these forms are referred to as 'model' forms. The following theorem clarifies the reasoning behind this naming convention:

Theorem 2.12 (Normal Forms). *Let V be a finite-dimensional real vector space equipped with either an inner product $\langle \cdot, \cdot \rangle$ or a symplectic form ω . Let $n = \dim V$. Then there exists a linear isomorphism $\phi : \mathbb{R}^n \rightarrow V$ such that the bilinear form on \mathbb{R}^n obtained by pulling back⁷ the form on V via ϕ coincides with the two model forms introduced earlier. Specifically:*

- If V has an inner product $\langle \cdot, \cdot \rangle$, $\phi^*\langle \cdot, \cdot \rangle$ is the model inner product on \mathbb{R}^n .
- If V has a symplectic form ω , $\phi^*\omega$ is the model symplectic form ω_0 on \mathbb{R}^n .

Proof. The proof proceeds by induction on the dimension $n = \dim V$. The base case $n = 0$ is trivial. Assume the theorem holds for dimensions less than n .

- Inner Product Case: $(V, \langle \cdot, \cdot \rangle)$. Assume $V \neq \{0\}$. Choose any non-zero vector $v_1 \in V$ and normalize it to obtain a unit vector $e_1 = \frac{v_1}{\|v_1\|}$, where $\|v_1\|^2 = \langle v_1, v_1 \rangle > 0$. Consider the orthogonal complement

$$V_1 = (\mathbb{R}e_1)^\perp = \{v \in V \mid \langle v, e_1 \rangle = 0\}.$$

Since the inner product is non-degenerate, we have the direct sum decomposition

$$V = \mathbb{R}e_1 + V_1,$$

and thus $\dim V_1 = n - 1$. The restriction of the inner product to V_1 , denoted by $\langle \cdot, \cdot \rangle|_{V_1}$, is clearly symmetric and positive-definite. Its non-degeneracy on V_1 follows from the direct sum decomposition and the non-degeneracy on V .

By the induction hypothesis, there exists an isomorphism $\phi_1 : \mathbb{R}^{n-1} \rightarrow V_1$ that pulls back $\langle \cdot, \cdot \rangle|_{V_1}$ to the standard Euclidean inner product on \mathbb{R}^{n-1} . Let $\{e_2, \dots, e_n\}$ be the orthonormal basis of V_1 corresponding via ϕ_1 to the standard basis of \mathbb{R}^{n-1} . Then $\{e_1, e_2, \dots, e_n\}$ is an orthonormal basis for V .

Finally, the isomorphism $\phi : \mathbb{R}^n \rightarrow V$, which maps the standard basis of \mathbb{R}^n to $\{e_1, e_2, \dots, e_n\}$, pulls back $\langle \cdot, \cdot \rangle$ to the standard Euclidean inner product on \mathbb{R}^n .

- Symplectic Case: (V, ω) . We do this next, see Corollary 2.16.

□

In the symplectic case, a basis that brings a general symplectic form ω into this standard structure is highly desirable.

Definition 2.13. A basis $\{e_1, \dots, e_m, f_1, \dots, f_m\}$ for a $2m$ -dimensional symplectic vector space (V, ω) is called a **symplectic basis** (or Darboux basis) if

$$\omega(e_i, e_j) = 0, \quad \omega(f_i, f_j) = 0, \quad \omega(e_i, f_j) = \delta_{ij}$$

for all $i, j \in \{1, \dots, m\}$.

In such a basis, the matrix of ω is precisely J_0 . The existence of such a basis for any symplectic vector space is a fundamental result.

Theorem 2.14. *Every finite-dimensional symplectic vector space (V, ω) admits a symplectic basis.*

⁷Recall that the pullback form $(\phi^*\beta)(x, y) = \beta(\phi(x), \phi(y))$ for $x, y \in \mathbb{R}^n$ and β a bilinear form on V .

Proof. The proof proceeds by induction on the dimension $2m$. The base case $m = 0$ is trivial. Assume the theorem holds for dimension $2(m - 1)$. Let V have dimension $2m$. Since ω is non-degenerate, it is non-zero. Pick any non-zero vector $e_1 \in V$. By non-degeneracy, there exists some $v \in V$ such that $\omega(e_1, v) \neq 0$. Rescale v to obtain f_1 such that $\omega(e_1, f_1) = 1$. Since $\omega(e_1, e_1) = 0$, e_1 and f_1 are linearly independent. Let $W = \text{span}\{e_1, f_1\}$. The restriction $\omega|_W$ is non-degenerate (its matrix is $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$). Consider the symplectic complement $W^\omega = \{v \in V \mid \omega(v, w) = 0 \text{ for all } w \in W\}$. One shows that $V = W + W^\omega$ and that $\omega|_{W^\omega}$ is non-degenerate, making $(W^\omega, \omega|_{W^\omega})$ a symplectic vector space of dimension $2(m - 1)$. By the induction hypothesis, W^ω has a symplectic basis $\{e_2, \dots, e_m, f_2, \dots, f_m\}$. Combining this with $\{e_1, f_1\}$ gives a symplectic basis for V . \square

An immediate corollary follows, but before we state it, we need to introduce one more definition.

Definition 2.15. A **symplectorphism** is an isomorphism that preserves the symplectic form.

Now we state the corollary:

Corollary 2.16. Any $2m$ -dimensional symplectic vector space (V, ω) is symplectomorphic to the standard symplectic space $(\mathbb{R}^{2m}, \omega_0)$.

Theorem 2.12 establishes that any finite-dimensional inner product space or symplectic vector space is isomorphic to a standard model space $(\mathbb{R}^n, \text{model form})$. This fundamental classification result has direct consequences for understanding the structure of the spaces of all such forms on a given vector space, specifically \mathbb{R}^n .

2.1.4 The Spaces $\text{Met}(\mathbb{R}^n)$ and $\text{Symp}(\mathbb{R}^n)$

We consider the action of the general linear group $\text{GL}(n, \mathbb{R})$ on these spaces. The standard (left) action of $G \in \text{GL}(n, \mathbb{R})$ on a bilinear form β on \mathbb{R}^n is given by pullback via the inverse transformation:

$$(G \cdot \beta)(x, y) = \beta(G^{-1}x, G^{-1}y) = ((G^{-1})^* \beta)(x, y).$$

Theorem 2.12, when applied to $V = \mathbb{R}^n$, guarantees that for any inner product g or symplectic form ω on \mathbb{R}^n , there exists $\phi \in \text{GL}(n, \mathbb{R})$ such that

$$\phi^* g = g_{\text{std}} \quad \text{or} \quad \phi^* \omega = \omega_0.$$

This implies that the action is transitive: given any two forms β_1, β_2 of the same type (both inner products or both symplectic), there exist $\phi_1, \phi_2 \in \text{GL}(n, \mathbb{R})$ such that

$$\phi_1^* \beta_1 = \beta_{\text{model}} \quad \text{and} \quad \phi_2^* \beta_2 = \beta_{\text{model}}.$$

Thus, we have $\phi_1^* \beta_1 = \phi_2^* \beta_2$. Applying $(\phi_1^{-1})^*$ yields

$$\beta_1 = (\phi_1^{-1})^* (\phi_2^* \beta_2) = (\phi_2 \circ \phi_1^{-1})^* \beta_2.$$

Let $A = \phi_2 \circ \phi_1^{-1} \in \text{GL}(n, \mathbb{R})$. Then $A^* \beta_2 = \beta_1$. If we define $G = A^{-1} = \phi_1 \circ \phi_2^{-1}$, then

$$(G^{-1})^* \beta_2 = \beta_1,$$

which implies that $G \cdot \beta_2 = \beta_1$. Thus, there exists $G \in \text{GL}(n, \mathbb{R})$ mapping β_2 to β_1 , establishing transitivity.

Corollary 2.17. *The group $\mathrm{GL}(2m, \mathbb{R})$ acts transitively on the space $\mathrm{Met}(\mathbb{R}^{2m})$ of inner products on \mathbb{R}^{2m} and on the space $\mathrm{Symp}(\mathbb{R}^{2m})$ of symplectic forms on \mathbb{R}^{2m} .*

This transitivity implies that the spaces $\mathrm{Met}(\mathbb{R}^n)$ and $\mathrm{Symp}(\mathbb{R}^n)$ can be identified with quotient spaces (homogeneous spaces) of $\mathrm{GL}(n, \mathbb{R})$. Consider the maps

$$\pi_{\mathrm{Met}} : \mathrm{GL}(n, \mathbb{R}) \longrightarrow \mathrm{Met}(\mathbb{R}^n), \quad A \longmapsto A^* g_{\mathrm{std}} \quad (2.1)$$

$$\pi_{\mathrm{Symp}} : \mathrm{GL}(n, \mathbb{R}) \longrightarrow \mathrm{Symp}(\mathbb{R}^n), \quad A \longmapsto A^* \omega_0 \quad (\text{for } n = 2m) \quad (2.2)$$

where g_{std} is the model inner product and ω_0 is the model symplectic form. Corollary 2.17 is equivalent to the statement that these maps are surjective: If $A^* \beta_{\mathrm{model}} = \beta$, then taking $G = A^{-1}$ gives $G \cdot \beta_{\mathrm{model}} = (G^{-1})^* \beta_{\mathrm{model}} = A^* \beta_{\mathrm{model}} = \beta$, showing that the orbit of β_{model} under the group action is the entire space.

The structure of these spaces is revealed by identifying the stabilizer subgroups of the model forms under the map $A \mapsto A^* \beta_{\mathrm{model}}$.

- For the standard inner product $g_{\mathrm{std}} = I_n$, the stabilizer consists of $A \in \mathrm{GL}(n, \mathbb{R})$ such that $A^* g_{\mathrm{std}} = g_{\mathrm{std}}$. This means $A^T I_n A = I_n$, or $A^T A = I_n$. These are precisely the orthogonal matrices, forming the **orthogonal group** $\mathrm{O}(n)$.
- For the standard symplectic form ω_0 (represented by matrix J_{2m}), the stabilizer consists of $A \in \mathrm{GL}(2m, \mathbb{R})$ such that $A^* \omega_0 = \omega_0$. This means $A^T J_{2m} A = J_{2m}$. These are the symplectic matrices, forming the **symplectic group** $\mathrm{Sp}(2m, \mathbb{R})$.

The spaces $\mathrm{Met}(\mathbb{R}^n)$ and $\mathrm{Symp}(\mathbb{R}^{2m})$ are smooth manifolds (in fact, open subsets of the vector spaces $\mathrm{SymBil}(\mathbb{R}^n; \mathbb{R})$ of symmetric bilinear forms and $\Lambda^2(\mathbb{R}^{2m})^*$ of skew-symmetric bilinear forms, respectively). The action of the Lie group $\mathrm{GL}(n, \mathbb{R})$ is smooth and transitive. By the Orbit-Stabilizer Theorem for smooth, transitive Lie group actions, the orbit map π induces a diffeomorphism between the quotient space $\mathrm{GL}(n, \mathbb{R})/\mathrm{Stab}(\beta_{\mathrm{model}})$ and the orbit $\pi(\mathrm{GL}(n, \mathbb{R}))$, which is the entire space $\mathrm{Met}(\mathbb{R}^n)$ or $\mathrm{Symp}(\mathbb{R}^{2m})$. This yields the following identifications:

$$\mathrm{GL}(n, \mathbb{R})/\mathrm{O}(n) \xrightarrow{\cong} \mathrm{Met}(\mathbb{R}^n), \quad [A] \longmapsto A^* g_{\mathrm{std}} \quad (2.3)$$

$$\mathrm{GL}(2m, \mathbb{R})/\mathrm{Sp}(2m, \mathbb{R}) \xrightarrow{\cong} \mathrm{Symp}(\mathbb{R}^{2m}), \quad [A] \longmapsto A^* \omega_0 \quad (2.4)$$

These are diffeomorphisms of homogeneous $\mathrm{GL}(n, \mathbb{R})$ -manifolds.

The distinct features of the stabilizer subgroups $\mathrm{O}(n)$ and $\mathrm{Sp}(2m, \mathbb{R})$ lead to big differences in the geometry and topology of the corresponding spaces of forms. A big difference lies in compactness:

- The orthogonal group $\mathrm{O}(n)$ is compact. Geometrically, this relates to the fact that orthogonal transformations preserve lengths and angles. Furthermore, the inclusion $\mathrm{O}(n) \hookrightarrow \mathrm{GL}(n, \mathbb{R})$ is a homotopy equivalence; $\mathrm{GL}(n, \mathbb{R})$ deformation retracts onto $\mathrm{O}(n)$. This retraction can be seen via the polar decomposition $A = OS$, where $A \in \mathrm{GL}(n, \mathbb{R})$, $O \in \mathrm{O}(n)$, and S is a unique symmetric positive-definite matrix. The map $A \mapsto O$ provides the retraction.
- The symplectic group $\mathrm{Sp}(2m, \mathbb{R})$ is non-compact. For instance, matrices of the form $\begin{pmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{pmatrix}$ belong to $\mathrm{Sp}(2, \mathbb{R})$ for any $\lambda \in \mathbb{R}^\times$, showing unboundedness. The inclusion $\mathrm{Sp}(2m, \mathbb{R}) \hookrightarrow \mathrm{GL}(2m, \mathbb{R})$ is not a homotopy equivalence.

These properties of the stabilizers have significant consequences for the topology of the quotient spaces $\text{Met}(\mathbb{R}^n)$ and $\text{Symp}(\mathbb{R}^{2m})$. The identification $\text{Met}(\mathbb{R}^n) \cong \text{GL}(n, \mathbb{R})/\text{O}(n)$ allows us to deduce its topology. The polar decomposition $A = OS$ relates elements of $\text{GL}(n, \mathbb{R})$ to pairs (O, S) where $O \in \text{O}(n)$ and S is symmetric positive-definite. The map $A \mapsto S = (A^T A)^{1/2}$ identifies the quotient space $\text{GL}(n, \mathbb{R})/\text{O}(n)$ with the space \mathcal{P}_n of $n \times n$ symmetric positive-definite matrices. \mathcal{P}_n is a convex cone within the vector space of symmetric matrices (see Remark 2.18 below) and is therefore contractible. The homotopy equivalence $\text{GL}(n, \mathbb{R}) \simeq \text{O}(n)$ is consistent with this, as the long exact sequence of homotopy groups for the fibration $\text{O}(n) \rightarrow \text{GL}(n, \mathbb{R}) \rightarrow \text{GL}(n, \mathbb{R})/\text{O}(n)$ implies that $\pi_k(\text{GL}(n, \mathbb{R})/\text{O}(n))$ is trivial for $k \geq 1$.

In contrast, since $\text{Sp}(2m, \mathbb{R})$ is non-compact and not homotopy equivalent to $\text{GL}(2m, \mathbb{R})$, the quotient space $\text{Symp}(\mathbb{R}^{2m}) \cong \text{GL}(2m, \mathbb{R})/\text{Sp}(2m, \mathbb{R})$ inherits non-trivial topology. The lowest dimensional case $m = 1$ ($n = 2$) illustrates this clearly. Here $\text{Sp}(2, \mathbb{R})$ coincides with the special linear group $\text{SL}(2, \mathbb{R})$, the group of 2×2 matrices with determinant 1. The standard symplectic form ω_0 on \mathbb{R}^2 is precisely the determinant (area) form. $\text{SL}(2, \mathbb{R})$ preserves this form. The determinant homomorphism $\det : \text{GL}(2, \mathbb{R}) \rightarrow \mathbb{R}^\times$ has kernel $\text{SL}(2, \mathbb{R})$. By the first isomorphism theorem for Lie groups, this induces a diffeomorphism $\text{GL}(2, \mathbb{R})/\text{SL}(2, \mathbb{R}) \cong \mathbb{R}^\times$. Therefore, $\text{Symp}(\mathbb{R}^2) \cong \mathbb{R}^\times$. Since $\mathbb{R}^\times = \mathbb{R} \setminus \{0\}$ is not contractible (it has two connected components and non-trivial π_0), the space $\text{Symp}(\mathbb{R}^2)$ is topologically distinct from the contractible space $\text{Met}(\mathbb{R}^2)$. The topology of $\text{Symp}(\mathbb{R}^{2m})$ becomes more complex for $m > 1$.

The following table summarizes the key distinctions:

Table 2: Comparison: Inner Product vs. Symplectic Vector Spaces

Feature	Inner Product Space (\mathbb{R}^n, g)	Symplectic Vector Space $(\mathbb{R}^{2m}, \omega)$
Bilinear Form Type	Symmetric, Positive-Definite	Skew-Symmetric, Non-Degenerate
Dimension	Any $n \geq 1$	Must be even, $n = 2m$
Model Space	Euclidean $(\mathbb{R}^n, \delta_{ij})$	Standard Symplectic $(\mathbb{R}^{2m}, \omega_0)$
Space of Forms	$\text{Met}(\mathbb{R}^n)$	$\text{Symp}(\mathbb{R}^{2m})$
Homogeneous Space	$\text{GL}(n, \mathbb{R})/\text{O}(n)$	$\text{GL}(2m, \mathbb{R})/\text{Sp}(2m, \mathbb{R})$
Stabilizer Group	$\text{O}(n)$ (Orthogonal Group)	$\text{Sp}(2m, \mathbb{R})$ (Symplectic Group)
Stabilizer Property	Compact	Non-Compact
Space Topology	Contractible	Non-Contractible (e.g., $\cong \mathbb{R}^\times$ for $m = 1$)

Remark 2.18.

1. The space $\text{Met}(\mathbb{R}^n)$ of inner products on \mathbb{R}^n can be identified with the set \mathcal{P}_n of symmetric positive-definite $n \times n$ matrices. This set is a subset of the vector space $\text{SymBil}(\mathbb{R}^n; \mathbb{R})$ of all symmetric bilinear forms (identified with symmetric matrices). \mathcal{P}_n is a convex subset: if $g_1, g_2 \in \mathcal{P}_n$ and $t \in (0, 1)$, then for any $v \neq 0$, $(tg_1 + (1-t)g_2)(v, v) = tg_1(v, v) + (1-t)g_2(v, v) > 0$ (since $g_1(v, v) > 0$, $g_2(v, v) > 0$, and $t, 1-t$ are non-negative with at least one positive if $t \in (0, 1)$). Furthermore, if $g \in \mathcal{P}_n$ and $c > 0$, then $cg \in \mathcal{P}_n$, so \mathcal{P}_n is an open convex cone. Any convex subset of a Euclidean space is contractible (e.g., via the linear homotopy

$H(g, t) = (1 - t)g + tg_0$ to a fixed point $g_0 \in \mathcal{P}_n$). This provides an alternative confirmation that $\text{Met}(\mathbb{R}^n)$ is contractible.

2. As established above, the contrast for $m = 1$ is sharp: $\text{Symp}(\mathbb{R}^2) \cong \text{GL}(2, \mathbb{R})/\text{SL}(2, \mathbb{R}) \cong \mathbb{R}^\times$, which is not contractible. The diffeomorphism $\text{GL}(2, \mathbb{R})/\text{SL}(2, \mathbb{R}) \cong \mathbb{R}^\times$ is realized by the determinant map. For higher dimensions ($m > 1$), the space $\text{Symp}(\mathbb{R}^{2m}) \cong \text{GL}(2m, \mathbb{R})/\text{Sp}(2m, \mathbb{R})$ continues to have non-trivial topology, which is considerably more complex than \mathbb{R}^\times .
3. The topological assertions made for the model spaces $\text{Met}(\mathbb{R}^n)$ and $\text{Symp}(\mathbb{R}^{2m})$ extend to the corresponding spaces $\text{Met}(V)$ and $\text{Symp}(V)$ for any n -dimensional (or $2m$ -dimensional) real vector space V . Theorem 2.12 provides an isomorphism $\phi : \mathbb{R}^n \rightarrow V$. This isomorphism induces diffeomorphisms $\phi^* : \text{Met}(V) \rightarrow \text{Met}(\mathbb{R}^n)$ and $\phi^* : \text{Symp}(V) \rightarrow \text{Symp}(\mathbb{R}^{2m})$ via pull-back. Since diffeomorphisms preserve topological properties like contractibility and homotopy type, the properties established for the model spaces hold for the spaces associated with any finite-dimensional vector space V . The details of this transport are deferred to later.

2.2 Euclidean Spaces and Affine Symplectic Spaces

While vector spaces provide the algebraic foundation for linear geometry, many geometric settings, particularly those arising in physics or as local models for manifolds, lack a naturally distinguished origin. Consider Euclidean space as taught in high school; although often identified with \mathbb{R}^n , its geometric properties related to points, lines, and parallelism are independent of where we place the origin. Similarly, the set of solutions to an inhomogeneous linear system $Ax = b$ (for $b \neq 0$) forms a geometric object (like a line or plane) that does not pass through the origin and hence is not a vector subspace, yet it shares the “flatness” property of vector subspaces. Affine spaces formalize this notion of a “vector space that has forgotten its origin.”

2.2.1 Linear Algebra, But For Affine Spaces

Definition 2.19. An **affine space** is a triple $(A, V, +)$, where A is a non-empty set whose elements are called **points**, V is a finite-dimensional real vector space called the **associated vector space** or the **space of translations**, and $+: A \times V \rightarrow A$ is a map, called the **action** of V on A , satisfying the following axioms:

- (A1) Identity: For every $a \in A$, $a + 0 = a$, where 0 is the zero vector in V .
- (A2) Associativity: For every $a \in A$ and all $u, v \in V$, $(a + u) + v = a + (u + v)$.
- (A3) Unique Translation: The action is **free and transitive**: for any pair of points $a, b \in A$, there exists a unique vector $v \in V$ such that $a + v = b$.

The unique vector v in (A3) is denoted by \vec{ab} or $b - a$. The **dimension** of the affine space A is defined as $\dim A = \dim V$.

Axiom (A3) encapsulates the idea that the vector space V acts as the group of translations on the set of points A . The action being transitive means any point can be reached from any other point by a unique translation vector. The action being free means that only the zero vector fixes any point. This definition formalizes the idea of A being a principal homogeneous space for the additive group of V .

It is important to distinguish between points $a \in A$ and vectors $v \in V$. While choosing an origin $o \in A$ allows us to identify any point a with the vector $\vec{o}a = a - o$, this identification depends entirely on the choice of o and is therefore not canonical. Geometric properties of affine spaces should be independent of such choices.

An alternative but equivalent definition uses the subtraction map.

Definition 2.20. An affine space is a pair (A, V) , where A is a non-empty set of points and V is a finite-dimensional real vector space, together with a map $- : A \times A \rightarrow V$, denoted $(b, a) \mapsto b - a$, satisfying:

- (B1) For every point $a \in A$ and every vector $v \in V$, there exists a unique point $b \in A$ such that $b - a = v$.
- (B2) Chasles' Relation: For all points $a, b, c \in A$, $(c - b) + (b - a) = c - a$.

The equivalence between these two definitions is straightforward. Given $A, V, +$, define $b - a$ as the unique vector v such that $a + v = b$. Axiom (A3) ensures existence and uniqueness, satisfying (B1). Axiom (A2) implies (B2). Conversely, given $A, V, -$, define $a + v$ as the unique point b such that $b - a = v$. Axiom (B1) ensures existence and uniqueness. Axioms (A1) and (A2) follow from (B1) and (B2).

An affine space of dimension 0 is a single point. An affine space of dimension 1 is called an **affine line**, and one of dimension 2 is called an **affine plane**. An affine subspace of dimension $\dim A - 1$ is an **affine hyperplane**.

In linear algebra, linear combinations are very fundamental and natural. However, attempting to form analogous combinations of points $\sum \lambda_i a_i$ in an affine space leads to frame dependence, unless the sum of the coefficients $\sum \lambda_i$ is equal to 1. This restriction leads to affine combinations.

The reason for the constraint $\sum \lambda_i = 1$ becomes clear when we examine independence from the origin. Let $o \in A$ be an arbitrary origin. We wish to define a point $p = \sum \lambda_i a_i$. Interpreting this relative to o , we might define the vector $\vec{o}p = \sum \lambda_i \vec{o}a_i$. If we choose a different origin o' , the corresponding vectors are $\vec{o}'a_i = \vec{o}a_i - \vec{o}o'$. The combination relative to o' would yield the vector

$$\vec{o}'p' = \sum \lambda_i \vec{o}'a_i = \sum \lambda_i (\vec{o}a_i - \vec{o}o') = \left(\sum \lambda_i \vec{o}a_i \right) - \left(\sum \lambda_i \right) \vec{o}o'.$$

For the resulting point p to be independent of the origin (i.e., $p = p'$), we require

$$\vec{o}'p' = \vec{o}p - \vec{o}o'.$$

Substituting the expressions gives

$$\left(\sum \lambda_i \vec{o}a_i \right) - \left(\sum \lambda_i \right) \vec{o}o' = \left(\sum \lambda_i \vec{o}a_i \right) - \vec{o}o'.$$

This equality holds if and only if

$$\left(\sum \lambda_i \right) \vec{o}o' = \vec{o}o',$$

which, for an arbitrary choice of o, o' , necessitates $\sum \lambda_i = 1$. This condition ensures that the definition of the combined point is belongs to the affine structure only, independent of any arbitrary origin choice.

Definition 2.21. Let (A, V) be an affine space. Let a_i for $i \in I$ be a finite family of points in A and λ_i for $i \in I$ be a corresponding family of scalars in \mathbb{R} such that $\sum_{i \in I} \lambda_i = 1$. The **affine combination** of the points a_i with weights λ_i is the point $p \in A$ defined by

$$p = o + \sum_{i \in I} \lambda_i (a_i - o)$$

for any arbitrary choice of origin $o \in A$. The definition ensures p is independent of the choice of o . We often write $p = \sum_{i \in I} \lambda_i a_i$.

Geometrically, the affine combination $\lambda_1 a_1 + \lambda_2 a_2$ with $\lambda_1 + \lambda_2 = 1$ represents a point on the line passing through a_1 and a_2 . If, in addition, $\lambda_1, \lambda_2 \geq 0$, the point lies on the line segment between a_1 and a_2 ; such combinations are called **convex combinations**.

Definition 2.22. An **affine frame** for an n -dimensional affine space A consists of a point $o \in A$ (the origin) and a basis $\{v_1, \dots, v_n\}$ for the associated vector space V . Any point $a \in A$ can be uniquely written as

$$a = o + \sum_{i=1}^n x^i v_i.$$

The scalars x^1, \dots, x^n are the **affine coordinates** of a relative to the frame $(o; \{v_i\})$.

Affine subspaces generalize the concepts of lines and planes within an affine space. They are subsets that are “flat” and closed under the formation of lines passing through any two of their points.

Definition 2.23. A subset $S \subseteq A$ of an affine space A is an **affine subspace** if either S is empty, or S is non-empty and for any family of points a_i for $i \in I$ in S and scalars λ_i for $i \in I$ with $\sum \lambda_i = 1$, the affine combination $\sum \lambda_i a_i$ is also in S .

This definition means affine subspaces are precisely the subsets closed under affine combinations. The empty set and singleton sets $\{a\}$ are affine subspaces. Any line through two distinct points is an affine subspace.

A fundamental characterization relates affine subspaces to translates of vector subspaces.

Proposition 2.24. A non-empty subset $S \subseteq A$ is an affine subspace if and only if there exists a point $a \in S$ and a unique vector subspace $U \subseteq V$ such that $S = a + U = \{a + u \mid u \in U\}$. The subspace U is called the **direction** of S , denoted \vec{S} .

Proof. Let $S \subseteq A$ be non-empty.

(\Rightarrow) Assume S is an affine subspace. Since S is non-empty, fix an arbitrary $a_0 \in S$. Define $U := \{a_0 p \mid p \in S\} = \{p - a_0 \mid p \in S\}$.

First, we show U is a vector subspace of V :

- $0_V = a_0 - a_0 \in U$ since $a_0 \in S$. So U is non-empty.
- Let $u_1, u_2 \in U$. Then $u_1 = p_1 - a_0$ and $u_2 = p_2 - a_0$ for some $p_1, p_2 \in S$. Consider the point $q = a_0 + (p_1 - a_0) + (p_2 - a_0) = p_1 + (p_2 - a_0)$. This can be written as an affine combination

$q = (1)p_1 + (1)p_2 + (-1)a_0$. Since $p_1, p_2, a_0 \in S$ and the coefficients sum to 1, $q \in S$. Then $u_1 + u_2 = (p_1 - a_0) + (p_2 - a_0) = q - a_0 \in U$. Thus, U is closed under addition.

- Let $u \in U$ and $c \in \mathbb{R}$. Then $u = p - a_0$ for some $p \in S$. Consider the point $q = (1 - c)a_0 + cp$. Since $a_0, p \in S$ and $(1 - c) + c = 1$, $q \in S$. Then $cu = c(p - a_0) = (cp - ca_0) = (cp + (1 - c)a_0 - a_0) = q - a_0 \in U$. Thus, U is closed under scalar multiplication.

So, U is a vector subspace of V .

Next, we show $S = a_0 + U$:

- If $p \in S$, then $p = a_0 + (p - a_0)$. Since $p - a_0 \in U$ by definition of U , $p \in a_0 + U$. Thus $S \subseteq a_0 + U$.
- If $x \in a_0 + U$, then $x = a_0 + u$ for some $u \in U$. By definition of U , $u = p' - a_0$ for some $p' \in S$. So $x = a_0 + (p' - a_0) = p'$. Since $p' \in S$, it follows that $x \in S$. Thus $a_0 + U \subseteq S$.

Hence $S = a_0 + U$.

Finally, we show U is unique for S . Suppose $S = a_0 + U_0$ and also $S = a_1 + U_1$ for some $a_0, a_1 \in S$ and vector subspaces U_0, U_1 . From $S = a_0 + U_0$, we have $U_0 = \{s - a_0 \mid s \in S\}$. From $S = a_1 + U_1$, we have $U_1 = \{s - a_1 \mid s \in S\}$. Since $a_1 \in S$, $a_1 = a_0 + u_0$ for some $u_0 \in U_0$. So $u_0 = a_1 - a_0$. Let $v \in U_0$. Then $v = s - a_0$ for some $s \in S$. We can write $v = (s - a_1) + (a_1 - a_0) = (s - a_1) + u_0$. Since $s \in S$ and $a_1 \in S$, $s - a_1 \in U_1$. To show $U_0 \subseteq U_1$: any $v \in U_0$ is $s - a_0$ for $s \in S$. Then $v = (s - a_1) - (a_0 - a_1)$. Since $s - a_1 \in U_1$ and $a_0 - a_1 \in U_1$ (as $U_1 = \{p - a_1 \mid p \in S\}$ and $a_0 \in S$), their difference $v \in U_1$. So $U_0 \subseteq U_1$. Similarly, $U_1 \subseteq U_0$. Thus $U_0 = U_1$. This unique subspace is denoted \vec{S} .

(\Leftarrow) Assume there exists $a_0 \in S$ and a vector subspace $U \subseteq V$ such that $S = a_0 + U$. We show S is an affine subspace (i.e., closed under affine combinations). Let $p_1, \dots, p_k \in S$ and $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ such that $\sum_{i=1}^k \lambda_i = 1$. Since $p_i \in S$, each $p_i = a_0 + u_i$ for some $u_i \in U$. The affine combination $p = \sum_{i=1}^k \lambda_i p_i$ can be expressed relative to a_0 as:

$$p = a_0 + \sum_{i=1}^k \lambda_i (p_i - a_0)$$

Substituting $p_i - a_0 = u_i$:

$$p = a_0 + \sum_{i=1}^k \lambda_i u_i$$

Since each $u_i \in U$ and U is a vector subspace, the linear combination $u_{\text{new}} = \sum_{i=1}^k \lambda_i u_i$ is also in U . Therefore, $p = a_0 + u_{\text{new}}$ where $u_{\text{new}} \in U$. This implies $p \in a_0 + U = S$. Thus, S is an affine subspace. \square

2.2.2 Affine Maps and the Affine Group

Affine maps are the structure-preserving morphisms between affine spaces. They preserve the essential affine properties like collinearity and parallelism.

Definition 2.25 (Affine Map). Let $(A, V, +)$ and $(A', V', +)$ be two affine spaces. A map $f : A \rightarrow A'$ is an **affine map** if there exists a linear map $\vec{f} : V \rightarrow V'$ such that for any $a \in A$ and $v \in V$,

$$f(a + v) = f(a) + \vec{f}(v)$$

The linear map \vec{f} is unique and is called the **associated linear map** (or linear part) of f . Equivalently, f is affine if for any points $a, b \in A$, $\vec{f}(\vec{ab}) = \vec{f}(a) - \vec{f}(b)$, i.e., $\vec{f}(b - a) = f(b) - f(a)$. A third equivalent condition is that f preserves affine combinations:

$$f\left(\sum \lambda_i a_i\right) = \sum \lambda_i f(a_i)$$

whenever $\sum \lambda_i = 1$.

Affine maps send lines to lines (or points), planes to planes (or lines, or points), and preserve parallelism and ratios of lengths along parallel lines. The composition of affine maps is affine. Any affine map f can be expressed, relative to a choice of origin $o \in A$, as the composition of its linear part \vec{f} (acting on V identified with A via o) followed by a translation by the vector $of(\vec{o})$:

$$f(x) = f(o) + \vec{f}(x - o)$$

Definition 2.26. An **affine transformation** is an invertible affine map from an affine space A to itself. The set of all affine transformations of A forms a group under composition, called the **affine group** of A , denoted $\text{Aff}(A)$. Notably, the inverse of an affine transformation is also affine, which is a nontrivial property: in general, the inverse of a structure-preserving map is not guaranteed to preserve the same structure (e.g., a continuous bijection need not be a homeomorphism in topology). The fact that affine maps are closed under inversion is essential for $\text{Aff}(A)$ to form a group.

How closely is the structure of the affine group related to the structure of the affine space? The answer is a lot.

Theorem 2.27. *Let A be an n -dimensional affine space with associated vector space V . The affine group $\text{Aff}(A)$ is isomorphic to the semidirect product $V \rtimes \text{GL}(V)$, where $\text{GL}(V)$ is the general linear group of V acting on V by its natural action (i.e., linear transformations). The group operation on $V \rtimes \text{GL}(V)$ is given by*

$$(v, f)(w, g) = (v + f(w), fg),$$

where $v, w \in V$ and $f, g \in \text{GL}(V)$. Thus,

$$\text{Aff}(A) \cong V \rtimes \text{GL}(V).$$

Proof. An affine transformation $f \in \text{Aff}(A)$ is determined by its associated linear map $\vec{f} \in \text{GL}(V)$ and the image $f(o)$ of a chosen origin $o \in A$. Let $v = of(\vec{o}) = f(o) - o$. We can represent f by the pair $(v, \vec{f}) \in V \times \text{GL}(V)$. Consider the composition $f_1 \circ f_2$, represented by (v_1, M_1) and (v_2, M_2) respectively, where $M_1 = \vec{f}_1$ and $M_2 = \vec{f}_2$. Let $a \in A$.

$$(f_1 \circ f_2)(a) = f_1(f_2(a)) = f_1(o + \vec{o}a + v_2 + M_2(\vec{o}a)) = f_1(f_2(o + \vec{o}a)) = f_1(f_2(o) + M_2(\vec{o}a))$$

After some algebraic manipulation, the result corresponds to the pair $(v_1 + M_1 v_2, M_1 M_2)$, which defines the group operation on $V \rtimes \text{GL}(V)$. This confirms the semidirect product structure. \square

The semidirect product structure reflects the geometric fact that translations are transformed by the linear part of an affine map. If we conjugate a translation $T_u(x) = x + u$ (represented by (u, I)) by an affine map $f(x) = Mx + v$ (represented by (v, M)), the result is another translation $T_{Mu}(x) = x + Mu$ (represented by (Mu, I)).

Affine transformations can be represented using $(n + 1) \times (n + 1)$ matrices acting on homogeneous coordinates. If $a \in A$ corresponds to $\begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix}$ and $f(a)$ corresponds to $\begin{pmatrix} \mathbf{x}' \\ 1 \end{pmatrix}$, then f represented by (v, M) corresponds to the matrix multiplication:

$$\begin{pmatrix} \mathbf{x}' \\ 1 \end{pmatrix} = \begin{pmatrix} M & v \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix}$$

Composition of affine transformations corresponds to multiplication of these augmented matrices.

Affine geometry provides a framework for parallelism and collinearity but lacks notions of distance, length, and angle. Introducing these metric concepts requires equipping the underlying vector space with an inner product. But before we get there, let's (rapidly) review some basic notions of Euclidean vector spaces from linear algebra.

Definition 2.28 (Inner Product). An **inner product** (or scalar product) on a real vector space V is a map $g : V \times V \rightarrow \mathbb{R}$ that is:

1. **Symmetric:** $g(u, v) = g(v, u)$ for all $u, v \in V$.
2. **Bilinear:** $g(au + bv, w) = ag(u, w) + bg(v, w)$ and $g(u, av + bw) = ag(u, v) + bg(u, w)$ for all $u, v, w \in V$ and $a, b \in \mathbb{R}$.
3. **Positive-definite:** $g(v, v) \geq 0$ for all $v \in V$, and $g(v, v) = 0$ if and only if $v = 0$.

A finite-dimensional real vector space equipped with an inner product (V, g) is called a **Euclidean vector space**.

The inner product induces a **norm** on V , defined by $\|v\| = \sqrt{g(v, v)}$. This norm satisfies the standard properties: $\|v\| \geq 0$ with equality if and only if $v = 0$, $\|\lambda v\| = |\lambda| \|v\|$, and the triangle inequality $\|u + v\| \leq \|u\| + \|v\|$.

Proposition 2.29. [Cauchy-Schwarz Inequality] For any u, v in a Euclidean vector space (V, g) ,

$$|g(u, v)| \leq \|u\| \|v\|.$$

Equality holds if and only if u and v are linearly dependent.

The inner product allows us to define orthogonality.

Definition 2.30 (Orthogonality). Two vectors $u, v \in V$ are **orthogonal**, denoted $u \perp v$, if $g(u, v) = 0$. A set of vectors $\{v_i\}$ is **orthogonal** if $g(v_i, v_j) = 0$ for all $i \neq j$, and **orthonormal** if $g(v_i, v_j) = \delta_{ij}$ (the Kronecker delta).

Euclidean geometry is the geometry of **Euclidean spaces**, which are vector spaces equipped with an inner product, allowing the measurement of distance, angles, and orthogonality.

2.2.3 Euclidean Affine Spaces

We now combine the affine structure with the Euclidean vector space structure.

Definition 2.31. A **Euclidean affine space** (or simply Euclidean space) E is an affine space whose associated vector space V is a Euclidean vector space (V, g) .

The inner product g on V allows us to define metric concepts on the affine space E .

Metric Tensor. The inner product g on V can be viewed as a constant **metric tensor field** on E . For any point $a \in E$, the tangent space $T_a E$ is canonically isomorphic to V (via $v \mapsto (a, v)$ in the trivialization $TE \cong E \times V$, or thinking of tangent vectors as equivalence classes of curves originating at a). We define the metric tensor g_a at a to be the inner product g on $T_a E \cong V$. The “constancy” means that under parallel transport (which is trivial in an affine space), the metric is invariant; equivalently, its components are constant in Cartesian coordinates.

Distance. The distance between two points $a, b \in E$ is defined as the norm of the unique translation vector connecting them:

$$d(a, b) = \|\vec{ab}\| = \|b - a\| = \sqrt{g(b - a, b - a)}$$

This distance function satisfies the axioms of a metric space: $d(a, b) \geq 0$ with equality if and only if $a = b$; $d(a, b) = d(b, a)$; and $d(a, c) \leq d(a, b) + d(b, c)$ (triangle inequality, follows from the norm property $\|u + v\| \leq \|u\| + \|v\|$).

Angles. The angle θ between two non-zero vectors $u, v \in V$ is defined via the inner product:

$$\cos \theta = \frac{g(u, v)}{\|u\| \|v\|}, \quad \theta \in [0, \pi]$$

The Cauchy-Schwarz inequality ensures that $-1 \leq \cos \theta \leq 1$. This definition extends to angles between intersecting affine subspaces (lines, planes) by considering the angles between vectors spanning their direction spaces.

Orthogonality. Orthogonality extends naturally from vectors to affine objects. Two vectors \vec{ab} and \vec{cd} are orthogonal if $g(\vec{ab}, \vec{cd}) = 0$. Two affine subspaces S_1, S_2 are orthogonal if their direction spaces \vec{S}_1, \vec{S}_2 are orthogonal, meaning $g(u, v) = 0$ for all $u \in \vec{S}_1, v \in \vec{S}_2$. This allows defining the **orthogonal projection** of a point onto an affine subspace: the projection p' of p onto $S = a + U$ is the unique point $p' \in S$ such that the vector $\vec{p'p} = p - p'$ is orthogonal to the direction space U . The distance from p to S is $d(p, p')$.

The existence of an inner product allows for the selection of particularly convenient bases and frames.

Definition 2.32. An **orthonormal affine frame** (or Euclidean frame) for E is a pair $(o; \{e_1, \dots, e_n\})$ where $o \in E$ is an origin and $\{e_1, \dots, e_n\}$ is an orthonormal basis for V .

Theorem 2.33. Every finite-dimensional Euclidean vector space (V, g) admits an orthonormal basis.

Proof. Gram-Schmidt orthogonalization process. Starting with an arbitrary basis $\{v_1, \dots, v_n\}$, one constructs an orthogonal basis $\{u_1, \dots, u_n\}$ via $u_1 = v_1$, $u_k = v_k - \sum_{j=1}^{k-1} \text{proj}_{u_j}(v_k)$, where $\text{proj}_u(v) = \frac{g(v, u)}{g(u, u)}u$. Then, an orthonormal basis is obtained by normalizing: $e_k = \frac{u_k}{\|u_k\|}$. \square

An orthonormal affine frame $(o; \{e_i\})$ defines a **Euclidean coordinate system** (or Cartesian coordinate system) x^1, \dots, x^n on E . Any point $a \in E$ has unique coordinates $x^1(a), \dots, x^n(a)$ such that:

$$a = o + \sum_{i=1}^n x^i(a) e_i$$

In such coordinates, the vector connecting two points a and b is:

$$\vec{ab} = b - a = \sum_{i=1}^n (x^i(b) - x^i(a)) e_i$$

The inner product of two vectors $u = \sum u^i e_i$ and $v = \sum v^j e_j$ takes the standard Euclidean form:

$$g(u, v) = g\left(\sum_i u^i e_i, \sum_j v^j e_j\right) = \sum_{i,j} u^i v^j g(e_i, e_j) = \sum_{i,j} u^i v^j \delta_{ij} = \sum_{i=1}^n u^i v^i$$

This leads to a important property regarding the metric tensor field g on E . The coordinate vector fields associated with a Euclidean coordinate system are $\frac{\partial}{\partial x^i}$, which correspond to the constant vector fields e_i under the identification $T_a E \cong V$. The components of the metric tensor in these coordinates are:

$$g_{ij} = g\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) = g(e_i, e_j) = \delta_{ij}$$

These components are constant functions on E .

Proposition 2.34. *In any Euclidean coordinate system x^1, \dots, x^n associated with an orthonormal affine frame, the metric tensor g takes the standard constant form:*

$$g = \sum_{i=1}^n dx^i \otimes dx^i$$

where $\{dx^i\}$ is the basis dual to $\{\frac{\partial}{\partial x^i}\}$. The matrix representation of g in this basis is the identity matrix I_n .

Recall that isometries are the symmetry transformations of Euclidean space – the maps that preserve its metric structure.

Definition 2.35 (Isometry). Let E be a Euclidean affine space. A map $f : E \rightarrow E$ is an **isometry** if it preserves distances:

$$d(f(a), f(b)) = d(a, b) \quad \text{for all } a, b \in E$$

Theorem 2.36. *The Euclidean group $Euc(n)$ is isomorphic to the semidirect product $\mathbb{R}^n \rtimes O(n)$, where \mathbb{R}^n represents the group of translations $T(n)$.*

$$Euc(n) \cong T(n) \rtimes O(n) \cong \mathbb{R}^n \rtimes O(n)$$

Similarly, the special Euclidean group $SE(n)$ is isomorphic to $\mathbb{R}^n \rtimes SO(n)$.

Proof. An isometry f is represented by (b, A) , where $b = f(0) - 0$ is the translation vector and $A = \vec{f} \in O(n)$ is the orthogonal linear part. The composition rule is

$$(b_1, A_1) \cdot (b_2, A_2) = (b_1 + A_1 b_2, A_1 A_2).$$

This is the semidirect product structure $\mathbb{R}^n \rtimes O(n)$ with the natural action of $O(n)$ on \mathbb{R}^n . The subgroup of translations

$$T(n) = \{(b, I) \mid b \in \mathbb{R}^n\}$$

is normal in $\text{Euc}(n)$. The subgroup $\text{SE}(n)$ corresponds to pairs (b, A) with $A \in \text{SO}(n)$. \square

The Euclidean group $\text{Euc}(n)$ is a Lie group of dimension

$$n + \dim(O(n)) = n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2}.$$

This finite dimensionality reflects the rigidity of Euclidean geometry; there are only a limited number of ways (translations, rotations, reflections) to move objects around without changing distances or angles.

2.2.4 Affine Symplectic Spaces

Previously, we explored symplectic vector spaces and affine vector spaces separately. It is now natural to consider how these structures can be combined. In what follows, we introduce the affine analogue of symplectic vector spaces.

Definition 2.37. An **affine symplectic space** is a pair (A, ω) where A is an affine space with associated vector space V , and (V, ω) is a symplectic vector space.

The symplectic form ω on V induces a **constant symplectic 2-form field** on A , also denoted by ω . At each point $a \in A$, the form ω_a on $T_a A \cong V$ is identified with ω . “Constant” means that the components ω_{ij} are constant in any canonical coordinate system. As a constant form on an affine space (which is diffeomorphic to \mathbb{R}^{2m}), ω is automatically closed, i.e., $d\omega = 0$. This contrasts with general symplectic manifolds where $d\omega = 0$ is a non-trivial condition.

Earlier, we discussed Darboux’s theorem, which roughly states that locally, all symplectic manifolds look the same. For the specific case of an affine symplectic space (A, ω) , where ω is constant, the local nature of Darboux’s theorem can be strengthened to a global statement, thanks to the existence of a global symplectic basis for the underlying vector space V .

Theorem 2.38. *Let (A, ω) be a $2m$ -dimensional affine symplectic space with associated vector space (V, ω) . There exists a global affine coordinate system $(q^1, \dots, q^m, p_1, \dots, p_m)$ on A such that the constant symplectic form ω is expressed as*

$$\omega = \sum_{i=1}^m dq^i \wedge dp_i$$

*everywhere on A . Such coordinates are called **canonical coordinates** or **Darboux coordinates**.*

Proof. By Theorem 2.14, the symplectic vector space (V, ω) admits a symplectic basis $\{e_1, \dots, e_m, f_1, \dots, f_m\}$ such that $\omega(e_i, e_j) = \omega(f_i, f_j) = 0$ and $\omega(e_i, f_j) = \delta_{ij}$. Choose an arbitrary origin $o \in A$. Define the affine coordinate functions $(q^1, \dots, q^m, p_1, \dots, p_m)$ such that for any point $a \in A$, the unique vector $\vec{oa} = a - o$ is given by

$$a - o = \sum_{i=1}^m q^i(a) e_i + \sum_{j=1}^m p_j(a) f_j.$$

This defines a global affine coordinate system on A . The coordinate vector fields are $\frac{\partial}{\partial q^i}$ (corresponding to e_i) and $\frac{\partial}{\partial p_j}$ (corresponding to f_j). The coordinate 1-forms dq^i and dp_j form the dual basis to these vector fields in each tangent space $T_a A \cong V$. Specifically,

$$dq^i(e_k) = \delta_k^i, \quad dq^i(f_k) = 0, \quad dp_j(e_k) = 0, \quad dp_j(f_k) = \delta_{jk}.$$

We evaluate the constant 2-form ω in these coordinates. For any pair of coordinate vector fields X, Y , $\omega(X, Y)$ is a constant function on A .

- $\omega\left(\frac{\partial}{\partial q^i}, \frac{\partial}{\partial q^k}\right) = \omega(e_i, e_k) = 0$
- $\omega\left(\frac{\partial}{\partial p_j}, \frac{\partial}{\partial p_l}\right) = \omega(f_j, f_l) = 0$
- $\omega\left(\frac{\partial}{\partial q^i}, \frac{\partial}{\partial p_j}\right) = \omega(e_i, f_j) = \delta_{ij}$

The 2-form $\sum_{k=1}^m dq^k \wedge dp_k$ acts on pairs of basis vectors as:

- $(\sum_k dq^k \wedge dp_k)\left(\frac{\partial}{\partial q^i}, \frac{\partial}{\partial q^l}\right) = 0$
- $(\sum_k dq^k \wedge dp_k)\left(\frac{\partial}{\partial p_j}, \frac{\partial}{\partial p_l}\right) = 0$
- $(\sum_k dq^k \wedge dp_k)\left(\frac{\partial}{\partial q^i}, \frac{\partial}{\partial p_j}\right) = \delta_{ij}$

Since ω and $\sum dq^k \wedge dp_k$ agree on all pairs of basis vectors, they are equal everywhere. Thus, the constant form ω takes the standard canonical form globally in these affine coordinates. \square

This global result highlights a key difference between affine symplectic spaces and general symplectic manifolds, where Darboux coordinates are only guaranteed locally. It implies that all affine symplectic spaces of dimension $2m$ are globally symplectomorphic (up to the choice of origin). While this suggests significant flexibility compared to the rigidity of Euclidean space (where global isometries exist but don't trivialize the structure), the existence of a global standard form simplifies the structure compared to general symplectic manifolds, which may have complex global topology preventing such a coordinate system.

Now, we discuss affine symplectomorphisms. The structure-preserving transformations of affine symplectic spaces are those affine maps whose linear part preserves the symplectic form.

Definition 2.39. An **affine symplectomorphism** of an affine symplectic space (A, ω) is an affine transformation $f : A \rightarrow A$ such that its associated linear map $\vec{f} : V \rightarrow V$ is a linear symplectomorphism, i.e., $\vec{f}^* \omega = \omega$. Equivalently, $\omega(\vec{f}(u), \vec{f}(v)) = \omega(u, v)$ for all $u, v \in V$.

Definition 2.40. The set of all affine symplectomorphisms of (A, ω) forms a group under composition, called the **affine symplectic group**, denoted $\text{Aff}_\omega(A)$. This group is a subgroup of the

general affine group $\text{Aff}(A)$, consisting of those affine transformations that preserve the symplectic structure.

In summary, here's a comparison of the key points we've discussed regarding Euclidean affine spaces and affine symplectic spaces.

Table 3: Succinct Comparison: Euclidean Affine vs. Affine Symplectic Spaces

Feature	Euclidean Affine (E)	Affine Symplectic ((A, ω))
Added Structure on V	Inner Product g (Symmetric, Bilinear, Positive-Definite)	Symplectic Form ω (Skew-Symmetric, Bilinear, Non-Degenerate)
Induced Field on A	Constant Metric Tensor g	Constant Symplectic Form ω ($d\omega = 0$)
Key Concepts	Distance, Angles, Orthogonality	Canonical Structure
Special Basis/Frame	Orthonormal ($g(e_i, e_j) = \delta_{ij}$)	Symplectic ($\omega(e_i, f_j) = \delta_{ij}$, others 0)
Form in Special Coords	$g = \sum dx^i \otimes dx^i$ (Cartesian)	$\omega = \sum dq^i \wedge dp_i$ (Darboux, Global)
Structure-Preserving Maps	Isometries (preserve distance d)	Affine Symplectomorphisms (preserve ω)
Linear Part \vec{f}	Orthogonal ($\vec{f} \in O(V, g)$)	Symplectic ($\vec{f} \in \text{Sp}(V, \omega)$)
Symmetry Group	$\text{Euc}(n) \cong V \rtimes O(V)$	$\text{Aff}_\omega(A)$ ($\vec{f} \in \text{Sp}(V, \omega)$)

2.2.5 Vector Fields, Flows, and Gradients on Affine Spaces

Let (A, V) be a real affine space of dimension $2n$, where V is the vector space of translations. We assume A is endowed with a symplectic form ω , transforming (A, V, ω) into an affine symplectic space. Recall that ω is a non-degenerate, skew-symmetric bilinear form on V . Let $U \subseteq A$ be an open subset.

The flatness of the affine space A permits a canonical identification between the tangent space $T_p A$ at any point $p \in A$ and the vector space V . Specifically, for any $p \in A$, the map $v \mapsto p + v$ provides an identification of V with A . The differential of this map identifies $T_0 V \cong V$ with $T_p A$. Consequently, the tangent bundle TA is trivial, $TA \cong A \times V$. This allows for a simplified definition of vector fields compared to the general manifold setting.

Definition 2.41 (Vector Field). A **vector field** on U is a smooth map $\xi : U \rightarrow V$. We denote the space of smooth vector fields on U by $\mathfrak{X}(U)$. At each point $p \in U$, $\xi(p) \in V$ is interpreted as a tangent vector attached to p , residing in the canonically identified space $T_p A \cong V$.

Similarly, we define differential forms by leveraging the canonical identification of cotangent spaces $T_p^* A$ with the dual space V^* .

Definition 2.42 (Differential Form). A **differential k -form** on U is a smooth map $\alpha : U \rightarrow \Lambda^k V^*$, where $\Lambda^k V^*$ denotes the k -th exterior power of the dual space V^* . We denote the space of smooth k -forms on U by $\Omega^k(U)$.

- For $k = 0$, $\Omega^0(U) = C^\infty(U, \mathbb{R})$ is the space of smooth real-valued functions on U .
- For $k = 1$, $\Omega^1(U)$ is the space of smooth **1-forms**. A 1-form α assigns to each $p \in U$ a covector $\alpha(p) \in V^*$.

Remark 2.43 (Exterior Derivative). *The exterior derivative $d : \Omega^k(U) \rightarrow \Omega^{k+1}(U)$ is defined as usual. For a function $f \in \Omega^0(U)$, its differential $df \in \Omega^1(U)$ is given at $p \in U$ by the linear map $(df)_p : V \rightarrow \mathbb{R}$ defined by $(df)_p(v) = Df(p)(v)$, the directional derivative of f at p in the direction v . In coordinates (x^1, \dots, x^{2n}) corresponding to a basis (e_1, \dots, e_{2n}) of V , $df = \sum_{i=1}^{2n} \frac{\partial f}{\partial x^i} dx^i$. Importantly, $d^2 = d \circ d = 0$.*

The symplectic form ω on V , being non-degenerate, establishes an isomorphism between V and its dual V^* . This contrasts with the isomorphism provided by a Euclidean inner product g .

The musical isomorphism (in the affine symplectic space case) allows us to associate a unique vector field to the differential (a 1-form) of any smooth function.

Earlier, we introduced the Lie algebra. Within the Lie algebra $\mathfrak{X}_\omega(N)$ lies a particularly significant subalgebra consisting of vector fields generated by functions.

Definition 2.44 (Symplectic Gradient). Let $f \in C^\infty(U)$. The **symplectic gradient** (or **Hamiltonian vector field**) of f is the unique vector field $\xi_f \in \mathfrak{X}(U)$ defined by the relation:

$$\iota_{\xi_f} \omega = -df$$

Equivalently, $\omega(\xi_f(p), v) = -(df)_p(v)$ for all $p \in U$ and $v \in V$. In terms of the musical isomorphisms, $\xi_f = -\sharp_\omega(df)$.

Such a function f is called a **Hamiltonian function** or **momentum function** (or simply momentum) generating ξ . The vector field ξ is then uniquely determined by df (and hence by f up to an additive constant) via the non-degeneracy of ω , and is denoted ξ_f . The space of Hamiltonian vector fields is denoted $\mathfrak{X}_{\text{Ham}}(N)$.

Remark 2.45 (Uniqueness of Momentum). *If $\xi = \xi_f$, then $\xi = \xi_{f+c}$ for any constant c (since $d(f+c) = df$). If N is connected, these are the only functions generating ξ_f ; the momentum function is determined up to an overall additive constant. This ambiguity corresponds physically to the fact that only potential energy differences are measurable.*

Remark 2.46 (Sign Convention). *The negative sign is conventional in Hamiltonian mechanics. It ensures that Hamilton's equations take their standard form. If H is the Hamiltonian function, the dynamics are governed by $\dot{x} = \xi_H(x)$. In canonical coordinates (q^i, p_i) where $\omega = \sum_i dq^i \wedge dp_i$, the definition $\iota_{\xi_H} \omega = -dH$ yields the familiar equations $\dot{q}^i = \frac{\partial H}{\partial p_i}$ and $\dot{p}_i = -\frac{\partial H}{\partial q^i}$.*

Example 2.47 (Euclidean Gradient - for contrast). If A were equipped with a Euclidean metric g , the standard **gradient** ∇f would be defined by $g(\nabla f, v) = df(v)$ for all $v \in V$, or $\nabla f = \sharp_g(df)$. Note the absence of the negative sign and the use of g instead of ω .

Vector fields prescribe infinitesimal motion. Integrating this motion yields trajectories and transformations known as flows.

Definition 2.48 (Integral Curve). An **integral curve** of a vector field $\xi \in \mathfrak{X}(U)$ starting at $p \in U$ is a smooth curve $\gamma : I \rightarrow U$, defined on some open interval I containing 0, such that $\gamma(0) = p$ and $\dot{\gamma}(t) = \xi(\gamma(t))$ for all $t \in I$. Here, $\dot{\gamma}(t)$ denotes the tangent vector to the curve at $\gamma(t)$, identified with an element of V .

Theorem 2.49 (Existence and Uniqueness of Local Flows). *Let $\xi \in \mathfrak{X}(U)$. For each $p \in U$, there exists a unique maximal integral curve $\gamma_p : I_p \rightarrow U$ starting at p . Furthermore, there exists an open neighborhood $W \subseteq \mathbb{R} \times U$ of $\{0\} \times U$ and a smooth map $\varphi : W \rightarrow U$, called the **local flow** of ξ , such that for each fixed $p \in U$, the map $t \mapsto \varphi(t, p)$ is the integral curve of ξ starting at p . We often write $\varphi_t(p)$ for $\varphi(t, p)$. For fixed t (where defined), $\varphi_t : U_t \rightarrow U'_t$ is a diffeomorphism between open subsets of U .*

Proof. This is a standard result from the theory of ordinary differential equations, applied to the system $\dot{x} = \xi(x)$. \square

Remark 2.50 (Affine Structure Distortion). *Unless the vector field ξ is constant (corresponding to parallel translation) or linear (corresponding to an affine transformation centered at some origin), its flow maps φ_t are generally not affine automorphisms of A . They typically distort distances, angles, and the affine structure itself.*

A central question is how flows interact with the underlying geometric structure, in our case, the symplectic form ω .

Definition 2.51 (Lie Derivative). Let $\xi \in \mathfrak{X}(U)$. The **Lie derivative** with respect to ξ acts on differential forms. For $\alpha \in \Omega^k(U)$, $\mathcal{L}_\xi \alpha \in \Omega^k(U)$ measures the rate of change of α along the flow of ξ . It can be defined via the flow:

$$\mathcal{L}_\xi \alpha = \left. \frac{d}{dt} \right|_{t=0} \varphi_t^* \alpha$$

where φ_t^* denotes the pullback operation by the diffeomorphism φ_t . Alternatively, Cartan's magic formula⁸ provides the expression:

$$\mathcal{L}_\xi \alpha = d(\iota_\xi \alpha) + \iota_\xi(d\alpha)$$

Continuous symmetries, such as those generated by Lie group actions, are often best understood through their infinitesimal generators – vector fields whose flows consist of symmetries.

Definition 2.52 (Symplectic Vector Field). An **infinitesimal symmetry** of (N, ω) is a smooth vector field $\xi \in \mathfrak{X}(N)$ such that its local flow $\{\varphi_t\}$ consists of (local) symplectomorphisms. This condition is equivalent to requiring that the Lie derivative of ω with respect to ξ vanishes:

$$\mathcal{L}_\xi \omega = 0$$

A vector field satisfying this condition is called **symplectic**. The space of all symplectic vector fields on N is denoted $\mathfrak{X}_\omega(N)$.

The condition $\mathcal{L}_\xi \omega = 0$ provides a powerful link to differential forms. Using Cartan's magic formula

⁸See Theorem 14.35 on page 372 of [20].

$\mathcal{L}_\xi = d\iota_\xi + \iota_\xi d$ and the fact that ω is closed ($d\omega = 0$), we have:

$$\mathcal{L}_\xi \omega = d(\iota_\xi \omega) + \iota_\xi(d\omega) = d(\iota_\xi \omega)$$

Therefore, a vector field ξ is symplectic if and only if the 1-form obtained by contracting ω with ξ is closed:

$$\xi \in \mathfrak{X}_\omega(N) \iff d(\iota_\xi \omega) = 0$$

This shows that the map $\xi \mapsto \iota_\xi \omega$ provides an isomorphism between the space $\mathfrak{X}_\omega(N)$ of symplectic vector fields and the space $Z^1(N)$ of closed 1-forms on N (assuming appropriate function spaces or interpreting the isomorphism via the non-degeneracy of ω).

Proposition 2.53. *The space $\mathfrak{X}_\omega(N)$ is a Lie subalgebra of $(\mathfrak{X}(N), [\cdot, \cdot])$.*

Proof. Let $\xi, \eta \in \mathfrak{X}_\omega(N)$. We must show $[\xi, \eta] \in \mathfrak{X}_\omega(N)$. We use the fundamental identity relating Lie derivatives and commutators: $\mathcal{L}_{[\xi, \eta]} = [\mathcal{L}_\xi, \mathcal{L}_\eta] = \mathcal{L}_\xi \mathcal{L}_\eta - \mathcal{L}_\eta \mathcal{L}_\xi$. Applying this identity to the form ω :

$$\mathcal{L}_{[\xi, \eta]} \omega = \mathcal{L}_\xi(\mathcal{L}_\eta \omega) - \mathcal{L}_\eta(\mathcal{L}_\xi \omega)$$

Since ξ and η are symplectic, $\mathcal{L}_\xi \omega = 0$ and $\mathcal{L}_\eta \omega = 0$. Substituting these into the equation yields:

$$\mathcal{L}_{[\xi, \eta]} \omega = \mathcal{L}_\xi(0) - \mathcal{L}_\eta(0) = 0 - 0 = 0$$

Thus, the commutator $[\xi, \eta]$ is also a symplectic vector field. □

The structure of this potentially infinite-dimensional Lie algebra $\mathfrak{X}_\omega(N)$ is complicated and reflects the geometry of (N, ω) .

The remarkable property of symplectic gradients is that they generate structure-preserving flows.

Theorem 2.54. *Every Hamiltonian vector field is symplectic. That is, if $f \in C^\infty(U)$ and ξ_f is its symplectic gradient, then $\mathcal{L}_{\xi_f} \omega = 0$.*

Proof. We apply Cartan's formula. Since ω is a form of constant coefficients on A (inherited from V), its exterior derivative $d\omega = 0$ (it is closed). By definition of the symplectic gradient, $\iota_{\xi_f} \omega = -df$. Therefore,

$$\mathcal{L}_{\xi_f} \omega = d(\iota_{\xi_f} \omega) + \iota_{\xi_f}(d\omega) = d(-df) + \iota_{\xi_f}(0) = -d^2 f = 0.$$

The result follows. □

Corollary 2.55. *The flow $\{\varphi_t\}$ generated by a Hamiltonian vector field ξ_f consists of **symplectomorphisms**, i.e., diffeomorphisms preserving ω .*

Remark 2.56 (Contrast with Euclidean Gradient). *The flow generated by the Euclidean gradient ∇f does not generally preserve the metric g . The condition for a vector field ξ to preserve g is $\mathcal{L}_\xi g = 0$, leading to Killing vector fields, which have a very different character from symplectic vector fields. Gradient flows typically move points towards regions of lower function value, orthogonal to level sets, while Hamiltonian flows move points along level sets of the Hamiltonian function (in many standard examples) in a manner that preserves symplectic area.*

Proposition 2.57 (Conservation of Energy). *The Hamiltonian function f is constant along the integral curves of its associated Hamiltonian vector field ξ_f . That is, $\mathcal{L}_{\xi_f} f = 0$.*

Proof. Using the definition of the Lie derivative for functions ($\mathcal{L}_\xi f = df(\xi)$) and the definition of ξ_f :

$$\mathcal{L}_{\xi_f} f = df(\xi_f) = (-\iota_{\xi_f} \omega)(\xi_f) = -\omega(\xi_f, \xi_f).$$

Since ω is skew-symmetric, $\omega(v, v) = 0$ for any vector v . Thus, $\mathcal{L}_{\xi_f} f = 0$. \square

This proposition is the geometric statement of energy conservation in Hamiltonian mechanics.

Earlier, we saw that every Hamiltonian vector field is symplectic. The converse holds if the domain U has trivial first de Rham cohomology, $H_{dR}^1(U) = 0$ (e.g., if U is star-shaped). On topologically non-trivial domains, there can exist symplectic vector fields that are not Hamiltonian.

Example 2.58 (Symplectic but not Hamiltonian Vector Field on T^2). The distinction between symplectic and Hamiltonian vector fields is not merely formal; it manifests concretely on manifolds with non-trivial first cohomology. Consider the 2-torus $M = T^2 = \mathbb{R}^2 / (2\pi\mathbb{Z})^2$, with angular coordinates (θ_1, θ_2) . Let the symplectic form be the standard area form $\omega = d\theta_1 \wedge d\theta_2$. This form is clearly closed ($d\omega = 0$) and non-degenerate.

Consider the vector field $\xi = \frac{\partial}{\partial \theta_1}$. This vector field generates the flow $\varphi_t(\theta_1, \theta_2) = (\theta_1 + t \pmod{2\pi}, \theta_2)$, representing rigid rotation along the first circle factor.

Is ξ symplectic? We compute the 1-form $\alpha_\xi = \iota_\xi \omega$:

$$\begin{aligned} \alpha_\xi &= \iota_{\partial/\partial \theta_1} (d\theta_1 \wedge d\theta_2) = (\iota_{\partial/\partial \theta_1} d\theta_1) \wedge d\theta_2 - d\theta_1 \wedge (\iota_{\partial/\partial \theta_1} d\theta_2) \\ &= (d\theta_1(\frac{\partial}{\partial \theta_1}))d\theta_2 - d\theta_1(d\theta_2(\frac{\partial}{\partial \theta_1})) \\ &= (1)d\theta_2 - d\theta_1(0) = d\theta_2. \end{aligned}$$

Since $d\alpha_\xi = d(d\theta_2) = 0$, the 1-form α_ξ is closed. Therefore, by definition, the vector field $\xi = \frac{\partial}{\partial \theta_1}$ is symplectic ($\mathcal{L}_\xi \omega = d(\iota_\xi \omega) = 0$).

Is ξ Hamiltonian? For ξ to be Hamiltonian, the 1-form $\alpha_\xi = d\theta_2$ would need to be exact. That is, there would have to exist a globally defined smooth function $f \in C^\infty(T^2)$ such that $\alpha_\xi = d\theta_2 = -df$. However, $d\theta_2$ is not exact on the torus T^2 . We can see this by integrating $d\theta_2$ over the closed loop $\gamma : [0, 2\pi] \rightarrow T^2$ defined by $\gamma(t) = (0, t)$:

$$\int_\gamma d\theta_2 = \int_0^{2\pi} (d\theta_2)_{\gamma(t)}(\dot{\gamma}(t)) dt$$

Since $\dot{\gamma}(t) = (0, 1)$ in velocity coordinates, corresponding to the vector $\frac{\partial}{\partial \theta_2}$ at $\gamma(t)$, the integrand is $(d\theta_2)(\frac{\partial}{\partial \theta_2}) = 1$. Thus,

$$\int_\gamma d\theta_2 = \int_0^{2\pi} 1 dt = 2\pi \neq 0.$$

Since the integral of $d\theta_2$ over a closed loop is non-zero, $d\theta_2$ cannot be exact (by Stokes' Theorem, or the definition of exactness via path integrals). Therefore, $\xi = \frac{\partial}{\partial \theta_1}$ is a symplectic vector field that is not Hamiltonian.

This aligns perfectly with the fundamental exact sequence. The first de Rham cohomology of the torus is $H_{dR}^1(T^2) \cong \mathbb{R}^2$, with basis classes $[d\theta_1]$ and $[d\theta_2]$. The map $\Phi : \mathfrak{X}_\omega(T^2) \rightarrow H_{dR}^1(T^2)$ sends our vector field $\xi = \frac{\partial}{\partial \theta_1}$ to the cohomology class $\Phi(\xi) = [\iota_\xi \omega] = [d\theta_2]$. Since $[d\theta_2]$ is a non-zero element of $H_{dR}^1(T^2)$, ξ lies in $\mathfrak{X}_\omega(T^2)$ but not in the kernel of Φ , which is $\mathfrak{X}_{\text{Ham}}(T^2)$. Similarly, the

vector field $\frac{\partial}{\partial \theta_2}$ is symplectic but not Hamiltonian, mapping via Φ to the class $[\iota_{\partial/\partial \theta_2} \omega] = [-d\theta_1]$, which is also non-zero. This example illustrates how the topology of the manifold ($H^1 \neq 0$) creates a distinction between the broader class of structure-preserving vector fields (symplectic) and the subclass generated by scalar functions (Hamiltonian).

Now we discuss the uniqueness of the Hamiltonian.

Remark 2.59 (Uniqueness of Hamiltonian). *If $\xi = \xi_f$ for some function f , then $\xi_f = \xi_{f+c}$ for any constant c , since $d(f+c) = df$. If U is connected, then any two Hamiltonian functions for the same Hamiltonian vector field differ by a constant.*

To summarize: the structure of an affine symplectic space provides the local model for symplectic manifolds. The concepts introduced here (Hamiltonian vector fields generated by functions via the symplectic form, and their flows which preserve this form) are fundamental. They form the bedrock of Hamiltonian mechanics, geometric quantization, and significant parts of modern symplectic topology. The key thing to remember is the relationship between functions (observables), vector fields (generators of dynamics), and the preservation of the symplectic structure (ω) under the resulting flows. This contrasts with Riemannian geometry, where gradients are tied to the metric and their flows typically do not preserve it. The conservation laws, particularly the preservation of ω and the constancy of the Hamiltonian function along its flow, are results of the symplectic setup.

2.3 Riemannian and Symplectic Manifolds

Our investigations so far within the framework of affine spaces have laid bare the essential algebraic structures underpinning Riemannian and symplectic geometry. While this affine setting provides the infinitesimal model, it falls short of capturing the complexities of global phenomena. Physical systems often evolve on spaces possessing non-trivial topology (e.g., configuration spaces involving angles) or inherent curvature, necessitating a transition to the broader landscape of smooth manifolds. Here, we undertake this transition, exploring how the local geometric structures inherited from V are woven into global tensor fields on a manifold M .

Let M be a smooth manifold of dimension n (which we require to be Hausdorff and second-countable). Recall that its smooth structure is defined by a maximal atlas $\mathcal{A} = \{(U_\alpha, \phi_\alpha)\}_{\alpha \in I}$, where $\{U_\alpha\}$ is an open cover of M and each chart map $\phi_\alpha : U_\alpha \rightarrow \phi_\alpha(U_\alpha) \subseteq \mathbb{R}^n$ is a homeomorphism onto an open subset of Euclidean space. Importantly, the transition maps $\phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(U_\alpha \cap U_\beta) \rightarrow \phi_\beta(U_\alpha \cap U_\beta)$ are required to be smooth (class C^∞) for all overlapping charts U_α, U_β . This compatibility condition ensures that notions of smoothness are well-defined globally on M .

The tangent bundle $\pi : TM \rightarrow M$ provides the natural arena for infinitesimal calculus on M . It can be constructed as the disjoint union $\coprod_{m \in M} T_m M$ quotiented by an equivalence relation identifying tangent vectors via the differentials of chart maps, or intrinsically via equivalence classes of curves or the space of derivations on the algebra of germs of smooth functions at each point. TM possesses the structure of a smooth vector bundle of rank n over M ; this means TM is itself a smooth manifold, π is a smooth map, and TM admits local trivializations $\Phi_\alpha : \pi^{-1}(U_\alpha) \rightarrow U_\alpha \times \mathbb{R}^n$ which are diffeomorphisms respecting the projection maps and restricting to linear isomorphisms on the fibers $\Phi_\alpha|_m : T_m M \rightarrow \{m\} \times \mathbb{R}^n \cong \mathbb{R}^n$.

Our objective is to endow M with global geometric structures that smoothly assign to each tangent

space $T_m M$ either an inner product (leading to Riemannian geometry) or a symplectic form (leading to symplectic geometry), mirroring the structures previously studied on the model vector space V .

2.3.1 Pointwise Structures, Fiber Bundles, and Global Tensor Fields

The path to global structures proceeds by first considering the desired algebraic structure on each fiber $T_m M$, then assembling these into a fiber bundle, and finally defining the global geometric structure as a smooth section of this bundle.

We have already defined an inner product and a symplectic form. Let's place these structures on the tangent space in the obvious way.

Definition 2.60. Let M be a smooth manifold.

1. A **Riemannian metric** on M is a smooth assignment of an inner product g_m on each tangent space $T_m M$, such that the map $g : TM \times_M TM \rightarrow \mathbb{R}$ is a smooth $(0, 2)$ -tensor field. That is, for each $m \in M$, the bilinear form $g_m : T_m M \times T_m M \rightarrow \mathbb{R}$ is symmetric and positive-definite.
2. If $\dim M = 2k$ is even, an **almost-symplectic form** on M is a smooth 2-form $\omega \in \Omega^2(M)$ such that ω_m is non-degenerate at each point $m \in M$. That is, for each $m \in M$, the bilinear form $\omega_m : T_m M \times T_m M \rightarrow \mathbb{R}$ is skew-symmetric and non-degenerate.

We can consider the spaces of such structures on a fixed vector space $V \cong \mathbb{R}^n$. As before, let $\text{Met}(V)$ denote the space of inner products on V , and $\text{Symp}(V)$ denote the space of symplectic forms on V (if $n = 2k$). These pointwise spaces can be bundled together over M .

Definition 2.61 (Associated Fiber Bundles). Associated to the tangent bundle $TM \rightarrow M$, we define:

1. The **bundle of metrics** $\text{Met}(M) \rightarrow M$, whose fiber over $m \in M$ is $\text{Met}(T_m M)$.
2. The **bundle of symplectic forms** $\text{Symp}(M) \rightarrow M$ (if $\dim M = 2k$), whose fiber over $m \in M$ is $\text{Symp}(T_m M)$.

These are smooth fiber bundles whose structure group is related to $\text{GL}(n, \mathbb{R})$ acting on the fibers $\text{Met}(\mathbb{R}^n)$ and $\text{Symp}(\mathbb{R}^n)$ respectively.

A global geometric structure is then precisely a choice of an element in the fiber over each point m , varying smoothly across M .

Definition 2.62 (Global Structures as Sections). A **Riemannian metric** on M is a smooth section $g : M \rightarrow \text{Met}(M)$ of the bundle $\rho_1 : \text{Met}(M) \rightarrow M$. Equivalently, it is a smooth tensor field $g \in \Gamma(T^*M \otimes T^*M)$ which is symmetric and positive-definite at each point $m \in M$. In local coordinates (x^1, \dots, x^n) , $g = \sum_{i,j=1}^n g_{ij}(x) dx^i \otimes dx^j$, where $(g_{ij}(x))$ is a smooth matrix-valued function such that each matrix $g(x)$ is symmetric and positive-definite.

An **almost symplectic structure** on M (which must be even-dimensional, $\dim M = 2k$) is a smooth section $\omega : M \rightarrow \text{Symp}(M)$ of the bundle $\rho_2 : \text{Symp}(M) \rightarrow M$. Equivalently, it is a smooth differential 2-form $\omega \in \Omega^2(M)$ which is non-degenerate at each point $m \in M$. In local coordinates (x^1, \dots, x^{2k}) , $\omega = \sum_{1 \leq i < j \leq 2k} \omega_{ij}(x) dx^i \wedge dx^j$, where the matrix $(\omega_{ij}(x))$ (with $\omega_{ii} = 0, \omega_{ji} = -\omega_{ij}$) is smooth and has non-zero determinant at each point x .

Remark 2.63 (Geometric Interpretation). *Infinitesimally, a Riemannian metric g_m allows us to measure lengths of tangent vectors ($\|v\|_m = \sqrt{g_m(v, v)}$) and angles between them ($\cos \theta = g_m(v, w) / (\|v\|_m \|w\|_m)$). An almost symplectic form ω_m allows us to measure the signed area of the parallelogram spanned by two tangent vectors v, w as $\omega_m(v, w)$. The non-degeneracy ensures that for any non-zero v , there exists some w such that this area is non-zero.*

2.3.2 Existence of Structures

The question of existence of these global structures reveals a big difference between Riemannian and symplectic geometry. The former demonstrates remarkable flexibility, while the latter encounters significant topological obstructions.

Theorem 2.64 (Universal Existence of Riemannian Metrics). *Every smooth manifold M admits a Riemannian metric.*

Detailed Sketch. The proof is based on the ability to patch together locally defined structures using partitions of unity, guaranteed by the paracompactness of M .

Local Existence: Let $\{(U_\alpha, \phi_\alpha)\}_{\alpha \in I}$ be an atlas for M . On each chart domain U_α , we can define a local Riemannian metric g_α . A standard choice is to pull back the Euclidean metric $g_{\text{std}} = \sum_{i=1}^n (dy^i)^2$ from the target space $\phi_\alpha(U_\alpha) \subseteq \mathbb{R}^n$: define $g_\alpha = \phi_\alpha^* g_{\text{std}}$. This yields a smooth, symmetric, positive-definite tensor field on U_α .

Partition of Unity: Since M is paracompact and Hausdorff, there exists a partition of unity subordinate to the open cover $\{U_\alpha\}$. This is a collection of smooth functions $\{\psi_\alpha : M \rightarrow [0, 1]\}_{\alpha \in I}$ such that:

- The support $\text{supp}(\psi_\alpha) = \overline{\{m \in M \mid \psi_\alpha(m) \neq 0\}}$ is compact and contained in U_α for each α .
- The collection of supports is locally finite (every point $m \in M$ has a neighborhood intersecting only finitely many supports).
- For every $m \in M$, $\sum_{\alpha \in I} \psi_\alpha(m) = 1$. (Note: the sum is finite due to local finiteness).

Gluing: Define a global symmetric tensor field g of type $(0, 2)$ by the formula:

$$g_m = \sum_{\alpha \in I} \psi_\alpha(m) (g_\alpha)_m$$

for each $m \in M$. Note that $(g_\alpha)_m$ is only defined if $m \in U_\alpha$; however, if $m \notin \text{supp}(\psi_\alpha)$, then $\psi_\alpha(m) = 0$, so the term vanishes. Thus, the sum is well-defined and involves only finitely many non-zero terms in a neighborhood of any m . Smoothness follows from the smoothness of ψ_α and g_α . Symmetry is clear as each g_α is symmetric.

Positive-Definiteness: For any $m \in M$ and any non-zero $v \in T_m M$, we need to show $g_m(v, v) > 0$. We have $g_m(v, v) = \sum_{\alpha \in I} \psi_\alpha(m) (g_\alpha)_m(v, v)$. Since $\sum \psi_\alpha(m) = 1$ and $\psi_\alpha(m) \geq 0$, there must be at least one index β such that $\psi_\beta(m) > 0$. For this β , $m \in \text{supp}(\psi_\beta) \subset U_\beta$, so $(g_\beta)_m$ is defined. Since $(g_\beta)_m$ is positive-definite and $v \neq 0$, we have $(g_\beta)_m(v, v) > 0$. Also, for all α where $\psi_\alpha(m) \neq 0$, we have $m \in U_\alpha$, so $(g_\alpha)_m(v, v) \geq 0$. Therefore, $g_m(v, v) = \sum \psi_\alpha(m) (g_\alpha)_m(v, v) \geq \psi_\beta(m) (g_\beta)_m(v, v) > 0$.

□

The situation for almost symplectic structures contrasts sharply with that of Riemannian metrics. While every smooth manifold admits a Riemannian metric (thanks to the convexity of the space of positive-definite inner products and the existence of partitions of unity), the existence of an almost symplectic form is not guaranteed—even on even-dimensional manifolds. This is because non-degeneracy is not a convex condition: convex combinations of non-degenerate 2-forms need not be non-degenerate. In contrast, positive-definiteness (a condition on Riemannian metrics) is convex, allowing smooth patching via partitions of unity.

Proposition 2.65 (Obstructions to Almost Symplectic Structures). *Let M be a smooth manifold of dimension $2k$. Necessary conditions for the existence of an almost symplectic structure ω include:*

1. **Even Dimension:** *Trivial by definition.*
2. **Orientability:** *An almost symplectic form ω determines a volume form $\Omega = \frac{(-1)^{k(k-1)/2}}{k!} \omega^k = \frac{(-1)^{k(k-1)/2}}{k!} \omega \wedge \cdots \wedge \omega$. Since ω is non-degenerate, Ω is nowhere vanishing. The existence of a nowhere-vanishing top-degree form is equivalent to the orientability of M . Thus, any almost symplectic manifold must be orientable.*
3. **Higher-Order Topological Obstructions:** *Even if M is orientable, further obstructions can exist, often related to characteristic classes of the tangent bundle TM . For example, the existence of an almost symplectic structure implies certain relations among Stiefel-Whitney classes and Pontryagin classes.*

Example 2.66 (Obstructions Illustrated).

- \mathbb{RP}^2 (Real Projective Plane): As a non-orientable 2-manifold, it cannot admit an almost symplectic structure due to obstruction (2). Topologically, $T(\mathbb{RP}^2)$ is non-trivial in a way that prevents a global non-degenerate skew form (related to $w_1(T\mathbb{RP}^2) \neq 0$).
- Orientable Surfaces (M^2): Every orientable surface admits an almost symplectic structure. As noted earlier, any area form suffices. For example, the sphere S^2 with its standard volume form is almost symplectic (and in fact, symplectic). The torus $T^2 = S^1 \times S^1$ with $d\theta_1 \wedge d\theta_2$ is another example.
- The 4-sphere S^4 : This is orientable ($H_4(S^4, \mathbb{Z}) \cong \mathbb{Z}$). However, S^4 famously does *not* admit an almost symplectic structure. This is a deeper result. It is related to the fact that TS^4 does not admit an *almost complex structure* (an endomorphism $J : TS^4 \rightarrow TS^4$ with $J^2 = -\text{id}$). While almost symplectic structures do not necessarily require almost complex structures, their obstructions are linked via characteristic classes. For spheres, only S^2 and S^6 admit almost complex structures (related to the existence of complex and octonion multiplications). The non-existence for S^4 can be shown using arguments involving Pontryagin classes or K-theory, demonstrating that topology beyond simple orientability plays a critical role. The argument using $H^2(S^4) = 0$, presented later, applies specifically to ruling out *closed* forms (symplectic), not just non-degenerate ones (almost symplectic).
- Complex Projective Space \mathbb{CP}^n : This space, of real dimension $2n$, *does* admit an almost symplectic structure. In fact, it admits a Kähler structure, which includes a symplectic form (the Fubini-Study form ω_{FS}) compatible with a complex structure J and a Riemannian metric g . The existence here highlights that suitable complex geometry often provides a route to constructing (almost) symplectic structures.

2.3.3 The Symplectic Condition: Integrability

The definition of an almost symplectic structure involves only the algebraic condition of non-degeneracy at each point, varying smoothly. Symplectic geometry properly imposes an additional, important differential condition: the closure of the 2-form ω .

Definition 2.67 (Symplectic Manifold). A **symplectic manifold** is a pair (M, ω) where M is a smooth manifold (necessarily even-dimensional and orientable) and $\omega \in \Omega^2(M)$ is a differential 2-form that is both:

1. **Non-degenerate** (pointwise): $\omega_m : T_m M \times T_m M \rightarrow \mathbb{R}$ is non-degenerate for all $m \in M$.
2. **Closed**: $d\omega = 0$.

The 2-form ω satisfying these conditions is called the **symplectic form** of M .

The condition $d\omega = 0$ fundamentally distinguishes symplectic geometry from Riemannian geometry. It acts as an *integrability condition*, ensuring a high degree of local uniformity, as captured by Darboux's Theorem. The closedness condition $d\omega = 0$ establishes a direct link to the topology of the manifold via de Rham cohomology.

Remark 2.68 (Cohomological Obstructions and Properties). *The condition $d\omega = 0$ means ω represents a cohomology class $[\omega] \in H_{dR}^2(M)$. This class, and its powers, carry significant information.*

- **Fundamental Class**: $[\omega]$ is a fundamental topological invariant of the symplectic structure. Symplectic diffeomorphisms must preserve this class.
- **Volume and Top Class**: If M is compact with $\dim M = 2k$, the non-degeneracy ensures the volume form $\omega^k/k!$ is nowhere zero. Since $d\omega = 0$, the volume form is also closed: $d(\omega^k) = k\omega^{k-1} \wedge d\omega = 0$. Its integral $\text{Vol}(M) = \int_M \omega^k/k!$ is non-zero. This implies the top cohomology class $[\omega^k] = [\omega]^k \in H_{dR}^{2k}(M)$ must be non-zero (as it pairs non-trivially with the fundamental homology class $[M]$).
- **Obstruction via H^2** : If $H_{dR}^2(M) = 0$, then any closed 2-form ω must be exact ($\omega = d\alpha$). As explored below for S^4 , this can prevent the existence of a symplectic structure on a compact manifold, as $[\omega]^k$ would necessarily be zero in cohomology if $[\omega] = 0$.
- **Hard Lefschetz Property**: On a compact Kähler manifold (a special class of symplectic manifolds admitting a compatible complex structure and Riemannian metric), the map $L : H^p(M) \rightarrow H^{p+2}(M)$ given by cup product with $[\omega]$, $L(\beta) = [\omega] \wedge \beta$, satisfies the Hard Lefschetz theorem: $L^{k-p} : H^p(M) \rightarrow H^{2k-p}(M)$ is an isomorphism for $p \leq k$. This property, while specific to Kähler geometry, highlights the deep structural role played by the symplectic class $[\omega]$. Symplectic manifolds satisfying Hard Lefschetz are often called Lefschetz manifolds.

Example 2.69 (S^4 revisited: Symplectic Obstruction via H^2). As established, S^4 does not even admit an almost symplectic structure. However, we can provide an independent argument showing it does not admit a symplectic structure using cohomology, which relies only on $d\omega = 0$ and non-degeneracy.

Suppose $\omega \in \Omega^2(S^4)$ were a symplectic form. It must be closed ($d\omega = 0$). The second de Rham cohomology of S^4 is trivial:

$$H_{dR}^2(S^4) = \{0\}.$$

By definition of cohomology, any closed form on a space with trivial cohomology must be exact. Thus,

$$\omega = d\alpha \quad \text{for some } \alpha \in \Omega^1(S^4).$$

Now consider the 4-form $\omega^2 = \omega \wedge \omega$. Since $\omega = d\alpha$,

$$\omega^2 = d\alpha \wedge d\alpha.$$

Using the property

$$d(A \wedge B) = dA \wedge B + (-1)^{\deg A} A \wedge dB,$$

we find

$$d(\alpha \wedge d\alpha) = d\alpha \wedge d\alpha - \alpha \wedge d^2\alpha = d\alpha \wedge d\alpha = \omega^2.$$

Thus, $\omega^2 = d(\alpha \wedge \omega)$ is an exact 4-form.

By Stokes' Theorem, the integral of an exact form over a compact manifold without boundary is zero:

$$\int_{S^4} \omega^2 = \int_{\partial S^4} (\alpha \wedge \omega) = 0.$$

However, for ω to be symplectic on the compact 4-manifold S^4 , ω^2 must be a volume form, and its integral

$$\int_{S^4} \omega^2$$

must be non-zero (twice the symplectic volume). This contradiction demonstrates that S^4 cannot admit a symplectic structure.

Compare this to the 4-torus T^4 , where

$$H_{\text{dR}}^2(T^4) \cong \mathbb{R}^6 \neq \{0\},$$

allowing for non-exact closed 2-forms like

$$d\theta_1 \wedge d\theta_2 + d\theta_3 \wedge d\theta_4,$$

which is a symplectic form.

So far, we've encountered several examples of manifolds that are not symplectic. Naturally, it's time to explore some sources of symplectic manifolds as well.

Example 2.70 (Sources of Symplectic Manifolds).

- **Orientable Surfaces (M^2):** Any area form ω is non-degenerate and automatically closed ($d\omega \in \Omega^3(M^2) = 0$). Examples include $(S^2, \omega_{\text{std}})$ and $(T^2, d\theta_1 \wedge d\theta_2)$.
- **Cotangent Bundles (T^*M):** As detailed below, these carry the canonical symplectic form $\omega_{\text{can}} = -d\theta$, providing phase spaces for mechanics.
- **Kähler Manifolds:** These are complex manifolds (M, J) equipped with a Riemannian metric g such that the associated Kähler form $\omega(X, Y) = g(JX, Y)$ is closed ($d\omega = 0$) and non-degenerate (which follows from g and J). Complex projective spaces $\mathbb{C}P^n$ with the Fubini-Study metric are prime examples. Kähler geometry lies at the intersection of Riemannian, complex, and symplectic geometry.

- **Coadjoint Orbits:** For a Lie group G with Lie algebra \mathfrak{g} , the coadjoint orbits in the dual space \mathfrak{g}^* carry canonical symplectic structures (the Kirillov-Kostant-Souriau form). These are important in representation theory and geometric quantization.

2.3.4 Canonical Example: The Cotangent Bundle T^*M

Among the sources of symplectic manifolds, the cotangent bundle holds a privileged position due to its universality and its role as the natural phase space in Hamiltonian mechanics.

Let M be any smooth manifold of dimension n . The cotangent bundle $\pi : T^*M \rightarrow M$ is the vector bundle whose fiber over $m \in M$ is the dual space $T_m^*M = (T_m M)^*$, the space of linear functionals (covectors) on the tangent space $T_m M$. A point $p \in T^*M$ can be thought of as a pair (m, α) where $m = \pi(p) \in M$ and $\alpha \in T_m^*M$.

Definition 2.71 (Tautological One-Form). The **tautological one-form** (or Liouville form) $\theta \in \Omega^1(T^*M)$ is defined as follows: For any point $p \in T^*M$ and any tangent vector $\eta \in T_p(T^*M)$ to the cotangent bundle at p ,

$$\theta_p(\eta) := \langle \alpha, \pi_* \eta \rangle = \alpha(\pi_* \eta)$$

where $\alpha \in T_{\pi(p)}^*M$ is the covector component represented by the point p , and $\pi_* : T_p(T^*M) \rightarrow T_{\pi(p)}M$ is the pushforward (differential) of the projection map π . The name “tautological” reflects the definition: θ at $p = (m, \alpha)$ evaluates the covector α on the projection (to the base manifold M) of tangent vectors η starting at p . It naturally extracts the “covector information” inherent in the point $p \in T^*M$ and applies it to the “base component” of tangent vectors at p .

To understand θ in coordinates, let (q^1, \dots, q^n) be local coordinates on an open set $U \subseteq M$. These induce canonical coordinates $(q^1, \dots, q^n, p_1, \dots, p_n)$ on the corresponding domain $T^*U = \pi^{-1}(U) \subseteq T^*M$. A point $z \in T^*U$ has coordinates $(q(z), p(z))$, representing the covector $\alpha = \sum_{i=1}^n p_i(z)(dq^i)_{q(z)} \in T_{q(z)}^*M$. Let $\eta \in T_z(T^*M)$. We can write $\eta = \sum a^j (\frac{\partial}{\partial q^j})_z + \sum b_j (\frac{\partial}{\partial p_j})_z$. The projection map is $\pi(q, p) = q$. Its differential is $\pi_*(\frac{\partial}{\partial q^j}) = \frac{\partial}{\partial q^j}$ (viewed as a vector at q) and $\pi_*(\frac{\partial}{\partial p_j}) = 0$. Thus, $\pi_* \eta = \sum a^j (\frac{\partial}{\partial q^j})_{q(z)}$. Applying the definition of θ :

$$\theta_z(\eta) = \alpha(\pi_* \eta) = \left(\sum_i p_i(z) dq^i \right) \left(\sum_j a^j \frac{\partial}{\partial q^j} \right) = \sum_i p_i(z) a^i.$$

Now consider the coordinate expression for $\theta = \sum A_i dq^i + \sum B^j dp_j$. Evaluating this on η : $\theta_z(\eta) = \sum A_i a^i + \sum B^j b_j$. Comparing, we must have $A_i = p_i(z)$ and $B^j = 0$. Therefore, in these canonical coordinates:

$$\theta = \sum_{i=1}^n p_i dq^i$$

Definition 2.72 (Canonical Symplectic Form on T^*M). The **canonical symplectic form** on the cotangent bundle T^*M is the 2-form $\omega_{\text{can}} \in \Omega^2(T^*M)$ defined by

$$\omega_{\text{can}} := -d\theta$$

The choice of sign is conventional in physics and ensures that Hamilton’s equations take their standard form when using ω_{can} and coordinates (q, p) . With this sign, $\omega_{\text{can}} = -d(\sum p_i dq^i) = -\sum dp_i \wedge dq^i = \sum dq^i \wedge dp_i$.

Theorem 2.73. *For any smooth manifold M , the cotangent bundle T^*M equipped with the canonical form $\omega_{\text{can}} = -d\theta$ is a symplectic manifold.*

Proof. Closedness: $\omega_{\text{can}} = -d\theta$. Therefore, $d\omega_{\text{can}} = -d(d\theta) = -d^2\theta = 0$, since the exterior derivative squares to zero. Closedness is automatic from the definition as an exact form (up to sign).

Non-degeneracy: We check this in the canonical local coordinates $(q^1, \dots, q^n, p_1, \dots, p_n)$ derived above. In these coordinates, $\omega_{\text{can}} = \sum_{i=1}^n dq^i \wedge dp_i$. The matrix representation of this bilinear form with respect to the basis $\{\frac{\partial}{\partial q^1}, \dots, \frac{\partial}{\partial q^n}, \frac{\partial}{\partial p_1}, \dots, \frac{\partial}{\partial p_n}\}$ for the tangent space $T_z(T^*M)$ is the $2n \times 2n$ block matrix:

$$J = \begin{pmatrix} 0_n & I_n \\ -I_n & 0_n \end{pmatrix}$$

where 0_n is the $n \times n$ zero matrix and I_n is the $n \times n$ identity matrix. The determinant of J is $\det(I_n) = 1$, which is non-zero. Since the determinant is non-zero, the bilinear form is non-degenerate. As this holds in local coordinates covering T^*M , ω_{can} is non-degenerate everywhere. Thus, $(T^*M, \omega_{\text{can}})$ is a symplectic manifold. \square

The existence of this canonical structure makes T^*M the prototypical example of a symplectic manifold and the natural setting for Hamiltonian mechanics, where q^i are generalized positions and p_i are conjugate momenta.

2.3.5 Dynamics on Symplectic Manifolds: Hamiltonian Formalism

The framework of vector fields, flows, and gradients developed in the affine setting finds its natural generalization and richest expression on symplectic manifolds.

Definition 2.74 (Symplectic and Hamiltonian Vector Fields on Manifolds). Let (M, ω) be a symplectic manifold.

1. A vector field $\xi \in \mathfrak{X}(M)$ is **symplectic** if it preserves the symplectic form under its flow, i.e., $\mathcal{L}_\xi \omega = 0$. As $d\omega = 0$, Cartan's formula $\mathcal{L}_\xi \omega = d(\iota_\xi \omega) + \iota_\xi(d\omega)$ reduces this condition to $d(\iota_\xi \omega) = 0$. Thus, ξ is symplectic if and only if the 1-form $\alpha_\xi := \iota_\xi \omega$ is closed. The set of all symplectic vector fields forms a Lie subalgebra $\mathfrak{X}_\omega(M) \subset \mathfrak{X}(M)$ under the standard commutator bracket.
2. A vector field $\xi \in \mathfrak{X}(M)$ is **Hamiltonian** if the closed 1-form $\alpha_\xi = \iota_\xi \omega$ is furthermore exact. That is, if there exists a smooth function $f \in C^\infty(M)$ (the **Hamiltonian function** or simply **Hamiltonian**) such that $\alpha_\xi = \iota_\xi \omega = -df$. Such a vector field is uniquely determined by f (up to addition of a locally constant function to f) and is denoted ξ_f . The space of Hamiltonian vector fields $\mathfrak{X}_{\text{Ham}}(M)$ is also a Lie subalgebra of $\mathfrak{X}(M)$.

Clearly, every Hamiltonian vector field is symplectic ($\iota_{\xi_f} \omega = -df \implies d(\iota_{\xi_f} \omega) = -d^2f = 0$). The converse depends on the topology of M , specifically its first de Rham cohomology group $H_{\text{dR}}^1(M)$.

Remark 2.75 (The Fundamental Exact Sequence). *The relationship between functions, Hamiltonian fields, symplectic fields, and topology is captured by the following sequence of vector spaces and linear maps:*

$$0 \longrightarrow \mathbb{R} \longrightarrow C_{\text{loc. const}}^\infty(M) \xrightarrow{i} C^\infty(M) \xrightarrow{\delta} \mathfrak{X}_{\text{Ham}}(M) \xrightarrow{j} \mathfrak{X}_\omega(M) \xrightarrow{\Phi} H_{\text{dR}}^1(M)$$

Here:

- $C_{\text{loc. const}}^\infty(M)$ are the locally constant functions (isomorphic to \mathbb{R} if M is connected). The map i is inclusion. The $0 \rightarrow \cdots \rightarrow C^\infty(M)$ part shows constants map to zero under δ .
- $\delta : C^\infty(M) \rightarrow \mathfrak{X}_{\text{Ham}}(M)$ is the map $f \mapsto \xi_f$, where ξ_f is defined by $\iota_{\xi_f}\omega = -df$. The kernel of δ consists precisely of the locally constant functions (since $df = 0$ if and only if f is locally constant).
- $j : \mathfrak{X}_{\text{Ham}}(M) \rightarrow \mathfrak{X}_\omega(M)$ is the inclusion map, as every Hamiltonian field is symplectic.
- $\Phi : \mathfrak{X}_\omega(M) \rightarrow H_{\text{dR}}^1(M)$ maps a symplectic vector field ξ to the de Rham cohomology class of the closed 1-form $\alpha_\xi = \iota_\xi\omega$, i.e., $\Phi(\xi) = [\iota_\xi\omega]$.

The sequence is exact at $C^\infty(M)$ ($\ker \delta = \text{im } i$) and at $\mathfrak{X}_{\text{Ham}}(M)$ ($\ker j = 0$, image of δ is $\mathfrak{X}_{\text{Ham}}(M)$). Exactness at $\mathfrak{X}_\omega(M)$ means $\ker \Phi = \text{im } j = \mathfrak{X}_{\text{Ham}}(M)$. This precisely states that a symplectic vector field ξ is Hamiltonian if and only if the cohomology class $[\iota_\xi\omega]$ is zero, i.e., if $\iota_\xi\omega$ is exact. The map Φ is sometimes related to the flux homomorphism. The sequence clarifies that $H_{\text{dR}}^1(M)$ measures the obstruction for a symplectic vector field to be Hamiltonian. If $H_{\text{dR}}^1(M) = 0$ (e.g., if M is simply connected), then every symplectic vector field is Hamiltonian. Furthermore, $\mathfrak{X}_{\text{Ham}}(M)$ is known to be an ideal in $\mathfrak{X}_\omega(M)$ under suitable conditions (e.g., M compact).

Example 2.76 (Symplectic but not Hamiltonian Vector Field on T^2). The distinction between symplectic and Hamiltonian vector fields is not merely formal; it manifests concretely on manifolds with non-trivial first cohomology. Consider the 2-torus $M = T^2 = \mathbb{R}^2/(2\pi\mathbb{Z})^2$, with angular coordinates (θ_1, θ_2) . Let the symplectic form be the standard area form $\omega = d\theta_1 \wedge d\theta_2$. This form is clearly closed ($d\omega = 0$) and non-degenerate.

Consider the vector field $\xi = \frac{\partial}{\partial \theta_1}$. This vector field generates the flow $\varphi_t(\theta_1, \theta_2) = (\theta_1 + t \pmod{2\pi}, \theta_2)$, representing rigid rotation along the first circle factor.

Is ξ symplectic? We compute the 1-form $\alpha_\xi = \iota_\xi\omega$:

$$\begin{aligned} \alpha_\xi &= \iota_{\partial/\partial \theta_1}(d\theta_1 \wedge d\theta_2) = (\iota_{\partial/\partial \theta_1}d\theta_1) \wedge d\theta_2 - d\theta_1 \wedge (\iota_{\partial/\partial \theta_1}d\theta_2) \\ &= (d\theta_1(\frac{\partial}{\partial \theta_1}))d\theta_2 - d\theta_1(d\theta_2(\frac{\partial}{\partial \theta_1})) \\ &= (1)d\theta_2 - d\theta_1(0) = d\theta_2. \end{aligned}$$

Since $d\alpha_\xi = d(d\theta_2) = 0$, the 1-form α_ξ is closed. Therefore, by definition, the vector field $\xi = \frac{\partial}{\partial \theta_1}$ is symplectic ($\mathcal{L}_\xi\omega = d(\iota_\xi\omega) = 0$).

Is ξ Hamiltonian? For ξ to be Hamiltonian, the 1-form $\alpha_\xi = d\theta_2$ would need to be exact. That is, there would have to exist a globally defined smooth function $f \in C^\infty(T^2)$ such that $\alpha_\xi = d\theta_2 = -df$. However, $d\theta_2$ is not exact on the torus T^2 . We can see this by integrating $d\theta_2$ over the closed loop $\gamma : [0, 2\pi] \rightarrow T^2$ defined by $\gamma(t) = (0, t)$:

$$\int_\gamma d\theta_2 = \int_0^{2\pi} (d\theta_2)_{\gamma(t)}(\dot{\gamma}(t))dt$$

Since $\dot{\gamma}(t) = (0, 1)$ in velocity coordinates, corresponding to the vector $\frac{\partial}{\partial \theta_2}$ at $\gamma(t)$, the integrand is $(d\theta_2)(\frac{\partial}{\partial \theta_2}) = 1$. Thus,

$$\int_\gamma d\theta_2 = \int_0^{2\pi} 1dt = 2\pi \neq 0.$$

Since the integral of $d\theta_2$ over a closed loop is non-zero, $d\theta_2$ cannot be exact (by Stokes' Theorem, or the definition of exactness via path integrals). Therefore, $\xi = \frac{\partial}{\partial\theta_1}$ is a symplectic vector field that is not Hamiltonian.

This aligns perfectly with the fundamental exact sequence. The first de Rham cohomology of the torus is $H_{\text{dR}}^1(T^2) \cong \mathbb{R}^2$, with basis classes $[d\theta_1]$ and $[d\theta_2]$. The map $\Phi : \mathfrak{X}_\omega(T^2) \rightarrow H_{\text{dR}}^1(T^2)$ sends our vector field $\xi = \frac{\partial}{\partial\theta_1}$ to the cohomology class $\Phi(\xi) = [\iota_\xi\omega] = [d\theta_2]$. Since $[d\theta_2]$ is a non-zero element of $H_{\text{dR}}^1(T^2)$, ξ lies in $\mathfrak{X}_\omega(T^2)$ but not in the kernel of Φ , which is $\mathfrak{X}_{\text{Ham}}(T^2)$. Similarly, the vector field $\frac{\partial}{\partial\theta_2}$ is symplectic but not Hamiltonian, mapping via Φ to the class $[\iota_{\partial/\partial\theta_2}\omega] = [-d\theta_1]$, which is also non-zero. This example illustrates how the topology of the manifold ($H^1 \neq 0$) creates a distinction between the broader class of structure-preserving vector fields (symplectic) and the subclass generated by scalar functions (Hamiltonian).

Just as functions generate Hamiltonian vector fields, they inherit an algebraic structure reflecting the symplectic geometry.

Definition 2.77 (Poisson Bracket). Let (M, ω) be a symplectic manifold. The **Poisson bracket** is a bilinear operation $\{\cdot, \cdot\} : C^\infty(M) \times C^\infty(M) \rightarrow C^\infty(M)$ defined for any $f, g \in C^\infty(M)$ by:

$$\{f, g\} := \omega(\xi_f, \xi_g)$$

Using the definition $\iota_{\xi_f}\omega = -df$ and properties of contraction and the Lie derivative, we derive equivalent expressions:

$$\{f, g\} = \iota_{\xi_g}(\iota_{\xi_f}\omega) = \iota_{\xi_g}(-df) = -df(\xi_g) = -\mathcal{L}_{\xi_g}f$$

$$\{f, g\} = -\iota_{\xi_f}(\iota_{\xi_g}\omega) = -\iota_{\xi_f}(-dg) = dg(\xi_f) = \mathcal{L}_{\xi_f}g$$

The Poisson bracket $\{f, g\}$ quantifies the infinitesimal rate of change of the function g along the flow generated by the Hamiltonian f .

Proposition 2.78 (Properties of the Poisson Bracket). *The Poisson bracket endows the vector space $C^\infty(M)$ with the structure of an infinite-dimensional Lie algebra over \mathbb{R} . Specifically, it satisfies:*

1. **Bilinearity:** $\{af_1 + bf_2, g\} = a\{f_1, g\} + b\{f_2, g\}$ and $\{f, ag_1 + bg_2\} = a\{f, g_1\} + b\{f, g_2\}$ for $a, b \in \mathbb{R}$.
2. **Skew-symmetry:** $\{f, g\} = -\{g, f\}$. (Follows directly from $\omega(\xi_f, \xi_g) = -\omega(\xi_g, \xi_f)$).
3. **Jacobi Identity:** $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$.
4. **Leibniz Rule (Derivation Property):** $\{f, gh\} = g\{f, h\} + \{f, g\}h$. This means that for fixed f , the map $g \mapsto \{f, g\} = \mathcal{L}_{\xi_f}g$ is a derivation of the associative algebra $(C^\infty(M), \cdot)$.

Furthermore, the map $\delta : f \mapsto \xi_f$ is a Lie algebra homomorphism from $(C^\infty(M), \{\cdot, \cdot\})$ to $(\mathfrak{X}_{\text{Ham}}(M), [\cdot, \cdot])$, where $[\cdot, \cdot]$ is the commutator of vector fields:

$$[\xi_f, \xi_g] = \xi_{\{f, g\}}$$

Sketch of Jacobi Identity. The Jacobi identity for the Poisson bracket is a direct consequence of the closedness of ω and the relationship between brackets. One way to see this is to use the homomorphism property $\xi_{\{f, g\}} = [\xi_f, \xi_g]$. Applying ξ again: $\xi_{\{f, \{g, h\}\}} = [\xi_f, \xi_{\{g, h\}}] = [\xi_f, [\xi_g, \xi_h]]$. Summing

cyclically and using the Jacobi identity for the vector field commutator $[\xi_f, [\xi_g, \xi_h]] + [\xi_g, [\xi_h, \xi_f]] + [\xi_h, [\xi_f, \xi_g]] = 0$, we get $\xi_{\{f, \{g, h\}\}} + \xi_{\{g, \{h, f\}\}} + \xi_{\{h, \{f, g\}\}} = 0$. Since the map $\delta : f \mapsto \xi_f$ has only locally constant functions as its kernel, this implies $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\}$ must be locally constant. A more direct calculation using $d\omega = 0$ shows it is indeed zero. \square

The Poisson bracket provides the algebraic foundation for Hamiltonian mechanics. Observables are represented by functions $f \in C^\infty(M)$, and the time evolution under a Hamiltonian H is given by Hamilton's equation: $\frac{df}{dt} = \{f, H\}$. This structure is also the starting point for deformation quantization, where the commutative algebra $(C^\infty(M), \cdot)$ is deformed into a non-commutative algebra whose commutator approximates the Poisson bracket as $\hbar \rightarrow 0$.

2.4 Symmetries and Moment Maps

The concept of symmetry provides a powerful lens through which to analyze mathematical structures and physical systems. Symmetries often reveal hidden simplicities, lead to fundamental conservation laws, and constrain the possible dynamics. Within the framework of symplectic geometry, which forms the foundations of Hamiltonian mechanics, the study of symmetries takes on a particular elegance and importance. It leads to the theory of moment maps, a construction that precisely quantifies the relationship between continuous symmetries and conserved quantities, thereby providing a geometric version of Noether's celebrated theorem. We provide an in-depth exploration of symmetries in the symplectic context, from the basic definitions to the construction, properties, obstructions, and interpretation of moment maps.

Throughout this section, let (N, ω) denote a connected symplectic manifold of dimension $2k$.

2.4.1 Finite Symmetries: Symplectomorphism Groups

The most direct notion of symmetry for a geometric object is a transformation that leaves the object invariant. For a symplectic manifold (N, ω) , this corresponds to diffeomorphisms of N that preserve the symplectic form ω .

Earlier, we introduced the notion of a symplectomorphism. Let's define it slightly more rigorously.

Definition 2.79 (Symplectomorphism). A diffeomorphism $\varphi : N \rightarrow N$ is termed a **symplectic diffeomorphism**, or **symplectomorphism**, if it preserves the symplectic form ω . Mathematically, this means the pullback of ω by φ coincides with ω itself:

$$\varphi^*\omega = \omega$$

Recall that the pullback acts on forms via $(\varphi^*\omega)_n(v, w) = \omega_{\varphi(n)}(d\varphi_n(v), d\varphi_n(w))$ for $n \in N$ and $v, w \in T_nN$.

Remark 2.80 (Consequences and Properties).

- **Volume Preservation:** Since ω is non-degenerate, its k -th exterior power $\omega^k = \omega \wedge \cdots \wedge \omega$ (up to a normalization factor $1/k!$) defines a volume form on N . As pullback commutes with the wedge product, $\varphi^*(\omega^k) = (\varphi^*\omega)^k$. If φ is a symplectomorphism, then $\varphi^*(\omega^k) = \omega^k$. This demonstrates that symplectomorphisms are necessarily volume-preserving diffeomorphisms. This is the geometric statement of Liouville's theorem asserting the conservation of phase space volume under Hamiltonian evolution (which, as we shall see, consists of symplectomorphisms).

- **The Symplectomorphism Group:** The set of all symplectomorphisms of (N, ω) forms a group under composition, denoted $\text{Symp}(N, \omega)$. This group captures the global symmetries of the symplectic structure. Unlike the isometry group of a compact Riemannian manifold, which is always a finite-dimensional Lie group (Myers-Steenrod theorem), the group $\text{Symp}(N, \omega)$ is typically infinite-dimensional. For instance, on $(\mathbb{R}^{2n}, \omega_{\text{std}})$, the group $\text{Symp}(\mathbb{R}^{2n})$ contains not only linear symplectic transformations but also Hamiltonian flows of arbitrary non-linear functions. Understanding the structure (e.g., topology, homotopy type) of $\text{Symp}(N, \omega)$ and its subgroups (like the group of Hamiltonian diffeomorphisms) is a major focus of modern symplectic topology.

Example 2.81 (Basic Symplectomorphisms).

- Let $N = \mathbb{R}^{2n}$ with $\omega = \sum dq^i \wedge dp_i$. Any translation $\varphi_a(q, p) = (q + a, p + b)$ is a symplectomorphism if and only if $b = 0$. Position translations $\varphi_a(q, p) = (q + a, p)$ are symplectomorphisms. Additionally, a linear map $\varphi_A(x) = Ax$ is a symplectomorphism if and only if $A \in \text{Sp}(2n, \mathbb{R})$, meaning $A^T J A = J$ where J is the matrix of ω .
- Let $N = S^2$ with its standard rotationally invariant area form $\omega = \sin \phi d\phi \wedge d\theta$ (in spherical coordinates). Any rotation $R \in \text{SO}(3)$ acts on S^2 and preserves ω , hence $\text{SO}(3) \subset \text{Symp}(S^2, \omega)$.

2.4.2 More on Symplectic Vector Fields

Earlier, we discussed symplectic and Hamiltonian vector fields. Naturally, this leads to Noether's theorem. But before that, we need one more motivating proposition.

Proposition 2.82. *The space $\mathfrak{X}_{\text{Ham}}(N)$ is a Lie subalgebra of $\mathfrak{X}_\omega(N)$. In fact, under suitable conditions (e.g., N compact), $\mathfrak{X}_{\text{Ham}}(N)$ is an ideal in $\mathfrak{X}_\omega(N)$.*

Sketch for subalgebra. We need to show that if $\xi_f, \xi_g \in \mathfrak{X}_{\text{Ham}}(N)$, then $[\xi_f, \xi_g] \in \mathfrak{X}_{\text{Ham}}(N)$. It can be shown (often using $d\omega = 0$) that the commutator of two Hamiltonian vector fields is itself Hamiltonian, generated by (minus) the Poisson bracket of the functions: $[\xi_f, \xi_g] = \xi_{\{f, g\}}$, where $\{f, g\} = \omega(\xi_f, \xi_g)$. Since $\{f, g\} \in C^\infty(N)$, the commutator is indeed Hamiltonian. \square

The connection between symmetries and conserved quantities, intuited by Noether, finds a precise formulation in the Hamiltonian setting.

Theorem 2.83 (Noether's Theorem - Hamiltonian Version). *Let (N, ω) be a symplectic manifold and let $H \in C^\infty(N)$ define a Hamiltonian system whose flow φ_t is generated by the vector field $-\xi_H$. Suppose $\xi \in \mathfrak{X}_\omega(N)$ is an infinitesimal symmetry of the system such that:*

1. *The symmetry preserves the Hamiltonian function: $\mathcal{L}_\xi H = 0$.*
2. *The symmetry is generated by a momentum function: $\xi = \xi_f$ for some $f \in C^\infty(N)$.*

Then the momentum function f is a conserved quantity along the flow φ_t generated by $-\xi_H$; that is, $f(\varphi_t(n)) = f(n)$ for all n and t . Equivalently, the Poisson bracket of f and H vanishes: $\{f, H\} = 0$.

Proof. We compute the time derivative of f along the flow φ_t generated by $-\xi_H$. This rate of change

is given by the Lie derivative $\mathcal{L}_{-\xi_H} f$.

$$\frac{d}{dt}(f \circ \varphi_t) = (\mathcal{L}_{-\xi_H} f) \circ \varphi_t$$

We need to show $\mathcal{L}_{-\xi_H} f = 0$. Using the definition of the Poisson bracket $\{f, H\} = \mathcal{L}_{\xi_f} H = -\mathcal{L}_{\xi_H} f$, we can write:

$$\mathcal{L}_{-\xi_H} f = -\mathcal{L}_{\xi_H} f = \{f, H\}$$

Now, we use the assumptions. Assumption (1) is that the symmetry ξ preserves H , i.e., $\mathcal{L}_\xi H = 0$. Assumption (2) is that $\xi = \xi_f$. Substituting this into the symmetry condition gives $\mathcal{L}_{\xi_f} H = 0$. Recalling the alternative definition of the Poisson bracket, $\{f, H\} = \mathcal{L}_{\xi_f} H$. Combining these, we get:

$$\mathcal{L}_{-\xi_H} f = \{f, H\} = \mathcal{L}_{\xi_f} H = 0$$

Since the time derivative of f along the flow is zero, f is constant along the integral curves of $-\xi_H$, meaning it is a conserved quantity of the Hamiltonian evolution. This theorem elegantly establishes that Hamiltonian symmetries correspond precisely to conserved quantities. \square

Example 2.84 (Linear and Angular Momentum Conservation).

- **Linear Momentum:** Consider a particle in \mathbb{R}^n with Hamiltonian $H(q, p) = \frac{|p|^2}{2m} + V(q)$ on $T^*\mathbb{R}^n$. Suppose the potential V is invariant under translation in the q^j direction, i.e., $\frac{\partial V}{\partial q^j} = 0$. The generator of infinitesimal translation in q^j is $\xi = \frac{\partial}{\partial q^j}$. We know $\xi = \xi_{-p_j}$ (using $\omega = \sum dq^i \wedge dp_i$). The momentum function is $f = -p_j$. We check if ξ preserves H : $\mathcal{L}_{\partial/\partial q^j} H = \frac{\partial H}{\partial q^j} = \frac{\partial V}{\partial q^j} = 0$. By Noether's theorem, the momentum $f = -p_j$ (or simply p_j) is conserved.
- **Angular Momentum:** Consider a particle in \mathbb{R}^3 with a central potential $V(q) = V(|q|)$. The Hamiltonian $H = \frac{|p|^2}{2m} + V(|q|)$ is invariant under rotations $R \in SO(3)$. Let $\xi \in \mathfrak{so}(3)$ generate an infinitesimal rotation, with corresponding fundamental vector field $\alpha(\xi)$ on $T^*\mathbb{R}^3$. We know $\mathcal{L}_{\alpha(\xi)} H = 0$. The momentum function corresponding to $\alpha(\xi)$ is $\mu_\xi = \langle L, \xi \rangle$, where $L = q \times p$ is the angular momentum vector. Noether's theorem implies that each component of angular momentum μ_ξ is conserved.

2.4.3 Lie Algebra Actions and the Moment Map Problem

Often, a system possesses a whole family of continuous symmetries described by a Lie group G .

Definition 2.85 (Symplectic G -Action and Infinitesimal Action). A (left) action $\rho : G \times N \rightarrow N$ of a Lie group G on (N, ω) is **symplectic** if $\varphi_g : n \mapsto g \cdot n$ is a symplectomorphism for all $g \in G$. This action induces an **infinitesimal action** $\alpha : \mathfrak{g} \rightarrow \mathfrak{X}_\omega(N)$ of the Lie algebra $\mathfrak{g} = T_e G$, defined by

$$\alpha(\xi)_n = \left. \frac{d}{dt} \right|_{t=0} (\exp(-t\xi) \cdot n)$$

where $\xi \in \mathfrak{g}$. The map α is a Lie algebra antihomomorphism: $\alpha([\xi, \eta]) = -[\alpha(\xi), \alpha(\eta)]$. The minus sign convention ensures that left group actions correspond to antihomomorphisms into vector fields. Importantly, $\alpha(\xi)$ is always a symplectic vector field if the G -action is symplectic.

Given such an action $\alpha : \mathfrak{g} \rightarrow \mathfrak{X}_\omega(N)$, we seek a way to systematically associate conserved quantities (momentum functions) to the generators $\xi \in \mathfrak{g}$.

The Moment Map Problem: Can we find a map $\mu : N \rightarrow \mathfrak{g}^*$ such that for each $\xi \in \mathfrak{g}$, the function $\mu_\xi := \langle \mu, \xi \rangle$ is a momentum function for the infinitesimal symmetry $\alpha(\xi)$? That is, does $\alpha(\xi) = \xi_{\mu_\xi}$ hold (requiring $\iota_{\alpha(\xi)}\omega = -d\mu_\xi$)? Furthermore, can this map μ be chosen to respect the algebraic structures involved, i.e., to be G -equivariant?

The existence and properties of such a map μ , called the moment map, are governed by the topology of N and the structure of \mathfrak{g} .

Definition 2.86 (Moment Map via Lift). Consider the infinitesimal action $\alpha : \mathfrak{g} \rightarrow \mathfrak{X}_\omega(N)$ and the fundamental exact sequence linking functions to Hamiltonian and symplectic vector fields:

$$0 \rightarrow H^0(N; \mathbb{R}) \rightarrow C^\infty(N) \xrightarrow{\delta} \mathfrak{X}_{\text{Ham}}(N) \hookrightarrow \mathfrak{X}_\omega(N) \xrightarrow{\Phi} H_{\text{dR}}^1(N) \rightarrow 0$$

where $\delta(f) = \xi_f$ (defined by $\iota_{\xi_f}\omega = -df$) and $\Phi(\eta) = [\iota_\eta\omega]$. The action α **admits a lift** to $C^\infty(N)$ if the image of α lies entirely within $\mathfrak{X}_{\text{Ham}}(N)$ (i.e., $\Phi(\alpha(\xi)) = 0$ for all ξ) and if there exists a linear map $\tilde{\alpha} : \mathfrak{g} \rightarrow C^\infty(N)$ such that $\delta \circ \tilde{\alpha} = \alpha$. Such a map $\tilde{\alpha}$ assigns a momentum function $\mu_\xi = \tilde{\alpha}(\xi)$ to each $\xi \in \mathfrak{g}$. The **moment map** is then the map $\mu : N \rightarrow \mathfrak{g}^*$ defined by duality:

$$\langle \mu(n), \xi \rangle := \mu_\xi(n)$$

Usually, one also requires $\tilde{\alpha}$ to be compatible with the Lie algebra structures (equivariance), possibly up to terms related to H^0 .

Remark 2.87 (Obstructions to Existence of Moment Maps). *As mentioned before, there are two main obstructions:*

1. **Obstruction in H^1 :** *The primary obstruction is the vanishing of the map $\Phi \circ \alpha : \mathfrak{g} \rightarrow H_{\text{dR}}^1(N)$. For a moment map to exist, every fundamental vector field $\alpha(\xi)$ must be Hamiltonian, which means the closed 1-form $\iota_{\alpha(\xi)}\omega$ must be exact, i.e., $[\iota_{\alpha(\xi)}\omega] = 0$ in $H_{\text{dR}}^1(N)$. If $H_{\text{dR}}^1(N) = 0$, this obstruction always vanishes.*

Example 2.88 (Revisiting Torus T^2 Obstruction). Let $N = T^2 = (\mathbb{R}/2\pi\mathbb{Z})^2$ with $\omega = d\theta_1 \wedge d\theta_2$. Consider the action of $\mathfrak{g} = \mathbb{R}$ by translation along the first factor: $\rho_t(\theta_1, \theta_2) = (\theta_1 + t, \theta_2)$. The generator is $\xi = 1 \in \mathbb{R}$, and the infinitesimal action is $\alpha(1) = \frac{\partial}{\partial \theta_1}$. We calculated $\iota_{\alpha(1)}\omega = d\theta_2$. The map $\Phi : \mathfrak{X}_\omega(T^2) \rightarrow H_{\text{dR}}^1(T^2)$ sends $\alpha(1)$ to $\Phi(\alpha(1)) = [d\theta_2]$. Since $H_{\text{dR}}^1(T^2) \cong \mathbb{R}^2$ with basis $[d\theta_1], [d\theta_2]$, the class $[d\theta_2]$ is non-zero. Thus, the H^1 obstruction does not vanish. The symplectic vector field $\alpha(1) = \partial/\partial\theta_1$ is not Hamiltonian, and therefore this action does not admit a moment map in the sense defined above (no function f exists such that $\xi_f = \partial/\partial\theta_1$).

2. **Obstruction in H^2 (Central Extension):** *Even if $\Phi \circ \alpha = 0$, ensuring that each $\alpha(\xi)$ is Hamiltonian ($\alpha(\xi) = \xi_{\mu_\xi}$ for some μ_ξ), we still need to know if the map $\tilde{\alpha} : \xi \mapsto \mu_\xi$ can be chosen to be a Lie algebra antihomomorphism (up to constants). The failure of this is measured by the Lie algebra cocycle $c(\xi, \eta) = \{\mu_\xi, \mu_\eta\} - \mu_{[\xi, \eta]_{\text{inf}}}$. If this cocycle is non-zero (but maps to constants), it defines a class in $H^2(\mathfrak{g}; H^0(N; \mathbb{R}))$ and corresponds to a non-trivial central extension:*

$$0 \longrightarrow H^0(N; \mathbb{R}) \longrightarrow \bar{\mathfrak{g}} \longrightarrow \mathfrak{g} \longrightarrow 0$$

A lift $\tilde{\alpha} : \mathfrak{g} \rightarrow C^\infty(N)/H^0$ respecting the Lie algebra structure exists if and only if this central extension splits (is trivial). For semisimple Lie algebras \mathfrak{g} , $H^2(\mathfrak{g}; \mathbb{R}) = 0$, so this obstruction often vanishes if $H^0(N; \mathbb{R}) = \mathbb{R}$.

2.4.4 Properties and Characterization of Moment Maps

While the definition via lifting is conceptually important, the following characterization is often more practical.

Definition 2.89 (Moment Map (via equation)). Let $\alpha : \mathfrak{g} \rightarrow \mathfrak{X}_\omega(N)$ be an infinitesimal action. A smooth map $\mu : N \rightarrow \mathfrak{g}^*$ is called a **moment map** for this action if for every $\xi \in \mathfrak{g}$, the corresponding component function $\mu_\xi := \langle \mu, \xi \rangle \in C^\infty(N)$ satisfies the differential equation:

$$d\mu_\xi = -\iota_{\alpha(\xi)}\omega$$

This equation elegantly encodes both the requirement that each $\alpha(\xi)$ is Hamiltonian and that μ_ξ is its corresponding momentum function (up to a globally consistent choice of constants inherent in μ).

A important property sought in moment maps is compatibility with the group action.

Definition 2.90 (Equivariance of Moment Maps). Let a Lie group G with Lie algebra \mathfrak{g} act symplectically on (N, ω) via ρ , inducing $\alpha : \mathfrak{g} \rightarrow \mathfrak{X}_\omega(N)$. A moment map $\mu : N \rightarrow \mathfrak{g}^*$ for α is **G -equivariant** if for all $g \in G$ and $n \in N$:

$$\mu(\rho(g, n)) = \text{Ad}_g^*(\mu(n))$$

where $\text{Ad}_g^* : \mathfrak{g}^* \rightarrow \mathfrak{g}^*$ is the coadjoint action of G on \mathfrak{g}^* . Infinitesimally, this corresponds to the condition:

$$\mathcal{L}_{\alpha(\xi)}\mu_\eta = \mu_{[\xi, \eta]_{\text{inf}}}$$

where $\mu_\xi = \langle \mu, \xi \rangle$ and $[\cdot, \cdot]_{\text{inf}}$ is the bracket associated with α (caution with signs depending on conventions). Using the Poisson bracket and the relation $\alpha(\xi) = \xi_{\mu_\xi}$, the infinitesimal condition, up to constants related to the H^2 obstruction, becomes:

$$\{\mu_\xi, \mu_\eta\} \approx \mu_{[\xi, \eta]}$$

More precisely, a moment map satisfying $d\mu_\xi = -\iota_{\alpha(\xi)}\omega$ is equivariant if and only if the map $\tilde{\alpha} : \xi \mapsto \mu_\xi$ satisfies $\tilde{\alpha}([\xi, \eta]) - \{\tilde{\alpha}(\xi), \tilde{\alpha}(\eta)\}$ is a constant function (related to the cocycle $c(\xi, \eta)$). If an equivariant moment map exists, this cocycle must vanish.

Example 2.91 (Moment Maps Calculations - Revisited).

- **Translations on $T^*\mathbb{R}^n$:** $\alpha(e_j) = \partial/\partial q^j$. $\omega = \sum dq^i \wedge dp_i$. Defining equation $d\mu_j = -\iota_{\partial/\partial q^j}\omega = dp_j$. Choose $\mu_j = p_j$. Moment map $\mu(q, p) = (p_1, \dots, p_n)$. Coadjoint action of $G = \mathbb{R}^n$ on $\mathfrak{g}^* = \mathbb{R}^n$ is trivial ($\text{Ad}_a^*(p) = p$). Equivariance requires $\mu(q + a, p) = \text{Ad}_a^*(\mu(q, p))$, i.e., $p = p$. It is equivariant. Infinitesimally, \mathfrak{g} is abelian, $[\xi, \eta] = 0$. We need $\{\mu_\xi, \mu_\eta\} = 0$. If $\xi = \sum a^j e_j, \eta = \sum b^k e_k$, then $\mu_\xi = \sum a^j p_j, \mu_\eta = \sum b^k p_k$. $\{\sum a^j p_j, \sum b^k p_k\} = \sum_{j,k} a^j b^k \{p_j, p_k\} = 0$ since $\{p_j, p_k\} = 0$. Equivariance holds.
- **Rotations on $T^*\mathbb{R}^3$:** $\alpha(\xi)$ is generator of rotation. $\mu(q, p) = q \times p$. We need $d\langle q \times p, \xi \rangle = -\iota_{\alpha(\xi)}\omega$. This requires calculation using exterior derivatives and properties of the cross product, confirming it holds. Equivariance $\mu(Rq, Rp) = \text{Ad}_R^*(\mu(q, p)) = R(q \times p)$ holds because $R(q \times p) = (Rq) \times (Rp)$ for $R \in SO(3)$. Infinitesimally, this corresponds to $\{\mu_\xi, \mu_\eta\} = \mu_{[\xi, \eta]}$ (where bracket is cross product in $\mathbb{R}^3 \cong \mathfrak{so}(3)$), which holds for angular momentum components.

2.4.5 Geometric Interpretation: Prequantization Bundles

A geometric understanding of the moment map emerges from the framework of geometric quantization, which seeks to construct quantum theories from classical Hamiltonian systems. A key preliminary step is “prequantization.”

Assume the symplectic form ω satisfies the **Weil integrality condition**: the cohomology class $[\omega/(2\pi i\hbar)]$ (often simplified to $[\omega/2\pi]$ by setting $\hbar = 1$ and ignoring i) is an integral class, i.e., lies in the image of $H^2(N; \mathbb{Z}) \rightarrow H_{\text{dR}}^2(N)$. This topological condition is necessary and sufficient for the existence of a **prequantization line bundle**, typically a principal S^1 -bundle (or \mathbb{R} -bundle if we ignore periodicity) $\pi : T \rightarrow N$, equipped with a connection 1-form $\Theta \in \Omega^1(T)$ whose curvature $F_\Theta = d\Theta$ satisfies

$$F_\Theta = d\Theta = \pi^*\omega$$

. The connection Θ provides the link between the geometry of the bundle T and the symplectic structure on the base N .

Now, consider a symplectic action of a Lie group G on (N, ω) . This action is said to **lift** to the prequantization bundle T if there is a G -action $\tilde{\rho} : G \times T \rightarrow T$ on the total space such that:

1. It covers the action on N : $\pi(\tilde{\rho}(g, z)) = \rho(g, \pi(z))$.
2. It preserves the connection form Θ : $\tilde{\rho}_g^*\Theta = \Theta$ for all $g \in G$.

The existence of such a lift is related to group cohomology $H^2(G; \dots)$.

If a lift exists, let $\tilde{\xi} \in \mathfrak{X}(T)$ be the fundamental vector field on T corresponding to $\xi \in \mathfrak{g}$ via the lifted action $\tilde{\rho}$. Since the action preserves the connection, the Lie derivative vanishes: $\mathcal{L}_{\tilde{\xi}}\Theta = 0$. Consider the function defined by contracting the connection with the lifted vector field: $\iota_{\tilde{\xi}}\Theta$. This function on T turns out to be constant along the fibers of π (it is “basic”) and therefore defines a unique function $\mu_\xi \in C^\infty(N)$ such that

$$\pi^*\mu_\xi = \iota_{\tilde{\xi}}\Theta$$

We can now verify that the map $\mu : N \rightarrow \mathfrak{g}^*$ defined by $\langle \mu, \xi \rangle = \mu_\xi$ is indeed a moment map. We compute the exterior derivative:

$$\begin{aligned} \pi^*(d\mu_\xi) &= d(\pi^*\mu_\xi) \quad (\text{pullback commutes with } d) \\ &= d(\iota_{\tilde{\xi}}\Theta) \quad (\text{by definition of } \mu_\xi) \\ &= \mathcal{L}_{\tilde{\xi}}\Theta - \iota_{\tilde{\xi}}(d\Theta) \quad (\text{Cartan's formula}) \\ &= 0 - \iota_{\tilde{\xi}}(\pi^*\omega) \quad (\text{using } \mathcal{L}_{\tilde{\xi}}\Theta = 0 \text{ and } d\Theta = \pi^*\omega) \\ &= -\pi^*(\iota_{\alpha(\xi)}\omega) \quad (\text{relating contractions under } \pi_*) \end{aligned}$$

Since π^* is injective on functions (forms pulled back from base), we deduce

$$d\mu_\xi = -\iota_{\alpha(\xi)}\omega$$

This recovers the defining equation for the moment map. This construction demonstrates that the momentum function μ_ξ can be interpreted as the vertical component (measured by Θ) of the lifted symmetry generator $\tilde{\xi}$ on the prequantization bundle. This perspective is important for geometric quantization, where wave functions are sections of T (or associated bundles) and operators corresponding to observables f involve both multiplication by f and differentiation along ξ_f , tied together by the connection Θ .

3 Hamiltonian and Lagrangian Mechanics

Abstract

We explore foundational structures of Hamiltonian and Lagrangian mechanics, emphasizing their geometric underpinnings. We begin by defining Hamiltonian systems through a symplectic manifold (N, ω) representing the phase space, and a one-parameter group of symplectic automorphisms φ_t dictating time evolution. Observables are functions on N , with their algebraic structure governed by the Poisson bracket $\{f, g\} = \omega(\xi_f, \xi_g)$, where ξ_f is the Hamiltonian vector field of f satisfying $i_{\xi_f}\omega = -df$. The framework is illustrated by deriving Hamilton's equations for particle motion in Euclidean space, $\dot{x} = y, \dot{y} = -V'(x)/m$, from the Hamiltonian $H = \frac{my^2}{2} + V(x)$ and symplectic form $\omega = m dy \wedge dx$. Generalizations to particle motion on Riemannian manifolds, where the phase space is the tangent bundle TM , are also considered.

We then transition to variational principles, initially examining geodesics in Euclidean space via the energy functional $E(x) = \int \frac{1}{2}|\dot{x}|^2 dt$, whose stationarity yields the geodesic equation $\ddot{x} = 0$. This principle is reformulated with greater generality and local insight using the calculus of differential forms on the space of paths $\mathcal{F} \times \mathbb{R}$. The Lagrangian 1-form $L = L_0(x, \dot{x}) dt \in \Omega^{0,1}$ and the variational 1-form $\gamma \in \Omega^{1,0}$ (e.g., $\gamma = m\langle \delta x, \dot{x} \rangle$ for a free particle) are introduced. The Euler-Lagrange equations are then rewritten locally as $\delta L + d\gamma = 0$ on the manifold of solutions, where δ and d are exterior derivatives on \mathcal{F} and \mathbb{R} respectively. This framework naturally describes general Lagrangian systems defined by a Lagrangian function $L_0(x, \dot{x})$.

The deep connection between Lagrangian and Hamiltonian formalisms then discussed. The canonical 2-form $\omega = \delta\gamma \in \Omega^{2,0}$, derived from the Lagrangian data, is shown to be δ -closed, d -closed (time-independent), and non-degenerate, thus defining the symplectic structure on the space of solutions. The Hamiltonian function $H = \iota_{\xi}(L + \gamma)$ is constructed as the Noether conserved quantity corresponding to time-translation invariance (where ξ generates time evolution), satisfying $\delta H = -\iota_{\xi}\omega$. Finally, a geometric interpretation of $\gamma(t)$ as a connection 1-form on the space of solutions is presented. The relation $\delta \int_{t_0}^{t_1} L = -(\gamma(t_1) - \gamma(t_0))$ reveals that the integrated Lagrangian acts as an isomorphism between these connections.

Contents

3.1	Hamiltonian Systems	69
3.1.1	Basic Definitions	69
3.1.2	Particle Motion in Euclidean Space	70
3.1.3	Remarks on Hamiltonian Systems	72
3.2	Variational Principles for Geodesics	73
3.2.1	Introduction	73
3.2.2	Via Differential Forms	74
3.3	Lagrangian Systems	76
3.3.1	The Principle of Stationary Action and Equations of Motion	77
3.3.2	Transition to Hamiltonian Mechanics	77
3.4	Geometric Interpretation of γ	80

References: For foundational treatments of classical mechanics, including Newtonian, Lagrangian, and Hamiltonian formalisms, Arnold's text [2] provides a mathematically sophisticated perspective, while Feynman's lectures [11] offer valuable physical intuition, particularly regarding the principle of least action.

3.1 Hamiltonian Systems

We now introduce the central objects of study in Hamiltonian mechanics: Hamiltonian systems. These provide a powerful geometric framework for describing the evolution of many physical systems, from classical particle mechanics to more abstract dynamical systems. The main object of this framework is a symplectic manifold, whose structure dictates the rules of motion.

3.1.1 Basic Definitions

Definition 3.1 (Hamiltonian System).

1. A **Hamiltonian system** consists of a symplectic manifold (N, ω) and a 1-parameter group φ_t of symplectic automorphisms of N .⁹ Here, N is the **phase space** and ω is the symplectic form.
2. A Hamiltonian system is **free** if (N, ω) is an affine symplectic space and φ_t is a 1-parameter group of affine symplectic automorphisms.

The symplectic manifold N represents the space of possible states of the system, and the symplectic form ω gives the Poisson bracket structure of the system. The 1-parameter group of symplectic automorphisms φ_t describes the time evolution of the system, mapping an initial state at time $t = 0$ to its state at time t .

A Hamiltonian system naturally gives rise to a mechanical system:

1. The space of pure states is the symplectic manifold N itself. The space of all states can be considered as the space of probability distributions on N , denoted by $\text{Prob}(N)$.
2. The algebra of observables is given by the space of complex-valued Borel functions on N , denoted by $\mathcal{O} = \text{Borel}(N; \mathbb{C})$. The real structure is given by complex conjugation. The subspace of smooth functions $\mathcal{O}^\infty = C^\infty(N; \mathbb{C})$ is dense in \mathcal{O} and carries a Lie algebra structure defined by the Poisson bracket:

$$\{f, g\} = \omega(\xi_f, \xi_g)$$

where ξ_f and ξ_g are the symplectic gradient vector fields of the functions f and g , respectively, defined by $df = \omega(\xi_f, \cdot)$ and $dg = \omega(\xi_g, \cdot)$. The real observables are given by the real-valued Borel functions $\mathcal{O}_{\mathbb{R}} = \text{Borel}(N; \mathbb{R})$, with a dense subspace of smooth real-valued functions $\mathcal{O}_{\mathbb{R}}^\infty = C^\infty(N; \mathbb{R}) = \Omega_N^0$.

3. The measurement pairing associates an observable $A : N \rightarrow \mathbb{R}$ and a state $\sigma \in \text{Prob}(N)$ to the pushforward probability measure $A_*\sigma \in \text{Prob}(\mathbb{R})$, which describes the probability distribution of the measured values of the observable A in the state σ .
4. A smooth observable $f \in \Omega_N^0$ has a symplectic gradient vector field $\xi_f \in \mathcal{X}_N$ defined by $df = \omega(\xi_f, \cdot)$. If ξ_f is complete (i.e., its integral curves exist for all time), it generates a 1-parameter flow ψ_t of symplectic diffeomorphisms. This flow induces a flow on states (by pushing forward the measure) and on complex observables (by pulling back the function), and these induced flows are compatible.
5. The chosen 1-parameter group φ_t of symplectic automorphisms in the definition of a Hamiltonian system is precisely the flow that generates the time evolution of the system on both states and observables.

⁹A symplectic automorphism is a diffeomorphism $\varphi : N \rightarrow N$ such that $\varphi^*\omega = \omega$.

Remark 3.2. *It is important to note that while every smooth observable has a symplectic gradient vector field, this vector field may not always generate a global flow. The existence of a global flow requires completeness of the vector field, which is not guaranteed in general. Therefore, the axiom regarding the existence of a global flow for all smooth observables is often something we hope for, rather than an axiom.*

3.1.2 Particle Motion in Euclidean Space

We consider the motion of a particle on the Euclidean line \mathbb{E}^1 . The dynamics of such a system are described within the Hamiltonian formalism.

Let N be the space of classical trajectories, meaning the space of functions $x : \mathbb{R} \rightarrow \mathbb{E}^1$ that satisfy Newton's second law for a given potential. For any specific trajectory, given an initial time $t_0 \in \mathbb{R}$ and initial conditions (position and velocity), a local solution for the trajectory exists. It is assumed that the potential V is sufficiently well-behaved, ensuring that solutions $x(t)$ exist for all time.

Under these conditions, for any chosen time $t_0 \in \mathbb{R}$, the evaluation map

$$N \longrightarrow \mathbb{E}^1 \times \mathbb{R}$$

which maps a trajectory $x(t)$ to its position $x(t_0)$ and velocity $\dot{x}(t_0)$ at time t_0 , is a diffeomorphism. This diffeomorphism allows an identification of the space of trajectories N with the phase space $\mathbb{E}^1 \times \mathbb{R}$.

We use coordinates (x, y) for this phase space, where ‘ x ’ is the coordinate on \mathbb{E}^1 representing position, and ‘ y ’ is the coordinate on \mathbb{R} representing velocity (i.e., $y = \dot{x}$). The symplectic form ω on this phase space is given by:

$$\omega = m \, dy \wedge dx$$

where m is the mass of the particle. This symplectic form corresponds to the canonical form $dp \wedge dq$ under the identification $q = x$ (position) and $p = my$ (canonical momentum).

The time evolution flow $\varphi_t : N \rightarrow N$ is generated by the Hamiltonian function H . For this system, H is the total mechanical energy, expressed as the sum of kinetic and potential energy:

$$H(x, y) = \frac{my^2}{2} + V(x)$$

where $V(x)$ is the potential energy function. The exterior derivative (or differential) of H is:

$$dH = \frac{\partial H}{\partial x} dx + \frac{\partial H}{\partial y} dy = V'(x) dx + my \, dy$$

where $V'(x) = \frac{dV}{dx}$.

The dynamics are governed by the Hamiltonian vector field ξ_H associated with H . The symplectic gradient is defined by the relation:

$$i_{\xi_H} \omega = -dH$$

where $i_{\xi_H} \omega$ is the interior product of ω with ξ_H . Let the components of ξ_H in our coordinate system be $A(x, y)$ and $B(x, y)$, such that:

$$\xi_H = A(x, y) \frac{\partial}{\partial x} + B(x, y) \frac{\partial}{\partial y}$$

We compute the interior product $i_{\xi_H}\omega$:

$$\begin{aligned} i_{\xi_H}\omega &= i_{A\frac{\partial}{\partial x}+B\frac{\partial}{\partial y}}(m\,dy\wedge dx) \\ &= A\cdot i_{\frac{\partial}{\partial x}}(m\,dy\wedge dx) + B\cdot i_{\frac{\partial}{\partial y}}(m\,dy\wedge dx) \end{aligned}$$

To evaluate the interior products with the basis vector fields:

$$\begin{aligned} i_{\frac{\partial}{\partial x}}(m\,dy\wedge dx) &= m\left((dy(\frac{\partial}{\partial x}))dx - (dx(\frac{\partial}{\partial x}))dy\right) \\ &= m(0\cdot dx - 1\cdot dy) = -m\,dy \end{aligned}$$

and

$$\begin{aligned} i_{\frac{\partial}{\partial y}}(m\,dy\wedge dx) &= m\left((dy(\frac{\partial}{\partial y}))dx - (dx(\frac{\partial}{\partial y}))dy\right) \\ &= m(1\cdot dx - 0\cdot dy) = m\,dx \end{aligned}$$

Substituting these results back into the expression for $i_{\xi_H}\omega$:

$$i_{\xi_H}\omega = A(-m\,dy) + B(m\,dx) = mB\,dx - mA\,dy$$

Now, we apply the defining equation $i_{\xi_H}\omega = -dH$:

$$\begin{aligned} mB\,dx - mA\,dy &= -(V'(x)dx + my\,dy) \\ mB\,dx - mA\,dy &= -V'(x)dx - my\,dy \end{aligned}$$

For this equality of 1-forms to hold, the coefficients of the basis 1-forms (dx and dy) on both sides must be equal:

- Equating coefficients of dx : $mB = -V'(x) \implies B = -\frac{1}{m}V'(x)$
- Equating coefficients of dy : $-mA = -my \implies A = y$

Thus, the Hamiltonian vector field is explicitly:

$$\xi_H = y\frac{\partial}{\partial x} - \frac{1}{m}V'(x)\frac{\partial}{\partial y}$$

Consequently, Hamilton's equations for the flow φ_t , which describe the time evolution of x and y , are obtained from the components of ξ_H (i.e., $\dot{x} = A$ and $\dot{y} = B$):

$$\begin{cases} \dot{x} = y \\ \dot{y} = -\frac{1}{m}V'(x) \end{cases}$$

These first-order differential equations are equivalent to Newton's second law, $m\ddot{x} = -V'(x)$. This can be verified by noting that $\dot{x} = y$, so differentiating with respect to time yields $\ddot{x} = \dot{y}$. Substituting this into the second Hamilton's equation gives $\ddot{x} = -\frac{1}{m}V'(x)$, or $m\ddot{x} = -V'(x)$.

3.1.3 Remarks on Hamiltonian Systems

Remark 3.3.

1. The transition from Newton's second-order differential equation to the Hamiltonian system of two first-order differential equations is a standard technique achieved by introducing the velocity (or momentum) as an independent variable.
2. The phase space $N = \mathbb{E}^1 \times \mathbb{R}$ is an affine space. For a free particle ($V = 0$), the Hamiltonian is $H = \frac{1}{2}my^2$, and the equations of motion are $\dot{x} = y$, $\dot{y} = 0$, which represent a flow by affine symplectic diffeomorphisms. For a harmonic oscillator ($V(x) = \frac{1}{2}kx^2$ with $k > 0$), the Hamiltonian is $H = \frac{1}{2}my^2 + \frac{1}{2}kx^2$, and the equations of motion are $\dot{x} = y$, $\dot{y} = -\frac{k}{m}x$, which also generate a flow by affine symplectic diffeomorphisms. Thus, both systems are free Hamiltonian systems.
3. Our example be generalized to the motion of a particle on a Riemannian manifold (M, g) under the influence of a potential $V : M \rightarrow \mathbb{R}$. In this case, the acceleration \ddot{x} is replaced by the acceleration computed using the Levi-Civita covariant derivative associated with the metric g . The phase space N becomes the tangent bundle TM of the configuration space M . Assuming that the equations of motion have solutions defined for all time, the map that takes a trajectory to its position and velocity at a given time is a diffeomorphism onto TM . The Riemannian metric g induces an isomorphism between the tangent bundle TM and the cotangent bundle T^*M . The symplectic form on TM is related to the pullback of the tautological symplectic form on T^*M . Specifically, using the metric to identify TM with T^*M , the symplectic form is m times the pullback of the tautological 2-form. The Hamiltonian function is given by:

$$H(\xi) = \frac{m}{2}|\xi|^2 + V(\pi(\xi))$$

where $\xi \in TM$, $|\xi|^2 = g(\xi, \xi)$ is the squared norm of the tangent vector with respect to the metric g , and $\pi : TM \rightarrow M$ is the projection map. The Riemannian manifold M is often referred to as the configuration space, while $N = TM$ is the phase space.

4. Later in this lecture, we will demonstrate how to derive the symplectic form $\omega = m dy \wedge dx$ and the Hamiltonian function $H = \frac{1}{2}my^2 + V(x)$ (and their generalizations to Riemannian manifolds) from a more fundamental starting point: the Lagrangian formalism.

Example 3.4. The phase space N of a Hamiltonian system does not necessarily have to be the tangent bundle of a manifold, and the 1-parameter group φ_t of symplectic automorphisms need not always be generated by a global Hamiltonian function. Consider the 2-torus $N = \mathbb{R}/2\pi\mathbb{Z} \times \mathbb{R}/2\pi\mathbb{Z}$ with coordinates (θ^1, θ^2) . The standard symplectic form on \mathbb{R}^2 is $\omega_0 = dx \wedge dy$. On the torus, we can consider the symplectic form $\omega = d\theta^1 \wedge d\theta^2$. The 1-parameter group of diffeomorphisms given by:

$$\varphi_t(\theta^1, \theta^2) = (\theta^1 + t, \theta^2)$$

is a symplectic automorphism because:

$$\varphi_t^*(d\theta^1 \wedge d\theta^2) = d(\theta^1 + t) \wedge d\theta^2 = d\theta^1 \wedge d\theta^2$$

This system can be seen as the quotient of a free system on \mathbb{R}^2 (with Hamiltonian $H(x, y) = y$) by the action of the discrete group $(2\pi\mathbb{Z} \times 2\pi\mathbb{Z})$ of automorphisms, where the quotient map is $(x, y) \mapsto (x \bmod 2\pi, y \bmod 2\pi)$.

3.2 Variational Principles for Geodesics

3.2.1 Introduction

Variational principles play a fundamental role in many areas of mathematics and physics, providing an alternative way to formulate equations of motion and geometric structures. They are particularly important in global analysis on manifolds, where they are used to study the existence and properties of critical points of functionals defined on infinite-dimensional spaces of maps.

Let V be a finite-dimensional real inner product space, and let A be an affine space over V (a Euclidean space). Fix real numbers $a < b$, and consider smooth paths $x : [a, b] \rightarrow A$. The length of the path is given by:

$$L(x) = \int_a^b |\dot{x}(t)| dt = \int_a^b \sqrt{\langle \dot{x}(t), \dot{x}(t) \rangle} dt$$

where $\dot{x}(t)$ is the velocity vector of the path at time t , and $|\cdot|$ denotes the norm induced by the inner product $\langle \cdot, \cdot \rangle$. The length of the path is independent of the parametrization.

Instead of directly working with the length functional, it is often more convenient to consider the energy functional:

$$E(x) = \int_a^b \frac{1}{2} |\dot{x}(t)|^2 dt = \int_a^b \frac{1}{2} \langle \dot{x}(t), \dot{x}(t) \rangle dt$$

A path that minimizes the energy functional (among paths with fixed endpoints) also minimizes the length functional, and vice versa, provided the length is non-zero.

To find the paths that make the energy functional stationary, we consider a variation of a given path $x(t)$. Let $x(t, s)$ be a smooth family of paths parametrized by $s \in (-\epsilon, \epsilon)$, such that $x(t, 0) = x(t)$ and the endpoints are fixed, i.e., $x(a, s) = x(a, 0)$ and $x(b, s) = x(b, 0)$ for all s . Let $\delta x(t) = \frac{\partial x}{\partial s}(t, 0)$ be the variation vector field along the path $x(t)$.

The variation of the energy functional is given by:

$$\delta E = \left. \frac{d}{ds} \right|_{s=0} E(x(\cdot, s)) = \left. \frac{d}{ds} \right|_{s=0} \int_a^b \frac{1}{2} \langle \frac{\partial x}{\partial t}(t, s), \frac{\partial x}{\partial t}(t, s) \rangle dt$$

We can interchange the derivative with respect to s and the integral:

$$\begin{aligned} \delta E &= \int_a^b \frac{1}{2} \frac{\partial}{\partial s} \langle \frac{\partial x}{\partial t}(t, s), \frac{\partial x}{\partial t}(t, s) \rangle \Big|_{s=0} dt \\ &= \int_a^b \langle \frac{\partial^2 x}{\partial s \partial t}(t, 0), \frac{\partial x}{\partial t}(t, 0) \rangle dt \end{aligned}$$

Since the partial derivatives commute, $\frac{\partial^2 x}{\partial s \partial t} = \frac{\partial^2 x}{\partial t \partial s} = \frac{\partial}{\partial t} (\frac{\partial x}{\partial s}) = \frac{\partial}{\partial t} (\delta x) = \delta \dot{x}$. Thus,

$$\delta E = \int_a^b \langle \delta \dot{x}(t), \dot{x}(t) \rangle dt$$

Using integration by parts, we have:

$$\int_a^b \langle \delta \dot{x}(t), \dot{x}(t) \rangle dt = \left[\langle \delta x(t), \dot{x}(t) \rangle \right]_a^b - \int_a^b \langle \delta x(t), \ddot{x}(t) \rangle dt$$

Since the endpoints are fixed, $\delta x(a) = 0$ and $\delta x(b) = 0$, so the boundary term vanishes. Therefore,

$$\delta E = - \int_a^b \langle \delta x(t), \ddot{x}(t) \rangle dt$$

For the energy functional to be stationary ($\delta E = 0$) for all variations δx with fixed endpoints, we must have $\ddot{x}(t) = 0$ for all $t \in [a, b]$. This equation describes a motion with constant velocity, which is a geodesic in the Euclidean space A .

3.2.2 Via Differential Forms

The variational principle, which yields the equations of motion such as the geodesic equation for a free particle, possesses an important local structure. This fundamental locality is most effectively revealed by re-studying the variational calculus in the language of differential forms on spaces of paths.

Remark 3.5 (On Boundary Conditions and Locality). *In the computation of the first variation of the energy functional $E(x)$, one encounters a boundary term arising from integration by parts:*

$$\int_{[a,b]} d_t \langle \delta x(t), \dot{x}(t) \rangle dt = \langle \delta x, \dot{x} \rangle|_{t=b} - \langle \delta x, \dot{x} \rangle|_{t=a}. \quad (3.1)$$

Let us denote $\gamma_0(t) = \langle \delta x(t), \dot{x}(t) \rangle$. This $\gamma_0(t)$ represents, for each time t , a function on the space of variations S (or, more generally, on the infinite-dimensional manifold of paths \mathcal{F}). The full variational 1-form γ discussed later will be $\gamma = \gamma_0 \in \Omega^{1,0}$. It is customary in variational problems to impose boundary conditions on the paths $x(t)$ such that these boundary terms vanish. For instance, one might fix the initial and final positions, $x(a)$ and $x(b)$, implying $\delta x(a) = \delta x(b) = 0$.

Importantly, however, the Euler-Lagrange equation itself (in this case, the geodesic equation $\ddot{x}(t) = 0$) is a local condition in time. Its derivation relies on the vanishing of the integrand of the variation for arbitrary variations $\delta x(t)$ supported away from the boundary. Thus, the specific choice of boundary conditions, while important for selecting a particular solution, does not affect the form of the local equations of motion.

To study this local structure more clearly, we transition our focus from the global energy functional to the local *Lagrangian function* $L_0(x, \dot{x}) = \frac{1}{2} \langle \dot{x}, \dot{x} \rangle$ and its associated *Lagrangian 1-form* (a 1-form with respect to the time variable)

$$L = L_0(x, \dot{x}) dt = \frac{1}{2} \langle \dot{x}, \dot{x} \rangle dt. \quad (3.2)$$

Let S be a smooth manifold parameterizing a family of paths $x : S \times [a, b] \rightarrow A$. Then L can be viewed as an element of $\Omega^{0,1}(S \times [a, b])$, i.e., a 0-form on S and a 1-form on $[a, b]$. The total energy is the pushforward $E = (pr_S)_*(L) \in \Omega^0(S)$, where $pr_S : S \times [a, b] \rightarrow S$ is the projection.

The natural arena for analyzing the local structure of variational problems is the product space $S \times [a, b]$ (or, more generally, $\mathcal{F} \times [a, b]$, where \mathcal{F} denotes the infinite-dimensional manifold of smooth paths $x : [a, b] \rightarrow A$). The calculus of variations can be elegantly formulated using the de Rham complex on this product space. We distinguish between variations of paths and evolution in time by introducing two differential operators:

- δ : the exterior derivative acting on forms with respect to the variables parameterizing the space of paths S (or \mathcal{F}).
- d : the exterior derivative acting on forms with respect to the time variable $t \in [a, b]$, so $d = dt \frac{d}{dt}$.

The total exterior derivative on $\Omega^\bullet(S \times [a, b])$ is $D = \delta + d$. These operators naturally endow the space of differential forms with a bicomplex structure, denoted $\Omega^{\bullet, \bullet}(S \times [a, b])$. A form $\omega \in \Omega^{p, q}(S \times [a, b])$ has degree p in the S -variables (often called the “variational degree”) and degree q in the time variable.

Within this bicomplex framework:

- The Lagrangian 1-form $L = L_0(x, \dot{x}) dt$ is manifestly an element of $\Omega^{0,1}(S \times [a, b])$.
- The variational 1-form $\gamma = \langle \delta x, \dot{x} \rangle$ (where δx is the 1-form on S obtained by applying δ to x) is an element of $\Omega^{1,0}(S \times [a, b])$. It is a 1-form with respect to variations on S and a 0-form in time.

Applying the exterior derivative δ to L yields $\delta L = (\delta L_0) dt = \langle \dot{x}, \delta \dot{x} \rangle dt \in \Omega^{1,1}(S \times [a, b])$. Similarly, applying d to γ gives $d\gamma = \left(\frac{d}{dt} \langle \delta x, \dot{x} \rangle\right) dt \in \Omega^{1,1}(S \times [a, b])$. The important identity derived from the standard integration by parts procedure is:

$$\langle \dot{x}, \delta \dot{x} \rangle = -\langle \delta x, \ddot{x} \rangle - \frac{d}{dt} \langle \delta x, \dot{x} \rangle. \quad (3.3)$$

Multiplying by dt , this identity translates into the language of forms as: $\delta L = -\langle \delta x, \ddot{x} \rangle dt - d\gamma$. This equation can be rewritten as $\delta L + d\gamma = -\langle \delta x, \ddot{x} \rangle dt$. The term $EL[\delta x] := -\langle \delta x, \ddot{x} \rangle dt$ is the Euler-Lagrange form associated with the variation δx . The requirement that $EL[\delta x]$ vanishes for arbitrary variations δx (supported away from boundaries) yields the Euler-Lagrange equation $\ddot{x} = 0$.

Let N be the manifold of solutions to the equations of motion (i.e., paths $x(t)$ satisfying $\ddot{x}(t) = 0$). These solutions are typically considered to be defined for all $t \in \mathbb{R}$. On the restricted space $N \times \mathbb{R}$, where $\ddot{x} = 0$ by definition, the Euler-Lagrange form $EL[\delta x]$ vanishes. Consequently, the relationship $\delta L + d\gamma = EL[\delta x]$ simplifies to:

$$\delta L + d\gamma = 0 \quad \text{on } N \times \mathbb{R}. \quad (3.4)$$

Here, $L \in \Omega^{0,1}(N \times \mathbb{R})$ and $\gamma \in \Omega^{1,0}(N \times \mathbb{R})$. Equation (3.4) expresses the vanishing of a $(1, 1)$ -form on the space of solutions extended over time. It is a very important piece of the covariant/geometric formulation of Lagrangian mechanics and field theory.

Furthermore, for the free particle Lagrangian, $L_0 = \frac{1}{2}|\dot{x}|^2$, $dL = d_t(\frac{1}{2}|\dot{x}|^2 dt) = \left(\frac{d}{dt}(\frac{1}{2}|\dot{x}|^2)\right) dt \wedge dt = 0$, since $dt \wedge dt = 0$. If one also assumes that γ is δ -closed on N (i.e., $\delta\gamma = 0$), which is often the case for forms constructed from fields and their first variations without involving higher derivatives of variations, then the combined object $L + \gamma$ is D -closed on $N \times \mathbb{R}$: $D(L + \gamma) = \delta L + dL + \delta\gamma + d\gamma = (\delta L + d\gamma) + dL + \delta\gamma = 0 + 0 + 0 = 0$. The condition $\delta L + d\gamma = 0$ is thus the important local statement encapsulating the equations of motion within this formalism.

Remark 3.6 (Further Aspects of the Variational Formalism).

1. *Origin of the variational 1-form γ :* The 1-form $\gamma = \langle \delta x, \dot{x} \rangle$ arises naturally from the integration by parts procedure. A more sophisticated geometric perspective, often developed in the context of the multisymplectic formalism or the calculus of variations on jet bundles, provides a

construction of such forms (and their generalizations) without explicit reliance on integration. This involves defining appropriate currents and forms on spaces of field histories.

2. *A priori justification for the locality of Euler-Lagrange equations:* The fact that the critical point condition for an action functional $E(x) = \int_a^b L_0(x(t), \dot{x}(t))dt$ yields a differential equation (local in time) can be understood intuitively. If a path x_0 minimizes (or extremizes) $E(x)$, it must necessarily do so for any sub-interval. Consider an infinitesimal interval $[t_0, t_0 + \varepsilon]$ for small $\varepsilon > 0$. The condition that x_0 extremizes $\int_{t_0}^{t_0+\varepsilon} L_0 dt$ for all such infinitesimal intervals, upon taking the limit $\varepsilon \rightarrow 0$ (after appropriate scaling), leads directly to the Euler-Lagrange equations evaluated at t_0 . This underscores the local nature of these equations.
3. *Significance of variational principles:* The formulation of physical laws or geometric problems as variational principles is exceptionally powerful. If a system of ordinary or partial differential equations can be identified as the Euler-Lagrange equations for some functional (the action or energy), this functional provides important insights.
 - It often serves as a guiding principle for constructing numerical methods to find solutions, e.g., via direct minimization or gradient flow techniques.
 - The functional itself can yield important a priori estimates (bounds, coercivity properties, etc.) essential for the analytical study of the existence, uniqueness, and regularity of solutions to the differential equations.
 - Symmetries of the action functional, via Noether's theorem, lead to conservation laws, which are fundamental in physical theories.

3.3 Lagrangian Systems

The Lagrangian approach to mechanics, rooted in the principle of stationary action (often attributed to Maupertuis, with significant contributions from Euler, Leibniz, and others), provides a powerful alternative to the Newtonian framework. While its historical development is rich, our focus here is on the formal mathematical structure, which underpins many areas of modern physics and geometry.

Remark 3.7.

1. *Scope:* It is noteworthy that not every Hamiltonian system admits a Lagrangian formulation. The existence of a Lagrangian description often implies additional structure or regularity.
2. *Analytical Power:* A Lagrangian formulation can lead to powerful estimates and greater analytical control over the system's dynamics, particularly when studying the existence and properties of solutions to the equations of motion.
3. *Geometric Structure:* As we will explore, the Lagrangian viewpoint often reveals deeper geometric structures associated with a mechanical system, including connections to symplectic and contact geometry.
4. *Quantum Transition:* The Lagrangian, specifically the action functional, is the fundamental starting point for Feynman's path integral formulation of quantum mechanics, providing a direct bridge from classical to quantum descriptions.

3.3.1 The Principle of Stationary Action and Equations of Motion

Consider a particle of mass m moving on a Riemannian manifold (M, g) , subject to a potential energy function $V : M \rightarrow \mathbb{R}$. A path of the particle is a smooth map $x : [a, b] \rightarrow M$. The *Lagrangian* $L_0 : TM \rightarrow \mathbb{R}$ for this system is defined as the difference between kinetic and potential energy:

$$L_0(x(t), \dot{x}(t)) = \frac{m}{2} g(\dot{x}(t), \dot{x}(t)) - V(x(t)), \quad (3.5)$$

where $\dot{x}(t) \in T_{x(t)}M$ is the velocity vector. The *action functional* S assigns to each path x the integral of the Lagrangian over the time interval:

$$S(x) = \int_a^b L_0(x(t), \dot{x}(t)) dt = \int_a^b \left[\frac{m}{2} |\dot{x}(t)|_g^2 - V(x(t)) \right] dt. \quad (3.6)$$

The principle of stationary action states that the actual paths taken by the system are those for which the action $S(x)$ is stationary with respect to variations of x that fix the endpoints $x(a)$ and $x(b)$.

Both the Lagrangian L_0 and the action S have physical dimensions of (mass)(length)²(time)⁻¹. These are referred to as units of *action*.

Remark 3.8 (On the Nature of the Lagrangian). *The Lagrangian is a somewhat magical quantity. Unlike energy, force, or momentum, which often have direct and intuitive physical interpretations, the Lagrangian's role is more abstract, serving as the integrand whose stationary integral yields the equations of motion. Its remarkable ability in describing a vast range of physical phenomena makes it fundamentally important, even if its direct interpretation remains subtle.*

We employ the variational bicomplex formalism introduced previously. Let \mathcal{F} be the infinite-dimensional manifold of smooth paths $x : \mathbb{R} \rightarrow M$. The relevant space for local analysis is $\mathcal{F} \times \mathbb{R}$. The Lagrangian 1-form $L \in \Omega^{0,1}(\mathcal{F} \times \mathbb{R})$ is $L = L_0(x, \dot{x})dt$. The variational 1-form $\gamma \in \Omega^{1,0}(\mathcal{F} \times \mathbb{R})$ associated with variations δx is given by the fiber derivative of L_0 with respect to velocity, paired with δx :

$$\gamma = m g(\dot{x}, \delta x). \quad (3.7)$$

The Euler-Lagrange equations, expressing the stationarity of the action, take the local form on the space $N \subset \mathcal{F}$ of solution paths:

$$\delta L + d\gamma = 0 \quad \text{on } N \times \mathbb{R}. \quad (3.8)$$

For the particle on (M, g) with potential V , this abstract equation translates into Newton's second law, expressed in terms of the Levi-Civita covariant derivative ∇ associated with g :

$$m \nabla_{\dot{x}} \dot{x} + \nabla V(x) = 0, \quad (3.9)$$

where ∇V is the gradient vector field of V , defined by $g(\nabla V, Y) = dV(Y)$ for any vector field Y on M .

3.3.2 Transition to Hamiltonian Mechanics

The Lagrangian formalism provides a natural pathway to Hamiltonian mechanics, revealing the underlying symplectic structure of the phase space.

On the space of solutions $N \times \mathbb{R}$, we define the *canonical 2-form* ω as the variation of γ :

Definition 3.9. The canonical 2-form ω is an element of $\Omega^{2,0}(N \times \mathbb{R})$ given by

$$\omega = \delta\gamma. \quad (3.10)$$

In the bicomplex notation, γ is a $(1, 0)$ -form (a 1-form in variations, 0-form in time), and ω is a $(2, 0)$ -form (a 2-form in variations, 0-form in time). This means that for each fixed time t , $\omega_t = \omega|_{N \times \{t\}}$ is a 2-form on the space of solutions N .

Lemma 3.10 (Properties of ω). *The 2-form $\omega = \delta\gamma$ satisfies the following properties on $N \times \mathbb{R}$:*

1. $D\omega = 0$, which implies $\delta\omega = 0$ (i.e., ω is δ -closed) and $d\omega = 0$ (i.e., ω is d -closed).
2. The restriction $\omega_t = \omega|_{N \times \{t\}}$ is independent of time t .
3. ω_t is nondegenerate on N (viewed as a manifold of initial conditions).

Proof.

1. Since $\omega = \delta\gamma$, we have $\delta\omega = \delta^2\gamma = 0$ because $\delta^2 = 0$. For $d\omega$, we use the identity $d\delta + \delta d = 0$ (which follows from $D^2 = (\delta + d)^2 = \delta^2 + d^2 + \delta d + d\delta = 0$) and the Euler-Lagrange equation $\delta L + d\gamma = 0$. $d\omega = d\delta\gamma = -\delta d\gamma$. From $\delta L + d\gamma = 0$, we have $d\gamma = -\delta L$. Thus, $d\omega = -\delta(-\delta L) = \delta^2 L = 0$. Since $D\omega = \delta\omega + d\omega$, we have $D\omega = 0 + 0 = 0$.
2. The condition $d\omega = 0$ means that ω has no dt component in its exterior derivative with respect to time. If we write $\omega = \sum_{i < j} \omega_{ij}(x, \delta x, t) \delta x^i \wedge \delta x^j$ locally (suppressing path indices for N), then $d\omega = \sum_{i < j} \frac{\partial \omega_{ij}}{\partial t} dt \wedge \delta x^i \wedge \delta x^j + \dots$. The vanishing of $d\omega$ implies $\frac{\partial \omega_{ij}}{\partial t} = 0$. Therefore, ω_t is independent of t .
3. For the particle system, $\gamma = m g(\dot{x}, \delta x)$. Then

$$\omega = \delta\gamma = m \delta(g(\dot{x}, \delta x)) = m (g(\delta\dot{x}, \delta x) - g(\dot{x}, \delta^2 x)) = m g(\delta\dot{x}, \delta x) \quad (3.11)$$

assuming a symmetric connection for g such that δ and covariant differentiation commute appropriately, or more directly, $\omega = m(\langle \delta\dot{x} \wedge \delta x \rangle_g)$ if interpreted on phase space.

More fundamentally, identifying the momentum $p_i = m g_{ij} \dot{x}^j$, γ corresponds to $p_i \delta x^i$. Then $\omega = \delta p_i \wedge \delta x^i$, which is the canonical symplectic form on the cotangent bundle T^*M (or a reduction thereof to N). This form is nondegenerate by definition of phase space. \square

Corollary 3.11. *For each $t \in \mathbb{R}$, the restriction $\omega_t = \omega|_{N \times \{t\}}$ is a symplectic form on N . Moreover, this symplectic structure is conserved in time.*

The Hamiltonian function H can be derived from the Lagrangian formalism, often interpreted as the conserved quantity associated with time-translation symmetry (Noether's theorem).

Let ξ be the vector field on $\mathcal{F} \times \mathbb{R}$ that generates the negative of time translation. It is characterized by its action on dt and variations δx : $\iota_\xi dt = -1$, and $\iota_\xi \delta x = \dot{x}$ when ξ acts on γ .

The second condition means that the component of ξ along \mathcal{F} (the space of paths) corresponds to the velocity along the path when contracted with the variation 1-form δx . More precisely, if $\xi = \xi_{\mathcal{F}} - \partial_t$, then $\iota_{\xi_{\mathcal{F}}} \delta x = \dot{x}$.

The Hamiltonian H is defined as a $(0,0)$ -form (a function) on $\mathcal{F} \times \mathbb{R}$:

Definition 3.12. The Hamiltonian function H is given by

$$H = \iota_\xi(L + \gamma). \quad (3.12)$$

Remark 3.13 (Noether's Theorem and the Hamiltonian). *The definition of H in (3.12) is directly motivated by Noether's theorem. If a Lagrangian system is invariant under a continuous symmetry, there is a corresponding conserved quantity. The Hamiltonian is precisely this conserved quantity associated with invariance under time translations. The term $L + \gamma$ can be seen as a component of a more general current.*

Lemma 3.14 (Properties of the Hamiltonian). *On the space of solutions $N \times \mathbb{R}$:*

1. *The Hamiltonian H is conserved in time, i.e., $dH = 0$ (where d is the time derivative component of D).*
2. *The variation of H is related to ω by $\delta H = -\iota_\xi \omega$.*

Proof. We use Cartan's magic formula $\mathcal{L}_\xi = D\iota_\xi + \iota_\xi D$. Thus, $DH = D\iota_\xi(L + \gamma) = \mathcal{L}_\xi(L + \gamma) - \iota_\xi D(L + \gamma)$.

1. For a time-invariant Lagrangian system, $L + \gamma$ is invariant under the flow of ξ (time translation), so $\mathcal{L}_\xi(L + \gamma) = 0$. The Lagrangian 1-form $L = L_0 dt$ has $dL = d(L_0 dt) = (\frac{dL_0}{dt}) dt \wedge dt = 0$. On $N \times \mathbb{R}$, we have $\delta L + d\gamma = 0$. Therefore, $D(L + \gamma) = \delta L + dL + \delta\gamma + d\gamma = (\delta L + d\gamma) + dL + \delta\gamma = 0 + 0 + \omega = \omega$. So, $DH = 0 - \iota_\xi \omega = -\iota_\xi \omega$. Since H is a function (a $(0,0)$ -form), $DH = \delta H + dH$. The form $\omega = \delta\gamma$ is a $(2,0)$ -form. Thus $\iota_\xi \omega$ is a $(1,0)$ -form (it only has a δ -component, no dt component, as $\xi = \xi_{\mathcal{F}} - \partial_t$ and ω has no dt). Specifically, $\iota_\xi \omega = \iota_{\xi_{\mathcal{F}}} \omega - \iota_{\partial_t} \omega = \iota_{\xi_{\mathcal{F}}} \omega$ since ω is type $(2,0)$. Comparing degrees in $DH = -\iota_\xi \omega$: The $(0,1)$ -component gives $dH = 0$, so H is constant in time.
2. The $(1,0)$ -component gives $\delta H = -\iota_\xi \omega$. This is Hamilton's equation in form notation, relating the Hamiltonian to the symplectic form and the generator of time evolution.

□

Example 3.15 (Particle on \mathbb{E}^1). Consider a particle of mass m moving on the Euclidean line \mathbb{E}^1 (so $g(\dot{x}, \dot{x}) = \dot{x}^2$) with potential $V(x)$. The Lagrangian 1-form is $L = (\frac{m}{2}\dot{x}^2 - V(x)) dt$. The variational 1-form is $\gamma = m\dot{x}\delta x$. So, $L + \gamma = (\frac{m}{2}\dot{x}^2 - V(x)) dt + m\dot{x}\delta x$. Using $\iota_\xi dt = -1$ and $\iota_\xi \delta x = \dot{x}$ (meaning $\iota_{\xi_{\mathcal{F}}} \delta x = \dot{x}$):

$$\begin{aligned} H &= \iota_\xi(L + \gamma) \\ &= \iota_\xi \left(\left(\frac{m}{2}\dot{x}^2 - V(x) \right) dt \right) + \iota_\xi(m\dot{x}\delta x) \\ &= - \left(\frac{m}{2}\dot{x}^2 - V(x) \right) + m\dot{x}(\iota_\xi \delta x) \\ &= -\frac{m}{2}\dot{x}^2 + V(x) + m\dot{x}^2 \\ &= \frac{m}{2}\dot{x}^2 + V(x). \end{aligned}$$

This is the familiar expression for the total energy (Hamiltonian) of the particle. Note the sign change of the potential $V(x)$ from the Lagrangian to the Hamiltonian.

3.4 Geometric Interpretation of γ

The variational approach to mechanics has deep connections to differential geometry and provides a powerful framework for understanding the underlying structures of physical theories.

On the space N of classical trajectories, the equation $\delta L = -d\gamma$ holds, which implies $\delta \int_{t_0}^{t_1} L = -(\gamma(t_1) - \gamma(t_0))$. This equation has a deep geometric interpretation in terms of connections on principal bundles.

Consider \mathbb{R} as a Lie group under addition. A connection on a trivial principal \mathbb{R} -bundle over N is a real-valued 1-form on N . For each $t \in \mathbb{R}$, the 1-form $\gamma(t)$ can be interpreted as an \mathbb{R} -connection on $N \times \{t\}$. The equation $\delta \int_{t_0}^{t_1} L = -(\gamma(t_1) - \gamma(t_0))$ implies that addition by $\int_{t_0}^{t_1} L$ provides an isomorphism of trivial principal \mathbb{R} -bundles with connection, mapping the connection $\gamma(t_0)$ to $\gamma(t_1)$. These isomorphisms are compatible for different time intervals.

By taking invariant sections in the \mathbb{R} -direction, we obtain a principal \mathbb{R} -bundle $T \rightarrow N$ equipped with a connection whose curvature is the symplectic form $\omega = \delta\gamma$. This provides a new way to construct the symplectic form on N as the curvature of a connection.

Remark 3.16. *The first step in the path integral formulation of quantum theory involves exponentiating this principal \mathbb{R} -bundle with its connection. Quantization can be intuitively understood as an exponentiation process. The units of γ are the units of action (ML^2/T). To perform the exponentiation, we need a constant with units of action, which is provided by Planck's constant \hbar . Applying the function $\exp(\frac{i}{\hbar} \cdot) : \mathbb{R} \rightarrow \mathbb{C}^\times$, we can form a principal \mathbb{C}^\times -bundle with connection associated to the \mathbb{R} -bundle. The associated complex line bundle with covariant derivative is often called the pre-quantum line bundle. This construction is canonical from a Lagrangian formulation and highlights the additional geometric structure inherent in the Lagrangian formalism.*

Example 3.17 (Particle on a Ring). Consider a particle moving on a ring, which can be modeled as the Euclidean circle $\mathbb{E}^1/2\pi L\mathbb{Z}$ of length $2\pi L$. Let $m > 0$ be the mass of the particle. We define a 1-parameter family of Lagrangian theories parametrized by $\theta \in \mathbb{R}$:

$$L = \left(\frac{m}{2} \dot{x}^2 + \frac{\theta}{2\pi} \dot{x} \right) dt$$

The first term is the usual kinetic energy, and the second term is locally exact but not globally exact on the circle since there is no global coordinate x that covers the entire circle smoothly.

The space N of classical solutions consists of motions with constant velocity. The Euler-Lagrange equation is $m\ddot{x} = 0$, so $x(t) = vt + x_0 \pmod{2\pi L}$. The phase space can be identified with the initial position and velocity, which are $(\theta^1, \theta^2) \in \mathbb{R}/2\pi L\mathbb{Z} \times \mathbb{R}$. The symplectic form can be derived from the Lagrangian.

The Hamiltonian derived from this Lagrangian is independent of θ . The Euler-Lagrange equation is $\frac{d}{dt}(m\dot{x} + \frac{\theta}{2\pi}) - 0 = 0$, so $m\ddot{x} = 0$. The solutions are $x(t) = vt + x_0 \pmod{2\pi L}$. The velocity v is constant. The momentum is $p = mv + \frac{\theta}{2\pi}$. The Hamiltonian is $H = p\dot{x} - L = (mv + \frac{\theta}{2\pi})v - (\frac{m}{2}v^2 + \frac{\theta}{2\pi}v) = \frac{m}{2}v^2 = \frac{1}{2m}(p - \frac{\theta}{2\pi})^2$.

The principal bundle with connection constructed from this Lagrangian depends on θ . The connection 1-form is related to $\gamma = (m\dot{x} + \frac{\theta}{2\pi})\delta x$. The curvature of this connection, which is the symplectic form, should be independent of θ .

The dependence on θ manifests in the corresponding quantum systems, as will be seen in later discussions of quantization.

Remark 3.18. *This example illustrates that different Lagrangian systems can give rise to the same Hamiltonian system. The Hamiltonian $H = \frac{p^2}{2m}$ on the phase space $T^*(\mathbb{R}/2\pi L\mathbb{Z}) \cong (\mathbb{R}/2\pi L\mathbb{Z}) \times \mathbb{R}$ (where p is the momentum conjugate to the angular coordinate) corresponds to free motion on the ring, independent of θ . However, the Lagrangian includes a term that affects the canonical momentum and the prequantum line bundle, leading to different quantum theories for different values of θ . This highlights that Lagrangian systems can contain more information than Hamiltonian systems.*

4 Spectral Theorems

Abstract

We present an introduction to spectral theory, beginning with linear operators on finite-dimensional complex vector spaces and progressing to the analysis of operators on infinite-dimensional Hilbert spaces. Emphasis is placed throughout on the algebraic and geometric structures underlying spectral decompositions.

The discussion starts with diagonalizable operators and their eigenspace decompositions, followed by the Jordan Normal Form for non-diagonalizable operators. These classical results are interpreted via module theory, where a linear operator defines a $\mathbb{C}[x]$ -module structure, and the primary decomposition corresponds to the Jordan structure. A sheaf-theoretic perspective is also briefly introduced, connecting operator spectra to sheaves over \mathbb{C} with support at eigenvalues.

We then introduce projection-valued measures (PVMs) as a framework for describing spectral decompositions of normal operators. For diagonalizable operators, we construct the associated PVM and use it to define a functional calculus $f \mapsto f(A)$, satisfying $A = \int \lambda d\pi_A(\lambda)$. Key properties of PVMs are established, including orthogonality, additivity, and their role in representing operators via spectral integrals.

The theory is extended to commuting families of normal operators on finite-dimensional inner product spaces, leading to the simultaneous diagonalization theorem. The spectral theorems for self-adjoint and unitary operators are formulated using PVMs supported on \mathbb{R} and \mathbb{T} , respectively, and the connection between them via the exponential map is made precise through the functional calculus. We briefly discuss extensions to unitary representations of abelian groups and Pontryagin duality.

In the final part, we turn to spectral theory in infinite-dimensional Hilbert spaces. We address unbounded self-adjoint operators, define their spectral measures, and state the spectral theorem in both the PVM and multiplication operator formulations. Stone's theorem on strongly continuous one-parameter unitary groups is included, connecting self-adjoint operators to unitary dynamics. We conclude with a brief overview of the C^* -algebraic approach for bounded normal operators and the Gelfand-Naimark theorem.

Contents

4.1	Linear Operators in Finite Dimensions	83
4.1.1	Diagonalizable Operators: Eigenspace Decomposition	83
4.1.2	The Jordan Normal Form: Canonical Structure for General Linear Operators	84
4.1.3	Geometric Perspectives: Module Theory and Sheaves	85
4.2	Projection-Valued Measures and Functional Calculus for Diagonalizable Operators	86
4.2.1	Projections, Orthogonal Projections, and Direct Sums	87
4.2.2	Projection-Valued Measures (PVMs) for Diagonalizable Operators	89
4.2.3	Connection to Probability Measures	92
4.2.4	The Functional Calculus for Diagonalizable Operators	96
4.3	Generalizations and Special Cases in Finite Dimensions	100
4.3.1	Simultaneous Diagonalization of Commuting Operators	101
4.3.2	Self-Adjoint Operators and the Spectral Theorem	103
4.3.3	Unitary Operators and the Spectral Theorem	105

4.3.4	Connection via Exponentiation and Unitary Representations	107
4.3.5	Generalization: Unitary Representations of Locally Compact Abelian Groups and Pontryagin Duality	109
4.4	Spectral Theory on Hilbert Space	113
4.4.1	Challenges in Infinite Dimensions: Unboundedness and Spectrum	113
4.4.2	Spectral Measures on Hilbert Space	115
4.4.3	Strongly Continuous One-Parameter Unitary Groups and Stone's Theorem	115
4.4.4	The Spectral Theorem for Self-Adjoint Operators on Hilbert Space	117
4.4.5	Connection to C^* -Algebras and the Functional Calculus for Bounded Normal Operators	119

References: We progress from finite-dimensional operator theory [17][9] to advanced spectral analysis. Key developments include projection-valued measures (PVMs), functional calculus for normal, self-adjoint, and unitary operators [5][16][28], and extensions to unitary representations of locally compact abelian groups via Pontryagin duality, detailed in harmonic analysis texts [12][29]. The core of the advanced material covers spectral theory for unbounded self-adjoint operators on Hilbert spaces, encompassing spectral measures, Stone's theorem on one-parameter unitary groups, and the spectral theorem's PVM and multiplication operator forms [28][30][18]. The chapter concludes with the C^* -algebraic approach to spectral theory for bounded normal operators and the Gelfand-Naimark theorem, explored in [25][6].

4.1 Linear Operators in Finite Dimensions

4.1.1 Diagonalizable Operators: Eigenspace Decomposition

Let W be a finite-dimensional complex vector space of positive dimension. Suppose $A : W \rightarrow W$ is a linear operator.

Definition 4.1. An *eigenspace* of A corresponding to a scalar $\lambda \in \mathbb{C}$ is the subspace

$$W_\lambda := \{\psi \in W : A\psi = \lambda\psi\} = \ker(A - \lambda \text{id}_W).$$

If $W_\lambda \neq \{0\}$, then λ is an *eigenvalue* of A , and any non-zero vector in W_λ is an *eigenvector*.

The algebraic closedness of \mathbb{C} guarantees that every operator A on a finite-dimensional complex vector space has at least one eigenvalue. There is a natural inclusion map from the direct sum of all eigenspaces into W :

$$\bigoplus_{\lambda \in \mathbb{C}} W_\lambda \longrightarrow W. \quad (4.1)$$

Note that the sum is finite, taken only over the eigenvalues $\lambda \in \text{Spec}(A)$.

Definition 4.2. The operator A is *diagonalizable* if the map (4.1) is an isomorphism. Equivalently, A is diagonalizable if W possesses a basis consisting entirely of eigenvectors of A .

When A is diagonalizable, its action is particularly simple: with respect to an eigenbasis, A is represented by a diagonal matrix whose entries are the eigenvalues, each repeated according to the dimension of its eigenspace (the geometric multiplicity). The operator A is determined by this eigenspace decomposition and the corresponding eigenvalues: A acts as multiplication by λ on W_λ .

However, not all operators are diagonalizable. This limitation necessitates a more general structure theory.

Example 4.3. Let $W = \mathbb{C}^3$ with the standard basis e_1, e_2, e_3 . Fix $\lambda \in \mathbb{C}$, and define $A : W \rightarrow W$ to be the operator represented by the matrix

$$\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

The characteristic polynomial is $(t - \lambda)^3$, so λ is the only eigenvalue. The eigenspace $W_\lambda = \ker(A - \lambda \text{id}_W)$ is found by solving

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

which yields $y = 0$ and $z = 0$. Thus, $W_\lambda = \mathbb{C} \cdot e_1$ is one-dimensional. Since $\dim W_\lambda = 1 < 3 = \dim W$, the operator A is not diagonalizable.

4.1.2 The Jordan Normal Form: Canonical Structure for General Linear Operators

When an operator A is not diagonalizable, the Jordan Normal Form (JNF) provides a canonical representation that is “as close to diagonal as possible.”

Definition 4.4. A subspace $W' \subset W$ is *A-invariant* if $A(W') \subset W'$. The operator A is *decomposable* if there exists a direct sum decomposition $W = W' \oplus W''$ into proper A -invariant subspaces. An operator is *indecomposable* if it is not decomposable.

Definition 4.5. A *Jordan block* is an indecomposable linear operator on a finite-dimensional complex vector space.

It is an important result that any Jordan block, in some basis, is represented by a matrix of the form (4.3) (allowing for different sizes $k \times k$). The scalar λ appearing on the diagonal is called the *generalized eigenvalue* of the Jordan block. If $\dim W = 1$, a Jordan block is simply multiplication by λ on the one-dimensional eigenspace. For the operator in Example 4.3, the domain of the map (4.1) is 1-dimensional, while the codomain W is 3-dimensional.

The structure of a Jordan block is tied to nilpotency. Let $A : W \rightarrow W$ be a Jordan block of size $k = \dim W$ with generalized eigenvalue λ . The operator $N := A - \lambda \text{id}_W$ is *nilpotent*, meaning $N^m = 0$ for some integer m . Specifically for a $k \times k$ Jordan block, $N^{k-1} \neq 0$ but $N^k = 0$. The kernels of successive powers of N form a *full flag* in W , which is a filtration by subspaces of increasing dimension:

$$\{0\} = W^{(0)} \subset W^{(1)} \subset W^{(2)} \subset \dots \subset W^{(k)} = W, \quad (4.2)$$

where $W^{(i)} := \ker(N^i)$ and $\dim W^{(i)} = i$ for $0 \leq i \leq k$. The line $W^{(1)} = \ker(N)$ is the unique 1-dimensional eigenspace of A with eigenvalue λ . Vectors in $W^{(i)} \setminus W^{(i-1)}$ for $i > 1$ are called *generalized eigenvectors*.

Theorem 4.6. *Let W be a finite-dimensional complex vector space, and suppose $A : W \rightarrow W$ is a linear operator. Then there exists an A -invariant decomposition*

$$W = \bigoplus_{i=1}^m W_i$$

where each W_i is a subspace on which the restriction $A|_{W_i}$ is a Jordan block. This decomposition is unique up to reordering the summands W_i . The generalized eigenvalues $\lambda_1, \dots, \lambda_m$ associated with the blocks $A|_{W_i}$ are the eigenvalues of A (possibly repeated). The operator A is diagonalizable if and only if $\dim W_i = 1$ for all i .

Proof. See standard linear algebra texts. □

4.1.3 Geometric Perspectives: Module Theory and Sheaves

The structure encapsulated by the Jordan Normal Form can be studied through the lens of algebraic geometry.

The Operator as a $\mathbb{C}[x]$ -Module: A linear operator $A : W \rightarrow W$ endows the vector space W with the structure of a module over the polynomial ring $R = \mathbb{C}[x]$. The action of a polynomial $p(x) = \sum c_j x^j \in R$ on a vector $w \in W$ is defined via $p(x) \cdot w := p(A)w = \sum c_j A^j w$. Since W is finite-dimensional over \mathbb{C} , it becomes a finitely generated module over $R = \mathbb{C}[x]$.

Importantly, $\mathbb{C}[x]$ is a Principal Ideal Domain (PID). The structure theorem for finitely generated modules over a PID states that any such module M is isomorphic to a direct sum of cyclic modules of the form R/I_j , where each ideal I_j is generated by a power of an irreducible element, $I_j = (p_j^{\alpha_j})$. In $R = \mathbb{C}[x]$, the irreducible polynomials are the linear factors $(x - \lambda)$ for $\lambda \in \mathbb{C}$. Thus, the R -module W associated with A decomposes uniquely (up to order) as:

$$W \cong \bigoplus_{i=1}^m \frac{\mathbb{C}[x]}{((x - \lambda_i)^{\alpha_i})}$$

Each cyclic summand $W_i \cong \mathbb{C}[x]/((x - \lambda_i)^{\alpha_i})$ corresponds precisely to an A -invariant subspace associated with a Jordan block $J_{\alpha_i}(\lambda_i)$ in the Jordan Normal Form of A . The action of x (representing A) on a standard basis for this cyclic module yields the Jordan matrix $J_{\alpha_i}(\lambda_i)$. This demonstrates how the JNF arises naturally from the study of modules.

Sheaf Interpretation: In algebraic geometry, we can associate to a commutative ring R its prime spectrum¹⁰, $\text{Spec } R$. For $R = \mathbb{C}[x]$, $\text{Spec } R$ is identified with the complex affine line \mathbb{C} , where each point $\lambda \in \mathbb{C}$ corresponds to the maximal ideal $(x - \lambda)$. An R -module M determines a sheaf¹¹ \mathcal{F}_M over $\text{Spec } R$. For the linear operator A , the corresponding $\mathbb{C}[x]$ -module W gives rise to a sheaf $\mathcal{F}_A \rightarrow \mathbb{C}$.

The *support* of this sheaf, the set of points $\lambda \in \mathbb{C}$ where the stalk $\mathcal{F}_{A,\lambda}$ is non-zero, is precisely the set of eigenvalues of A , $\text{Spec}(A)$. For any open set $E \subset \mathbb{C}$, the sections of \mathcal{F}_A over E form the vector space

$$\Gamma_E(\mathcal{F}_A) = \bigoplus_{\lambda_i \in E \cap \text{Spec}(A)} W_i, \quad (4.3)$$

¹⁰The *prime spectrum* of R is the set $\text{Spec}(R) := \{\mathfrak{p} \subseteq R \mid \mathfrak{p} \text{ is a prime ideal}\}$.

¹¹We discuss sheaves in a later lecture; they are only used in passing here.

where W_i is the subspace corresponding to the i -th Jordan block (associated with eigenvalue λ_i) in the decomposition (4.6). This is depicted in Figure 1.

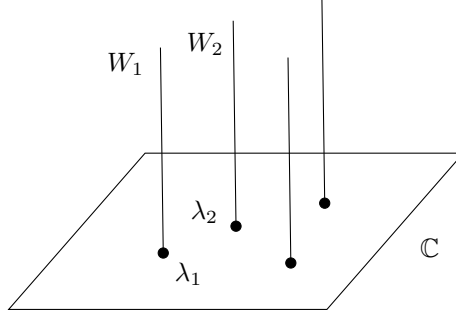


Figure 1: The sheaf $\mathcal{F}_A \rightarrow \mathbb{C}$ of the linear operator $A : W \rightarrow W$. The support is the finite set $\text{Spec}(A) = \{\lambda_1, \dots, \lambda_m\}$. The stalk over λ_i corresponds to the generalized eigenspace W_i .

Remark 4.7. If A is diagonalizable, \mathcal{F}_A is a skyscraper sheaf, supported only at the eigenvalues, with the stalk at λ_i being the eigenspace W_{λ_i} . However, the decomposition (4.6) depicted in Figure 1 does not determine the operator A in general.

If there is a Jordan block of dimension $\alpha_i > 1$, the structure of the sheaf \mathcal{F}_A in a formal (infinitesimal) neighborhood of the corresponding eigenvalue λ_i encodes the nilpotent action $N = A - \lambda_i \text{id}_W$ on the generalized eigenspace W_i . This nilpotent structure, reflected in the flag (4.2), is captured by the non-trivial structure of the stalk $\mathcal{F}_{A, \lambda_i} \cong W_i$ as a module over the local ring $\mathcal{O}_{\mathbb{C}, \lambda_i}$ or its completion. The nilpotency is geometrically represented by this infinitesimal sheaf-theoretic behavior.

4.2 Projection-Valued Measures and Functional Calculus for Diagonalizable Operators

Henceforth in this section, we restrict our attention exclusively to linear operators $A : W \rightarrow W$ that are **diagonalizable** on a finite-dimensional complex vector space W . Recall that an operator A is diagonalizable if and only if W admits a basis consisting of eigenvectors of A , or equivalently, if its minimal polynomial has distinct roots. As we saw earlier, this is also equivalent to the geometric condition that the vector space W decomposes as the direct sum of the eigenspaces of A . Let $\text{Spec}(A) = \{\lambda_1, \dots, \lambda_m\} \subset \mathbb{C}$ denote the spectrum of A (the set of distinct eigenvalues), and let $W_{\lambda_j} = \text{Ker}(A - \lambda_j \text{id}_W)$ be the eigenspace corresponding to the eigenvalue λ_j . By the definition of diagonalizability, the eigenspace decomposition provides an isomorphism:

$$W \cong \bigoplus_{\lambda \in \text{Spec}(A)} W_{\lambda}.$$

The action of the operator A is completely determined by this decomposition and the associated eigenvalues: for any vector $w \in W$, if we decompose it uniquely as $w = \sum_{\lambda_j \in \text{Spec}(A)} w_j$ with $w_j \in W_{\lambda_j}$, then $A(w) = \sum_{\lambda_j \in \text{Spec}(A)} A(w_j) = \sum_{\lambda_j \in \text{Spec}(A)} \lambda_j w_j$. Thus, A simply acts as scalar multiplication by λ_j on the subspace W_{λ_j} . This picture is captured by the interpretation of A as the collection of its actions on the “stalks” W_{λ} over the discrete space $\text{Spec}(A)$.

While intuitive, this description relies on the direct sum structure. We now develop an equivalent and powerful formalism that encodes such “multiplication operators” purely in terms of projection

operators associated with the eigenspaces. This projection-valued measure approach offers significant advantages, particularly for generalization.

4.2.1 Projections, Orthogonal Projections, and Direct Sums

We begin by recalling and elaborating on the important properties of projection operators, which are the building blocks of spectral measures.

Definition 4.8 (Projection Operator). A linear operator $P : W \rightarrow W$ on a vector space W is called a **projection** (or idempotent operator) if it satisfies the condition $P^2 = P$. The set of all projection operators on W is denoted by $\text{Proj}(W)$.

Projections are linked to direct sum decompositions of the vector space.

Proposition 4.9 (Projections and Direct Sums Correspondence).

1. If $P \in \text{Proj}(W)$, then W decomposes as the direct sum of the image (range) and kernel (null space) of P :

$$W = \text{Im}(P) \oplus \text{Ker}(P). \quad (4.4)$$

Furthermore, P acts as the identity on $\text{Im}(P)$ and as the zero operator on $\text{Ker}(P)$. The operator $Q = \text{id}_W - P$ is also a projection, called the complementary projection, satisfying $\text{Im}(Q) = \text{Ker}(P)$, $\text{Ker}(Q) = \text{Im}(P)$, and $PQ = QP = 0$.

2. Conversely, given any direct sum decomposition $W = W'' \oplus W'$ into subspaces W'' and W' , there exists a unique projection operator $P : W \rightarrow W$ such that $\text{Im}(P) = W''$ and $\text{Ker}(P) = W'$. This operator is called the **projection onto W'' along W'** , and we denote it by P'' . That is,

$$P'' : W \longrightarrow W, \quad \text{with } \text{Im}(P'') = W'', \text{ and } \text{Ker}(P'') = W'. \quad (4.5)$$

Proof.

1. For any $w \in W$, we can write

$$w = Pw + (w - Pw).$$

Clearly, $Pw \in \text{Im}(P)$. Let $w' = w - Pw$. Then

$$Pw' = P(w - Pw) = Pw - P^2w = Pw - Pw = 0,$$

so $w' \in \text{Ker}(P)$. This shows that

$$W = \text{Im}(P) + \text{Ker}(P).$$

To show the sum is direct, let $v \in \text{Im}(P) \cap \text{Ker}(P)$. Since $v \in \text{Im}(P)$, there exists some $u \in W$ such that $v = Pu$. Since $v \in \text{Ker}(P)$, we also have $Pv = 0$. Therefore,

$$P^2u = 0,$$

which implies $Pu = 0$, since $P^2 = P$. Hence, $v = 0$, proving the sum is direct.

If $v \in \text{Im}(P)$, then $v = Pu$, so $Pv = P^2u = Pu = v$. If $v \in \text{Ker}(P)$, then $Pv = 0$.

For $Q = \text{id}_W - P$, we compute

$$Q^2 = (\text{id}_W - P)(\text{id}_W - P) = \text{id}_W - 2P + P^2 = \text{id}_W - 2P + P = \text{id}_W - P = Q.$$

Thus, $Q^2 = Q$.

Next, we find the kernel and image of Q :

$$\text{Ker}(Q) = \{w \mid (\text{id}_W - P)w = 0\} = \{w \mid w = Pw\} = \text{Im}(P),$$

and

$$\text{Im}(Q) = \{(\text{id}_W - P)w \mid w \in W\} = \text{Ker}(P),$$

as shown in the decomposition. Moreover, we have

$$PQ = P(\text{id}_W - P) = P - P^2 = 0,$$

and

$$QP = (\text{id}_W - P)P = P - P^2 = 0.$$

2. Given $W = W'' \oplus W'$, any $w \in W$ has a unique decomposition

$$w = w'' + w',$$

where $w'' \in W''$ and $w' \in W'$. Define $P(w) = w''$. This map is clearly linear. Its image is W'' , and its kernel is

$$\{w'' + w' \mid w'' = 0\} = W'.$$

We now check idempotency:

$$P^2(w) = P(w'') = w'' = P(w),$$

since the decomposition of $w'' \in W''$ is $w'' = w'' + 0$.

Uniqueness follows because any projection P' with $\text{Im}(P') = W''$ and $\text{Ker}(P') = W'$ must map $w = w'' + w'$ to $P'(w'') + P'(w')$. Since $w'' \in \text{Im}(P')$, we have $P'(w'') = w''$. Since $w' \in \text{Ker}(P')$, we have $P'(w') = 0$. Thus,

$$P'(w) = w'' = P(w).$$

□

When W possesses an inner product structure $\langle \cdot, \cdot \rangle$, a special class of projections arises, characterized by compatibility with the inner product.

Definition 4.10 (Orthogonal Projection). Let $(W, \langle \cdot, \cdot \rangle)$ be a finite-dimensional complex inner product space (i.e., a Hilbert space). A projection $P \in \text{Proj}(W)$ is called an **orthogonal projection** if its image and kernel are orthogonal subspaces with respect to the inner product: $\text{Im}(P) \perp \text{Ker}(P)$. That is, $\langle Pw_1, (\text{id}_W - P)w_2 \rangle = 0$ for all $w_1, w_2 \in W$.

Orthogonal projections admit a important algebraic characterization involving the adjoint operator P^* , defined by $\langle Pw, v \rangle = \langle w, P^*v \rangle$ for all $v, w \in W$.

Proposition 4.11 (Characterization of Orthogonal Projections). *A linear operator $P : W \rightarrow W$ on a finite-dimensional Hilbert space W is an orthogonal projection if and only if it satisfies both $P^2 = P$ (idempotency) and $P^* = P$ (self-adjointness).*

Proof. (\Rightarrow) Assume that P is an orthogonal projection. We already know that $P^2 = P$. We need to show that $P^* = P$.

Since $W = \text{Im}(P) \oplus \text{Ker}(P)$ is an orthogonal decomposition, any $w, v \in W$ can be uniquely written as:

$$w = w_1 + w_2 \quad \text{and} \quad v = v_1 + v_2,$$

where $w_1, v_1 \in \text{Im}(P)$ and $w_2, v_2 \in \text{Ker}(P)$.

Therefore, we have:

$$Pw = w_1 \quad \text{and} \quad Pv = v_1.$$

We now compute the inner products:

$$\langle Pw, v \rangle = \langle w_1, v_1 + v_2 \rangle = \langle w_1, v_1 \rangle + \langle w_1, v_2 \rangle = \langle w_1, v_1 \rangle,$$

since $v_2 \in \text{Ker}(P)$ and is orthogonal to $w_1 \in \text{Im}(P)$.

Similarly:

$$\langle w, Pv \rangle = \langle w_1 + w_2, v_1 \rangle = \langle w_1, v_1 \rangle + \langle w_2, v_1 \rangle = \langle w_1, v_1 \rangle,$$

since $w_2 \in \text{Ker}(P)$ and is orthogonal to $v_1 \in \text{Im}(P)$.

Since we have:

$$\langle Pw, v \rangle = \langle w, Pv \rangle \quad \text{for all } w, v \in W,$$

it follows that $P^* = P$.

(\Leftarrow) Now assume that $P^2 = P$ and $P^* = P$. We know that $W = \text{Im}(P) \oplus \text{Ker}(P)$. Let $u \in \text{Im}(P)$ and $v \in \text{Ker}(P)$. We want to show that $\langle u, v \rangle = 0$.

Since $u \in \text{Im}(P)$, we have $u = Pu$. Since $v \in \text{Ker}(P)$, we have $Pv = 0$. Therefore:

$$\langle u, v \rangle = \langle Pu, v \rangle = \langle u, P^*v \rangle \quad (\text{by the definition of adjoint}).$$

Since $P^* = P$, we obtain:

$$\langle u, v \rangle = \langle u, Pv \rangle.$$

Since $Pv = 0$, it follows that:

$$\langle u, v \rangle = \langle u, 0 \rangle = 0.$$

Thus, $\text{Im}(P) \perp \text{Ker}(P)$, and therefore, P is an orthogonal projection. □

4.2.2 Projection-Valued Measures (PVMs) for Diagonalizable Operators

We now return to our setting of a diagonalizable operator $A : W \rightarrow W$ on a finite-dimensional complex vector space W , with the eigenspace decomposition $W = \bigoplus_{\lambda \in \text{Spec}(A)} W_\lambda$. Let $\pi_{\lambda_j} \in \text{Proj}(W)$ denote the unique projection onto the eigenspace W_{λ_j} along the direct sum of the other eigenspaces $\bigoplus_{\lambda_k \neq \lambda_j} W_{\lambda_k}$. These individual projections satisfy important orthogonality relations.

Lemma 4.12 (Properties of Eigenspace Projections). *Let A be diagonalizable with distinct eigenvalues $\text{Spec}(A) = \{\lambda_1, \dots, \lambda_m\}$, and let π_{λ_j} be the projection onto W_{λ_j} along $\bigoplus_{k \neq j} W_{\lambda_k}$. Then:*

1. $\sum_{j=1}^m \pi_{\lambda_j} = \text{id}_W$ (Resolution of Identity).
2. $\pi_{\lambda_j} \pi_{\lambda_k} = \delta_{jk} \pi_{\lambda_j}$ (Orthogonality of Projections).

Proof.

(1) Any $w \in W$ decomposes uniquely as $w = \sum_{k=1}^m w_k$ with $w_k \in W_{\lambda_k}$. By definition, $\pi_{\lambda_j}(w) = w_j$. Therefore, $(\sum_{j=1}^m \pi_{\lambda_j})(w) = \sum_{j=1}^m \pi_{\lambda_j}(w) = \sum_{j=1}^m w_j = w = \text{id}_W(w)$. Since this holds for all w , $\sum \pi_{\lambda_j} = \text{id}_W$.

(2) Consider $\pi_{\lambda_j} \pi_{\lambda_k}$. The image of π_{λ_k} is W_{λ_k} . If $j \neq k$, then W_{λ_k} is contained in the kernel of π_{λ_j} (since $\text{Ker}(\pi_{\lambda_j}) = \bigoplus_{l \neq j} W_{\lambda_l}$). Therefore, for $j \neq k$, $\pi_{\lambda_j} \pi_{\lambda_k} = 0$. If $j = k$, then $\pi_{\lambda_j} \pi_{\lambda_j} = (\pi_{\lambda_j})^2 = \pi_{\lambda_j}$ since π_{λ_j} is a projection. Combining these, $\pi_{\lambda_j} \pi_{\lambda_k} = \delta_{jk} \pi_{\lambda_j}$. \square

These projections allow us to define a set function whose values are operators, associating subsets of the complex plane to projections onto corresponding sums of eigenspaces.

Definition 4.13 (Projection-Valued Measure (Finite-Dimensional Case)). Let $A : W \rightarrow W$ be a diagonalizable operator on a finite-dimensional complex vector space W . The **projection-valued measure (PVM)** or **spectral measure** associated with A is the map

$$\pi_A : \mathcal{P}(\mathbb{C}) \longrightarrow \text{Proj}(W) \quad (4.6)$$

(where $\mathcal{P}(\mathbb{C})$ is the power set of \mathbb{C}) defined by assigning to each subset $E \subseteq \mathbb{C}$ the projection onto the sum of eigenspaces corresponding to eigenvalues lying within E :

$$E \mapsto \pi_A(E) := \sum_{\lambda_j \in E \cap \text{Spec}(A)} \pi_{\lambda_j}$$

where π_{λ_j} is the projection onto the eigenspace W_{λ_j} along $\bigoplus_{k \neq j} W_{\lambda_k}$. If $E \cap \text{Spec}(A)$ is empty, the sum is empty and defined as the zero operator.

Remark 4.14 (Dependence on Spectrum). *Observe that although the domain is formally written as $\mathcal{P}(\mathbb{C})$, the map π_A is constant on sets E having the same intersection with the finite set $\text{Spec}(A)$. It is non-zero only for sets E that contain at least one eigenvalue of A . Effectively, π_A is supported on $\text{Spec}(A)$.*

Remark 4.15 (Orthogonality for Normal Operators). *If W is equipped with an inner product $\langle \cdot, \cdot \rangle$ and the operator A is **normal** (i.e., A commutes with its adjoint, $AA^* = A^*A$), then A is diagonalizable, and its eigenspaces corresponding to distinct eigenvalues are mutually orthogonal: $W_{\lambda_j} \perp W_{\lambda_k}$ for $j \neq k$. In this case, the direct sum $W = \bigoplus_{\lambda_j \in \text{Spec}(A)} W_{\lambda_j}$ is an orthogonal direct sum. Consequently, each individual projection π_{λ_j} is an orthogonal projection ($\pi_{\lambda_j}^* = \pi_{\lambda_j}$), and the PVM $\pi_A(E) = \sum_{\lambda_j \in E \cap \text{Spec}(A)} \pi_{\lambda_j}$ is also an orthogonal projection for every $E \subseteq \mathbb{C}$. This holds, in particular, if A is self-adjoint ($A^* = A$) or unitary ($A^*A = AA^* = \text{id}_W$).*

The map π_A exhibits properties strongly analogous to those of a classical (scalar-valued) measure, justifying its name.

Proposition 4.16 (Properties of the PVM π_A). *The map $\pi_A : \mathcal{P}(\mathbb{C}) \rightarrow \text{Proj}(W)$ defined above satisfies the following properties:*

(i) **Null Empty Set:** $\pi_A(\emptyset) = 0$.

(ii) **Normalization:** $\pi_A(\mathbb{C}) = \pi_A(\text{Spec}(A)) = \text{id}_W$.

(iii) **Finite Additivity:** If $E_1, E_2 \subseteq \mathbb{C}$ are disjoint subsets, then

$$\pi_A(E_1 \cup E_2) = \pi_A(E_1) + \pi_A(E_2).$$

More generally, for pairwise disjoint E_1, \dots, E_n , $\pi_A(\cup_{i=1}^n E_i) = \sum_{i=1}^n \pi_A(E_i)$.

(iv) **Multiplicativity/Intersection Property:** For any $E_1, E_2 \subseteq \mathbb{C}$,

$$\pi_A(E_1 \cap E_2) = \pi_A(E_1)\pi_A(E_2).$$

These properties imply, in particular, that projections associated with disjoint sets commute and their product is zero: if $E_1 \cap E_2 = \emptyset$, then $\pi_A(E_1)\pi_A(E_2) = \pi_A(E_1 \cap E_2) = \pi_A(\emptyset) = 0$.

Proof. Let $S = \text{Spec}(A)$.

(i) $\pi_A(\emptyset) = \sum_{\lambda \in \emptyset \cap S} \pi_\lambda = 0$ (sum over empty set).

(ii) $\pi_A(\mathbb{C}) = \sum_{\lambda \in \mathbb{C} \cap S} \pi_\lambda = \sum_{\lambda \in S} \pi_\lambda = \text{id}_W$ by Lemma 4.12(1). Similarly for $\pi_A(S)$.

(iii) If $E_1 \cap E_2 = \emptyset$, then $E_1 \cap S$ and $E_2 \cap S$ are disjoint subsets of S . The set of eigenvalues in the union is the disjoint union of the eigenvalues in each set: $(E_1 \cup E_2) \cap S = (E_1 \cap S) \cup (E_2 \cap S)$. Therefore,

$$\pi_A(E_1 \cup E_2) = \sum_{\lambda \in (E_1 \cup E_2) \cap S} \pi_\lambda = \sum_{\lambda \in (E_1 \cap S)} \pi_\lambda + \sum_{\lambda \in (E_2 \cap S)} \pi_\lambda = \pi_A(E_1) + \pi_A(E_2).$$

(iv) Using the definition and the orthogonality property $\pi_\lambda \pi_\mu = \delta_{\lambda\mu} \pi_\lambda$ from Lemma 4.12(2):

$$\begin{aligned} \pi_A(E_1)\pi_A(E_2) &= \left(\sum_{\lambda \in E_1 \cap S} \pi_\lambda \right) \left(\sum_{\mu \in E_2 \cap S} \pi_\mu \right) \\ &= \sum_{\lambda \in E_1 \cap S} \sum_{\mu \in E_2 \cap S} \pi_\lambda \pi_\mu \\ &= \sum_{\lambda \in E_1 \cap S} \sum_{\mu \in E_2 \cap S} \delta_{\lambda\mu} \pi_\lambda \\ &= \sum_{\lambda \in (E_1 \cap S) \cap (E_2 \cap S)} \pi_\lambda \quad (\text{only terms with } \lambda = \mu \text{ survive}) \\ &= \sum_{\lambda \in (E_1 \cap E_2) \cap S} \pi_\lambda \\ &= \pi_A(E_1 \cap E_2). \end{aligned}$$

□

The set of projections $\{\pi_{\lambda_j}\}_{\lambda_j \in \text{Spec}(A)}$ associated with individual eigenvalues forms the core of the PVM. They constitute a complete set of mutually orthogonal (in the sense $\pi_j \pi_k = 0$ for $j \neq k$) projections summing to the identity, often called a **resolution of the identity** associated with A .

4.2.3 Connection to Probability Measures

The designation of the map $\pi_A : \mathcal{P}(\mathbb{C}) \longrightarrow \text{Proj}(W)$ as a projection valued measure invites a comparison to the more familiar concept of classical, scalar-valued measures, particularly probability measures. While PVMs are operator-valued, their mathematical properties enable the construction of classical probability measures associated with vectors in the underlying space, thereby providing a general framework.

We first recall the basics of classical measure theory and probability spaces.

Definition 4.17 (Measurable Space and σ -Algebra). A **measurable space** is a pair (X, Σ) , where X is a non-empty set, often called the “sample space” or “universe of outcomes,” and Σ is a **σ -algebra** (or σ -field) of subsets of X . Σ is a collection of subsets of X satisfying:

1. $X \in \Sigma$ (the entire space is measurable).
2. If $E \in \Sigma$, then its complement $X \setminus E$ is also in Σ (closure under complementation).
3. If $\{E_i\}_{i=1}^\infty$ is a countable collection of sets in Σ , then their union $\cup_{i=1}^\infty E_i$ is in Σ (closure under countable unions).

The elements of Σ are called “measurable sets” or, in probabilistic contexts, “events.”

Definition 4.18 (Positive Measure). A **positive measure** on a measurable space (X, Σ) is a function $\mu : \Sigma \rightarrow [0, \infty]$ (the set of non-negative extended real numbers) such that:

1. $\mu(\emptyset) = 0$ (the measure of the empty set is zero).
2. (Countable Additivity / σ -additivity) For any countable sequence $\{E_i\}_{i=1}^\infty$ of pairwise disjoint sets in Σ (i.e., $E_i \cap E_j = \emptyset$ for $i \neq j$),

$$\mu\left(\bigcup_{i=1}^\infty E_i\right) = \sum_{i=1}^\infty \mu(E_i).$$

From these axioms, several standard properties of measures follow, including:

- Finite Additivity: For any finite collection of pairwise disjoint sets $E_1, \dots, E_n \in \Sigma$, $\mu(\cup_{k=1}^n E_k) = \sum_{k=1}^n \mu(E_k)$.
- Monotonicity: If $E, F \in \Sigma$ and $E \subseteq F$, then $\mu(E) \leq \mu(F)$.
- Subadditivity: For any countable collection $\{E_i\}_{i=1}^\infty$ of sets in Σ (not necessarily disjoint), $\mu(\cup_{i=1}^\infty E_i) \leq \sum_{i=1}^\infty \mu(E_i)$.
- Continuity from below: If $E_1 \subseteq E_2 \subseteq \dots$ is an increasing sequence of sets in Σ and $E = \cup_{n=1}^\infty E_n$, then $\mu(E) = \lim_{n \rightarrow \infty} \mu(E_n)$.
- Continuity from above: If $E_1 \supseteq E_2 \supseteq \dots$ is a decreasing sequence of sets in Σ , $E = \cap_{n=1}^\infty E_n$, and $\mu(E_1) < \infty$, then $\mu(E) = \lim_{n \rightarrow \infty} \mu(E_n)$.

Definition 4.19 (Finite Measures and Probability Measures).

1. A positive measure μ is **finite** if $\mu(X) < \infty$.

2. A positive measure P on (X, Σ) is a **probability measure** if $P(X) = 1$. The triple (X, Σ, P) is then termed a **probability space**.

More generally, one defines signed measures (real-valued) and complex measures, both satisfying countable additivity.

In the context of a PVM π_A for a diagonalizable operator A on a finite-dimensional space W , the spectrum $\text{Spec}(A)$ is a finite set. We can take $X = \text{Spec}(A)$ and $\Sigma = \mathcal{P}(\text{Spec}(A))$ (the power set of the spectrum). In this scenario, any finitely additive set function automatically satisfies countable additivity because any countable disjoint union can only contain finitely many non-empty sets.

The PVM π_A , as defined in Definition 4.13 has many parallels to the axiomatic structure of a probability measure, especially concerning its behavior under set operations (Proposition 4.16).

- (i) **Null Empty Set:** $\pi_A(\emptyset) = 0$. This is structurally identical to $P(\emptyset) = 0$. The PVM assigns the zero operator to the empty set of spectral values, signifying that the “spectral content” associated with no eigenvalues is trivial (projection onto the zero-dimensional subspace $\{0_W\}$).
- (ii) **Normalization / Total Measure:** $\pi_A(\mathbb{C}) = \pi_A(\text{Spec}(A)) = \text{id}_W$. This parallels $P(X) = 1$. The identity operator id_W serves as the operatorial analogue of “total probability” or “certainty.” It confirms that the entire space W is accounted for by the spectral decomposition $W = \bigoplus_{\lambda \in \text{Spec}(A)} W_\lambda$. The sum of projections onto all distinct eigenspaces, $\sum_{\lambda \in \text{Spec}(A)} \pi_\lambda = \text{id}_W$ (see Lemma 4.12).
- (iii) **Finite Additivity:** $\pi_A(E_1 \cup E_2) = \pi_A(E_1) + \pi_A(E_2)$ **for disjoint** $E_1, E_2 \subseteq \mathbb{C}$. This mirrors the additivity $P(E_1 \cup E_2) = P(E_1) + P(E_2)$ for mutually exclusive events. The PVM property follows from its definition as a sum over individual eigenspace projections:

$$\pi_A(E_1 \cup E_2) = \sum_{\lambda \in (E_1 \cap S) \cup (E_2 \cap S)} \pi_\lambda = \sum_{\lambda \in E_1 \cap S} \pi_\lambda + \sum_{\lambda \in E_2 \cap S} \pi_\lambda = \pi_A(E_1) + \pi_A(E_2),$$

where $S = \text{Spec}(A)$ and the disjointness of E_1, E_2 implies disjointness of $E_1 \cap S$ and $E_2 \cap S$. The sum is one of linear operators. For this sum $\pi_A(E_1) + \pi_A(E_2)$ to itself be a projection (which $\pi_A(E_1 \cup E_2)$ is), a necessary and sufficient condition is that $\pi_A(E_1)\pi_A(E_2) = 0$ and $\pi_A(E_2)\pi_A(E_1) = 0$. This condition is indeed satisfied due to property (iv) below, as $E_1 \cap E_2 = \emptyset \implies \pi_A(E_1)\pi_A(E_2) = \pi_A(\emptyset) = 0$. The image of $\pi_A(E_1 \cup E_2)$ is then the direct sum of the images of $\pi_A(E_1)$ and $\pi_A(E_2)$.

The major difference from classical probability measures is that $\pi_A(E)$ is an operator (a projection) and not a scalar value in $[0, 1]$. However, this operator-valued structure provides a way to generate classical scalar measures, including probability measures, when an inner product and specific vectors (states) are introduced. This is particularly transparent for normal operators, where the PVM consists of orthogonal projections.

Let W be a finite-dimensional complex inner product space, and let A be a normal operator on W . Then its PVM $\pi_A(E)$ consists of orthogonal projections (i.e., $\pi_A(E)^* = \pi_A(E)$ and $\pi_A(E)^2 = \pi_A(E)$). For any vector $\psi \in W$, consider the map $\mu_\psi : \mathcal{P}(\text{Spec}(A)) \rightarrow [0, \infty)$ defined by:

$$\mu_\psi(E) := \langle \psi, \pi_A(E)\psi \rangle.$$

Since $\pi_A(E)$ is an orthogonal projection,

$$\begin{aligned}\langle \psi, \pi_A(E)\psi \rangle &= \langle \psi, \pi_A(E)^* \pi_A(E)\psi \rangle \\ &= \langle \pi_A(E)\psi, \pi_A(E)\psi \rangle \\ &= \|\pi_A(E)\psi\|^2 \geq 0.\end{aligned}$$

This map μ_ψ is a finite positive measure on $\text{Spec}(A)$:

1. $\mu_\psi(\emptyset) = \|\pi_A(\emptyset)\psi\|^2 = \|0\psi\|^2 = 0$.
2. (Finite Additivity) For disjoint $E_1, E_2 \subseteq \text{Spec}(A)$: The projections $\pi_A(E_1)$ and $\pi_A(E_2)$ are orthogonal and their images are orthogonal subspaces because $E_1 \cap E_2 = \emptyset \implies \pi_A(E_1)\pi_A(E_2) = 0$. This means that $\pi_A(E_1)\psi$ is orthogonal to $\pi_A(E_2)\psi$. Thus,

$$\begin{aligned}\mu_\psi(E_1 \cup E_2) &= \|\pi_A(E_1 \cup E_2)\psi\|^2 = \|(\pi_A(E_1) + \pi_A(E_2))\psi\|^2 \\ &= \|\pi_A(E_1)\psi + \pi_A(E_2)\psi\|^2 \\ &= \|\pi_A(E_1)\psi\|^2 + \|\pi_A(E_2)\psi\|^2 \quad (\text{by the Pythagorean theorem}) \\ &= \mu_\psi(E_1) + \mu_\psi(E_2).\end{aligned}$$

The total measure is $\mu_\psi(\text{Spec}(A)) = \|\pi_A(\text{Spec}(A))\psi\|^2 = \|\text{id}_W \psi\|^2 = \|\psi\|^2$. Crucially, if the vector ψ is normalized such that $\|\psi\| = 1$, then $\mu_\psi(\text{Spec}(A)) = 1$. In this case, μ_ψ becomes a classical **probability measure** on the finite set $\text{Spec}(A)$. Thus, the PVM π_A , while operator-valued, naturally induces a family of probability measures, parameterized by the choice of a normalized vector ψ . Each such μ_ψ describes a probability distribution over the possible spectral values of A .

More generally, for any pair of vectors $\phi, \psi \in W$, the function $\mu_{\phi, \psi}(E) := \langle \phi, \pi_A(E)\psi \rangle$ defines a complex measure on $\text{Spec}(A)$. These are essential for defining the integral $\langle \phi, f(A)\psi \rangle = \int f(\lambda) d\mu_{\phi, \psi}(\lambda)$ for complex-valued functions f .

Once we have the probability measure μ_ψ associated with a normalized vector ψ and a self-adjoint operator A (whose spectrum is real), we can define the expectation of a function $f : \text{Spec}(A) \rightarrow \mathbb{R}$ in the classical sense:

$$\mathbb{E}_{\mu_\psi}[f] := \int_{\text{Spec}(A)} f(\lambda) d\mu_\psi(\lambda) = \sum_{\lambda_j \in \text{Spec}(A)} f(\lambda_j) \mu_\psi(\{\lambda_j\}).$$

Using the functional calculus $f(A) = \sum_{\lambda_j \in \text{Spec}(A)} f(\lambda_j) \pi_{\lambda_j}$ and $\mu_\psi(\{\lambda_j\}) = \langle \psi, \pi_{\lambda_j} \psi \rangle = \|\pi_{\lambda_j} \psi\|^2$, we find:

$$\begin{aligned}\langle \psi, f(A)\psi \rangle &= \left\langle \psi, \sum_{\lambda_j \in \text{Spec}(A)} f(\lambda_j) \pi_{\lambda_j} \psi \right\rangle \\ &= \sum_{\lambda_j \in \text{Spec}(A)} f(\lambda_j) \langle \psi, \pi_{\lambda_j} \psi \rangle \\ &= \sum_{\lambda_j \in \text{Spec}(A)} f(\lambda_j) \mu_\psi(\{\lambda_j\}) \\ &= \mathbb{E}_{\mu_\psi}[f].\end{aligned}$$

In particular, for $f(\lambda) = \lambda$, the “average value” $\langle \psi, A\psi \rangle$ is precisely the expectation $\mathbb{E}_{\mu_\psi}[\lambda]$, where λ is viewed as a random variable distributed according to μ_ψ . Similarly, the variance can be defined:

$$\text{Var}_{\mu_\psi}(\lambda) = \mathbb{E}_{\mu_\psi}[(\lambda - \mathbb{E}_{\mu_\psi}[\lambda])^2] = \langle \psi, (A - \langle \psi, A\psi \rangle \text{id}_W)^2 \psi \rangle.$$

This shows how the operatorial PVM framework connects directly to classical statistical concepts through the vector-dependent probability measures.

The PVM property (iv) $\pi_A(E_1 \cap E_2) = \pi_A(E_1)\pi_A(E_2)$ is another major difference from the properties of typical probability measures. For probabilities, $P(E_1 \cap E_2) = P(E_1)P(E_2)$ holds if and only if the events E_1 and E_2 are **statistically independent**. Independence is a specific relationship between events contingent on a particular probability measure P , not an axiom of probability measures themselves.

In contrast, for a PVM π_A derived from a single operator A , the multiplicative property $\pi_A(E_1 \cap E_2) = \pi_A(E_1)\pi_A(E_2)$ is an inherent structural property. It follows from the idempotency and mutual orthogonality of the elementary spectral projections $\pi_\lambda \pi_\mu = \delta_{\lambda\mu} \pi_\lambda$. Since all $\pi_A(E)$ are sums of these π_λ , they all commute: $\pi_A(E_1)\pi_A(E_2) = \pi_A(E_2)\pi_A(E_1)$. This commutativity of projection values for any two sets is a key feature. The property reflects the underlying logic of spectral decomposition: applying a “filter” for eigenvalues in E_2 and then a “filter” for eigenvalues in E_1 is equivalent to applying a single “filter” for eigenvalues in $E_1 \cap E_2$. This robust algebraic property is essential for ensuring that the functional calculus $f \mapsto f(A)$ defines an algebra homomorphism (specifically, $(fg)(A) = f(A)g(A)$).

The true power of this measure-theoretic perspective, especially its connection to probability, becomes fully apparent in the context of infinite-dimensional Hilbert spaces and operators with continuous spectra (as discussed in Section 4.4).

1. The PVM π_A associated with a self-adjoint operator A is defined on the Borel σ -algebra $\mathcal{B}(\sigma(A))$ of its spectrum.
2. The crucial property is **countable additivity (strong operator topology-convergence)**: for any sequence $\{E_i\}_{i=1}^\infty$ of pairwise disjoint Borel sets,

$$\pi_A\left(\bigcup_{i=1}^\infty E_i\right)\psi = \sum_{i=1}^\infty \pi_A(E_i)\psi \quad (\text{for all } \psi \in \mathcal{H}).$$

This SOT-countable additivity of the PVM π_A is precisely the condition required to ensure that each derived scalar function $\mu_\psi(E) = \|\pi_A(E)\psi\|^2$ is a true, countably additive probability measure on $(\sigma(A), \mathcal{B}(\sigma(A)))$ when $\|\psi\| = 1$. This mathematical structure supports the probabilistic interpretation of quantum measurements, even for observables with continuous spectra like position or momentum, where sums turn into integrals against probability measures determined by the quantum state.

To summarize, while PVMs are fundamentally operator-valued, they share core axiomatic traits with scalar measures, particularly concerning normalization and additivity for disjoint sets. For normal operators on an inner product space, the PVM, being composed of orthogonal projections, naturally allows the construction of a family of classical positive scalar measures $\mu_\psi(E) = \|\pi_A(E)\psi\|^2$. When the vector ψ is normalized, μ_ψ becomes a probability measure on the spectrum of A . This provides a purely mathematical pathway to probabilistic interpretations. The expectation values formed with respect to these probability measures align perfectly with the inner product expressions like

$\langle \psi, f(A)\psi \rangle$. The distinctive multiplicative property of PVMs, $\pi_A(E_1 \cap E_2) = \pi_A(E_1)\pi_A(E_2)$, is an algebraic feature that has no direct counterpart in general probability theory beyond the special case of statistical independence; for PVMs, it is important for the homomorphism property of the functional calculus. Overall, PVMs offer a mathematical framework in which operator-valued measures can produce classical probability distributions when applied to specific states.

4.2.4 The Functional Calculus for Diagonalizable Operators

The primary purpose of the PVM formalism is to provide a unified way to define functions of the operator A . This is known as the functional calculus. Applying functions to operators naturally arises in various contexts, such as solving systems of linear differential equations (e^{tA}), defining norms or condition numbers of matrices, and formulating observables and their evolution in quantum mechanics.

First, the operator A itself can be reconstructed from its eigenvalues and the corresponding projections. Since A acts as $\lambda_j \text{id}_{W_{\lambda_j}}$ on the image of π_{λ_j} , and $\sum \pi_{\lambda_j} = \text{id}_W$, we can write A as a linear combination of these projections, weighted by the eigenvalues.

Proposition 4.20 (Spectral Decomposition of A). *Let A be a diagonalizable operator with distinct eigenvalues $\text{Spec}(A) = \{\lambda_1, \dots, \lambda_m\}$ and corresponding eigenspace projections π_{λ_j} . Then A can be expressed as:*

$$A = \sum_{j=1}^m \lambda_j \pi_{\lambda_j}. \quad (4.7)$$

This is often written suggestively as an integral with respect to the PVM:

$$A = \int_{\mathbb{C}} \lambda d\pi_A(\lambda) \quad \text{or} \quad A = \int_{\text{Spec}(A)} \lambda d\pi_A(\lambda).$$

Proof. Let $w \in W$. Decompose $w = \sum_{k=1}^m w_k$ where

$$w_k = \pi_{\lambda_k}(w) \in W_{\lambda_k}.$$

Then

$$A(w) = \sum_{k=1}^m A(w_k) = \sum_{k=1}^m \lambda_k w_k.$$

Now compute the action of the sum:

$$\left(\sum_{j=1}^m \lambda_j \pi_{\lambda_j} \right) (w) = \sum_{j=1}^m \lambda_j \pi_{\lambda_j} \left(\sum_{k=1}^m w_k \right) = \sum_{j=1}^m \sum_{k=1}^m \lambda_j \pi_{\lambda_j} w_k.$$

Since $w_k \in W_{\lambda_k}$, we have

$$\pi_{\lambda_j} w_k = \delta_{jk} w_k.$$

Thus, the sum becomes

$$\sum_{j=1}^m \sum_{k=1}^m \lambda_j \delta_{jk} w_k = \sum_{k=1}^m \lambda_k w_k.$$

This equals $A(w)$. Since the operators agree on all w , they are equal. The integral notation emphasizes that A is synthesized from its spectral values λ weighted by the spectral measure π_A concentrated at those values.

□

This spectral decomposition naturally suggests how to define $f(A)$ for functions f . If $p(z) = \sum_{k=0}^N c_k z^k$ is a polynomial, then using $A^k = (\sum \lambda_j \pi_{\lambda_j})^k = \sum \lambda_j^k \pi_{\lambda_j}$ (due to $\pi_j \pi_l = \delta_{jl} \pi_j$), we find

$$p(A) = \sum_{k=0}^N c_k A^k = \sum_{k=0}^N c_k \left(\sum_{j=1}^m \lambda_j^k \pi_{\lambda_j} \right) = \sum_{j=1}^m \left(\sum_{k=0}^N c_k \lambda_j^k \right) \pi_{\lambda_j} = \sum_{j=1}^m p(\lambda_j) \pi_{\lambda_j}.$$

This calculation shows that applying a polynomial p to the operator A is equivalent to applying p to each eigenvalue λ_j and summing the results weighted by the corresponding projection π_{λ_j} . This motivates extending the definition from polynomials to arbitrary functions defined on the spectrum of A .

Definition 4.21 (Functional Calculus). Let $A : W \rightarrow W$ be a diagonalizable operator with spectral measure π_A . Let $f : \text{Spec}(A) \rightarrow \mathbb{C}$ be any function defined on the spectrum of A . (We can extend f to all of \mathbb{C} arbitrarily, e.g., by setting $f(\lambda) = 0$ for $\lambda \notin \text{Spec}(A)$, as only its values on $\text{Spec}(A)$ affect the sum). The operator $f(A) : W \rightarrow W$ is defined as:

$$f(A) := \sum_{\lambda_j \in \text{Spec}(A)} f(\lambda_j) \pi_{\lambda_j}. \quad (4.8)$$

This can also be written using the integral notation:

$$f(A) = \int_{\text{Spec}(A)} f(\lambda) d\pi_A(\lambda).$$

This map $f \mapsto f(A)$ is called the **functional calculus** for the diagonalizable operator A .

This definition provides a consistent and computationally effective way to apply functions to operators.

Example 4.22 (Applying Functions to Functional Calculus).

- The operator exponential e^{tA} (important for solving linear ODE systems $\dot{x} = Ax$) is given by

$$e^{tA} = \sum_{j=1}^m e^{t\lambda_j} \pi_{\lambda_j}.$$

- If $0 \notin \text{Spec}(A)$, meaning A is invertible, we can define $f(\lambda) = 1/\lambda$ for $\lambda \in \text{Spec}(A)$. Then the inverse is

$$A^{-1} = f(A) = \sum_{j=1}^m \frac{1}{\lambda_j} \pi_{\lambda_j}.$$

We can verify this:

$$AA^{-1} = \left(\sum_{k=1}^m \lambda_k \pi_{\lambda_k} \right) \left(\sum_{j=1}^m \frac{1}{\lambda_j} \pi_{\lambda_j} \right) = \sum_{j,k=1}^m \frac{\lambda_k}{\lambda_j} \pi_{\lambda_k} \pi_{\lambda_j} = \sum_{j=1}^m \frac{\lambda_j}{\lambda_j} \pi_{\lambda_j} = \sum_{j=1}^m \pi_{\lambda_j} = \text{id}_W.$$

- Consider the matrix $A = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$. The spectrum of A is given by

$$\text{Spec}(A) = \{2, -1\}.$$

Let

$$W_2 = \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}, \quad W_{-1} = \text{span} \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}.$$

The corresponding projection operators are

$$\pi_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \pi_{-1} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Now, let $f(x) = x^3$. We can compute $f(A)$ as follows:

$$f(A) = f(2)\pi_2 + f(-1)\pi_{-1} = 8\pi_2 + (-1)\pi_{-1} = \begin{pmatrix} 8 & 0 \\ 0 & -1 \end{pmatrix},$$

which is equal to A^3 .

Next, let $g(x) = \sqrt{x}$ (defined for positive eigenvalues). If

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix},$$

then we compute $g(A)$ as follows:

$$g(A) = \sqrt{4}\pi_4 + \sqrt{9}\pi_9 = 2\pi_4 + 3\pi_9 = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix},$$

which is one possible square root of A .

The functional calculus possesses important algebraic properties, formalizing the idea that calculations with $f(A)$ mirror calculations with the function f itself on the spectrum. Let $\mathcal{F}(\text{Spec}(A))$ be the algebra of complex-valued functions on the finite set $\text{Spec}(A)$ equipped with pointwise addition, scalar multiplication, and pointwise multiplication. Let $\mathcal{L}(W)$ be the algebra of linear operators on W .

Theorem 4.23 (Properties of the Functional Calculus). *The functional calculus map $\Psi : \mathcal{F}(\text{Spec}(A)) \rightarrow \mathcal{L}(W)$ defined by $\Psi(f) = f(A)$ is an algebra homomorphism. That is, for $f, g \in \mathcal{F}(\text{Spec}(A))$ and $c \in \mathbb{C}$:*

1. $\Psi(f + g) = (f + g)(A) = f(A) + g(A) = \Psi(f) + \Psi(g)$ (Additivity).
2. $\Psi(cf) = (cf)(A) = cf(A) = c\Psi(f)$ (Homogeneity).
3. $\Psi(fg) = (fg)(A) = f(A)g(A) = \Psi(f)\Psi(g)$ (Multiplicativity).
4. $\Psi(\mathbf{1}) = \mathbf{1}(A) = \text{id}_W$, where $\mathbf{1}$ is the constant function $f(\lambda) = 1$.
5. $\Psi(\text{id}_{\mathbb{C}}) = \text{id}_{\mathbb{C}}(A) = A$, where $\text{id}_{\mathbb{C}}$ is the identity function $f(\lambda) = \lambda$.

If W is equipped with an inner product and A is a normal operator, then Ψ is furthermore a $*$ -homomorphism, meaning it respects the adjoint operation (involution):

6. $\Psi(\bar{f}) = \bar{f}(A) = (f(A))^* = (\Psi(f))^*$, where \bar{f} is the complex conjugate function $\bar{f}(\lambda) = \overline{f(\lambda)}$.

Proof. Properties (1) and (2) follow directly from the linearity of the summation in the definition (4.8):

$$\begin{aligned} (f+g)(A) &= \sum (f+g)(\lambda_j)\pi_j \\ &= \sum (f(\lambda_j) + g(\lambda_j))\pi_j \\ &= \sum f(\lambda_j)\pi_j + \sum g(\lambda_j)\pi_j \\ &= f(A) + g(A) \end{aligned}$$

and

$$\begin{aligned} (cf)(A) &= \sum (cf)(\lambda_j)\pi_j \\ &= \sum cf(\lambda_j)\pi_j \\ &= c \sum f(\lambda_j)\pi_j \\ &= cf(A). \end{aligned}$$

(3) Multiplicativity:

$$\begin{aligned} f(A)g(A) &= \left(\sum_{j=1}^m f(\lambda_j)\pi_{\lambda_j} \right) \left(\sum_{k=1}^m g(\lambda_k)\pi_{\lambda_k} \right) \\ &= \sum_{j,k=1}^m f(\lambda_j)g(\lambda_k)(\pi_{\lambda_j}\pi_{\lambda_k}) \\ &= \sum_{j,k=1}^m f(\lambda_j)g(\lambda_k)(\delta_{jk}\pi_{\lambda_j}) \quad (\text{using Lemma 4.12(2)}) \\ &= \sum_{j=1}^m f(\lambda_j)g(\lambda_j)\pi_{\lambda_j} \\ &= \sum_{j=1}^m (fg)(\lambda_j)\pi_{\lambda_j} = (fg)(A). \end{aligned}$$

(4) $\mathbf{1}(A) = \sum \mathbf{1}(\lambda_j)\pi_{\lambda_j} = \sum 1 \cdot \pi_{\lambda_j} = \text{id}_W$ by Lemma 4.12(1). (5) $\text{id}_\mathbb{C}(A) = \sum \text{id}_\mathbb{C}(\lambda_j)\pi_{\lambda_j} = \sum \lambda_j\pi_{\lambda_j} = A$ by Proposition 4.20. (6) Assume W is a Hilbert space and A is normal. Then each π_{λ_j} is an orthogonal projection, so $\pi_{\lambda_j}^* = \pi_{\lambda_j}$ by Proposition 4.11 and Remark 4.15. The adjoint of a sum is the sum of adjoints, and the adjoint of a scalar multiple is the conjugate scalar multiple of

the adjoint:

$$\begin{aligned}
(f(A))^* &= \left(\sum_{j=1}^m f(\lambda_j) \pi_{\lambda_j} \right)^* \\
&= \sum_{j=1}^m \overline{f(\lambda_j)} (\pi_{\lambda_j})^* \quad (\text{properties of Hilbert space adjoint}) \\
&= \sum_{j=1}^m \overline{f(\lambda_j)} \pi_{\lambda_j} \quad (\text{since } \pi_{\lambda_j}^* = \pi_{\lambda_j}) \\
&= \sum_{j=1}^m \bar{f}(\lambda_j) \pi_{\lambda_j} = \bar{f}(A). \quad \square
\end{aligned}$$

4.3 Generalizations and Special Cases in Finite Dimensions

So far, we've seen how the framework of projection-valued measures provides a powerful, basis-independent, and elegant way to understand diagonalizable operators on finite-dimensional vector spaces. By shifting focus from eigenvectors to projections onto eigenspaces, we obtain a spectral measure π_A supported on the spectrum $\text{Spec}(A)$. This measure satisfies properties analogous to classical measures but takes values in the space of projection operators. The resolution of identity, $\sum \pi_{\lambda_j} = \text{id}_W$, allows the reconstruction of the operator itself via its spectral decomposition, $A = \sum \lambda_j \pi_{\lambda_j}$.

This decomposition, in turn, forms the foundation for functional calculus, $f \mapsto f(A) = \sum f(\lambda_j) \pi_{\lambda_j}$. This calculus allows the application of arbitrary functions (defined on the spectrum) to the operator A in a manner consistent with polynomial functions and respecting the algebraic structure (as formalized by the homomorphism properties in Theorem 4.23). When dealing with normal operators on Hilbert spaces, the PVM consists of orthogonal projections, and the functional calculus further respects the adjoint operation, making it a $*$ -homomorphism.

While developed here for simplicity in finite dimensions, the true significance of the projection-valued measure and functional calculus perspective lies in its generalization to infinite-dimensional Hilbert spaces. The Spectral Theorem extends this framework to normal operators (bounded or unbounded). It asserts the existence of a PVM π_A defined on the Borel σ -algebra of \mathbb{C} (supported on the spectrum $\sigma(A)$, which may now be continuous) satisfying countable additivity, such that the operator and functions of it can be represented as integrals:

$$A = \int_{\sigma(A)} \lambda d\pi_A(\lambda) \quad \text{and} \quad f(A) = \int_{\sigma(A)} f(\lambda) d\pi_A(\lambda).$$

This infinite-dimensional spectral theory is very valuable for quantum mechanics, where observables are modeled by self-adjoint operators A . The spectral measure π_A then dictates the possible outcomes of measurements (the spectrum $\sigma(A)$) and their probabilities (via $\langle \psi, \pi_A(E) \psi \rangle$ for a state ψ and outcome range E). The finite-dimensional theory, therefore, serves as an essential conceptual launchpad for understanding these deeper and more widely applicable results.

So far, we've seen the theory for a single diagonalizable operator. Now, we explore how to deal with families of operators that can be simultaneously diagonalizable, which is important for analyzing systems with multiple commuting observables or symmetries. This section explores these important

specializations and generalizations within the finite-dimensional complex inner product space setting, laying essential groundwork for infinite-dimensional spectral theory and applications in quantum mechanics.

Throughout this section, unless otherwise specified, W denotes a finite-dimensional complex vector space equipped with an inner product $\langle \cdot, \cdot \rangle$, making it a finite-dimensional Hilbert space. We assume the inner product is linear in the second argument and conjugate-linear in the first.

4.3.1 Simultaneous Diagonalization of Commuting Operators

In many physical and mathematical contexts, one encounters systems described by multiple operators acting on the same space. A natural question is whether there exists a single basis in which all these operators take a simple (diagonal) form. The ability to simultaneously diagonalize a family of operators signifies the existence of a complete set of compatible observables or states distinguished by a common set of eigenvalues. Commutativity plays the central role.

Theorem 4.24 (Simultaneous Diagonalizability). *Let $\mathcal{F} = \{A_1, \dots, A_n\}$ be a finite set of linear operators $A_i : W \rightarrow W$ on a finite-dimensional complex vector space W .*

1. **(Necessity of Commutation)** *If the operators in \mathcal{F} are simultaneously diagonalizable (i.e., there exists a basis $\mathcal{B} = \{w_1, \dots, w_d\}$ for W such that each w_k is an eigenvector for every $A_i \in \mathcal{F}$), then the operators commute pairwise: $A_i A_j = A_j A_i$ for all $1 \leq i, j \leq n$.*
2. **(Sufficiency with Diagonalizability)** *Conversely, if the operators in \mathcal{F} commute pairwise ($A_i A_j = A_j A_i$ for all i, j) and each operator A_i is individually diagonalizable, then the operators in \mathcal{F} are simultaneously diagonalizable.*

Proof. (1) Assume there exists a basis $\mathcal{B} = \{w_k\}$ such that $A_i w_k = \lambda_{ik} w_k$ for all i, k . For any $w_k \in \mathcal{B}$, we have:

$$\begin{aligned} A_i A_j w_k &= A_i (\lambda_{jk} w_k) = \lambda_{jk} (A_i w_k) = \lambda_{jk} \lambda_{ik} w_k \\ A_j A_i w_k &= A_j (\lambda_{ik} w_k) = \lambda_{ik} (A_j w_k) = \lambda_{ik} \lambda_{jk} w_k \end{aligned}$$

Since complex numbers commute ($\lambda_{jk} \lambda_{ik} = \lambda_{ik} \lambda_{jk}$), we have $A_i A_j w_k = A_j A_i w_k$ for all basis vectors w_k . By linearity, this extends to all vectors in W , so $A_i A_j = A_j A_i$. (Alternatively, using matrices: If $A_i = S D_i S^{-1}$ where D_i are diagonal matrices representing the operators in the common eigenbasis \mathcal{B} (and S is the change-of-basis matrix), then $A_i A_j = (S D_i S^{-1})(S D_j S^{-1}) = S D_i D_j S^{-1}$. Since diagonal matrices always commute, $D_i D_j = D_j D_i$, we have $A_i A_j = S D_j D_i S^{-1} = (S D_j S^{-1})(S D_i S^{-1}) = A_j A_i$.)

(2) We prove this by induction on the dimension $d = \dim W$. The base case $d = 1$ is trivial. Assume the theorem holds for spaces of dimension less than d . Let A_1, \dots, A_n be commuting and individually diagonalizable operators on W . If all A_i are scalar multiples of the identity, $A_i = c_i \text{id}_W$, then any basis of W is a simultaneous eigenbasis, and we are done. Assume at least one operator, say A_n , is not a scalar multiple of the identity. Since A_n is diagonalizable, it has at least two distinct eigenvalues. Let $\text{Spec}(A_n) = \{\lambda_1, \dots, \lambda_m\}$ with $m \geq 2$. The space W decomposes into a direct sum of eigenspaces: $W = \bigoplus_{j=1}^m W_{\lambda_j}(A_n)$, where $W_{\lambda_j}(A_n) = \text{Ker}(A_n - \lambda_j \text{id}_W)$. Since A_n is not scalar, at least one eigenspace $W_{\lambda_j}(A_n)$ must be a proper subspace of W , meaning $\dim W_{\lambda_j}(A_n) < d$.

Now, we use the commutativity. For any A_i ($1 \leq i \leq n$) and any $v \in W_{\lambda_j}(A_n)$, we have $A_n v = \lambda_j v$. Then, using $A_i A_n = A_n A_i$:

$$A_n(A_i v) = (A_n A_i) v = (A_i A_n) v = A_i(A_n v) = A_i(\lambda_j v) = \lambda_j(A_i v)$$

This calculation shows that if $v \in W_{\lambda_j}(A_n)$, then $A_i v$ is also in $W_{\lambda_j}(A_n)$. Therefore, each eigenspace $W_{\lambda_j}(A_n)$ is invariant under all operators A_1, \dots, A_n .

Consider the restriction of the operators A_1, \dots, A_n to a specific eigenspace $W_{\lambda_j}(A_n)$. Let $A'_i = A_i|_{W_{\lambda_j}(A_n)} : W_{\lambda_j}(A_n) \rightarrow W_{\lambda_j}(A_n)$. These restricted operators A'_1, \dots, A'_n still commute with each other. Furthermore, since the original operators A_i were diagonalizable on W , their restrictions A'_i to an invariant subspace are also diagonalizable (this relies on the fact that the minimal polynomial of the restriction divides the minimal polynomial of the original operator, which has distinct roots). Note that $A'_n = A_n|_{W_{\lambda_j}(A_n)}$ is just $\lambda_j \text{id}_{W_{\lambda_j}(A_n)}$.

Since $\dim W_{\lambda_j}(A_n) < d$, we can apply the induction hypothesis to the commuting, diagonalizable operators A'_1, \dots, A'_{n-1} acting on $W_{\lambda_j}(A_n)$. This guarantees the existence of a basis \mathcal{B}_j for $W_{\lambda_j}(A_n)$ consisting of vectors that are simultaneous eigenvectors for A'_1, \dots, A'_{n-1} . Any vector $w \in \mathcal{B}_j$ is also an eigenvector for A'_n (with eigenvalue λ_j), and thus it is a simultaneous eigenvector for all A_1, \dots, A_n .

By constructing such a basis \mathcal{B}_j for each eigenspace $W_{\lambda_j}(A_n)$ and taking their union $\mathcal{B} = \bigcup_{j=1}^m \mathcal{B}_j$, we obtain a basis for $W = \bigoplus W_{\lambda_j}(A_n)$ consisting of simultaneous eigenvectors for all operators A_1, \dots, A_n . \square

Remark 4.25 (Importance of Diagonalizability). *Commutativity alone is insufficient. Consider*

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, AB = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}, BA = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix},$$

so A and B commute. However, neither A nor B is diagonalizable (their only eigenvalue is 1, but the eigenspace is only 1-dimensional, spanned by $(1, 0)^T$). Consequently, they cannot be simultaneously diagonalized. The condition that each operator be individually diagonalizable is essential for the second part of the theorem.

Corollary 4.26 (Simultaneous Unitary Diagonalization). *If W is a finite-dimensional complex inner product space and $\mathcal{F} = \{A_1, \dots, A_n\}$ is a family of commuting **normal** operators ($A_i A_i^* = A_i^* A_i$), then they are simultaneously unitarily diagonalizable. That is, there exists an orthonormal basis for W consisting of simultaneous eigenvectors for all $A_i \in \mathcal{F}$.*

Proof. Normal operators on a finite-dimensional complex inner product space are always diagonalizable (Spectral Theorem, see below). Since the A_i commute and are individually diagonalizable, Theorem 4.24(2) guarantees they are simultaneously diagonalizable. Furthermore, the proof shows that we can find an eigenbasis within each joint eigenspace. Since eigenspaces of normal operators corresponding to distinct eigenvalues are orthogonal, and within a degenerate eigenspace we can always choose an orthonormal basis, the resulting simultaneous eigenbasis for W can be chosen to be orthonormal. \square

The PVM formalism generalizes naturally to infinite dimensions, providing the core object for the spectral theorem.

Definition 4.27 (Spectral Measure / PVM). Let (X, Σ) be a measurable space (where Σ is a σ -algebra of subsets of X) and let \mathcal{H} be a complex separable Hilbert space. Let $\text{Proj}^\perp(\mathcal{H})$ denote the set of orthogonal projection operators on \mathcal{H} . A map

$$\pi : \Sigma \rightarrow \text{Proj}^\perp(\mathcal{H})$$

is called a **spectral measure** (or **projection-valued measure**, PVM) if it satisfies:

1. **Normalization:** $\pi(\emptyset) = 0$ (the zero operator) and $\pi(X) = \text{id}_{\mathcal{H}}$ (the identity operator).
2. **Countable Additivity (Strong):** If $\{E_i\}_{i=1}^{\infty}$ is a sequence of pairwise disjoint sets in Σ ($E_i \cap E_j = \emptyset$ for $i \neq j$), then for any vector $\psi \in \mathcal{H}$, the series $\sum_{i=1}^{\infty} \pi(E_i)\psi$ converges in the norm topology of \mathcal{H} to $\pi(\cup_{i=1}^{\infty} E_i)\psi$:

$$\pi\left(\bigcup_{i=1}^{\infty} E_i\right)\psi = \sum_{i=1}^{\infty} \pi(E_i)\psi \quad (\text{convergence in } \mathcal{H}).$$

This is equivalent to requiring convergence in the strong operator topology.

3. **Multiplicativity:** For any $E_1, E_2 \in \Sigma$,

$$\pi(E_1 \cap E_2) = \pi(E_1)\pi(E_2).$$

Remark 4.28 (Geometric Picture and PVM on V^*). *Consider a family of commuting, diagonalizable operators $\{A_i\}_{i \in I}$ (where I is some index set). We can view this family as arising from a representation of an abelian structure. For instance, let $V = \mathbb{C}^n$ and consider a linear map $A : V \rightarrow \text{End}(W)$ defined by $A(c_1, \dots, c_n) = \sum_{i=1}^n c_i A_i$, where $\{A_1, \dots, A_n\}$ is the commuting, diagonalizable family. Since the A_i commute and are diagonalizable, the space W decomposes into a direct sum of simultaneous eigenspaces, $W = \bigoplus_{j=1}^l W^{(j)}$. On each subspace $W^{(j)}$, every operator A_i acts as scalar multiplication: $A_i|_{W^{(j)}} = \lambda_{ij} \text{id}_{W^{(j)}}$. Consequently, any linear combination $A(\xi)$ for $\xi = (c_1, \dots, c_n) \in V$ also acts as a scalar on $W^{(j)}$:*

$$A(\xi)|_{W^{(j)}} = \left(\sum_{i=1}^n c_i A_i \right) |_{W^{(j)}} = \sum_{i=1}^n c_i (A_i|_{W^{(j)}}) = \sum_{i=1}^n c_i \lambda_{ij} \text{id}_{W^{(j)}} = \theta_j(\xi) \text{id}_{W^{(j)}}$$

where $\theta_j : V \rightarrow \mathbb{C}$ is the linear functional defined by $\theta_j(c_1, \dots, c_n) = \sum_{i=1}^n c_i \lambda_{ij}$. This functional θ_j belongs to the dual space $V^* = \text{Hom}(V, \mathbb{C})$. The tuple of eigenvalues $(\lambda_{1j}, \dots, \lambda_{nj})$ thus defines a point θ_j in the dual space V^* .

This structure admits an interpretation as a projection-valued measure on the dual space V^* . Let $\pi^{(j)}$ be the projection onto the joint eigenspace $W^{(j)}$ along the sum of the other joint eigenspaces. Then $\{\pi^{(j)}\}_{j=1}^l$ forms a resolution of the identity: $\sum_{j=1}^l \pi^{(j)} = \text{id}_W$ and $\pi^{(j)}\pi^{(k)} = \delta_{jk}\pi^{(j)}$. The action of any operator $A(\xi)$ can be written using this PVM:

$$A(\xi) = \sum_{j=1}^l \theta_j(\xi) \pi^{(j)} = \int_{V^*} \theta(\xi) d\pi_A(\theta)$$

where the PVM π_A is supported on the finite set $\{\theta_1, \dots, \theta_l\} \subset V^*$ and $d\pi_A(\theta)$ assigns the projection $\pi^{(j)}$ to the point θ_j . This perspective elegantly encodes the simultaneous spectral decomposition of the entire commuting family via a single PVM on the dual space parameterizing the eigenvalues. It mirrors the skyscraper sheaf interpretation where the operator is determined by its scalar actions on stalks (eigenspaces) located at specific points (eigenvalue tuples/functionals) in a base space (V^*).

4.3.2 Self-Adjoint Operators and the Spectral Theorem

We now specialize to operators that respect the inner product structure in a specific way: self-adjoint operators, which play a important role as observables in quantum mechanics. Assume W is a finite-dimensional complex inner product space.

Definition 4.29 (Adjoint and Self-Adjoint Operator). Given a linear operator $A : W \rightarrow W$, its **adjoint** $A^* : W \rightarrow W$ is the unique linear operator satisfying

$$\langle Av, w \rangle = \langle v, A^*w \rangle \quad \text{for all } v, w \in W.$$

An operator A is called **self-adjoint** (or **Hermitian**) if $A = A^*$.

Remark 4.30 (Properties of Adjoint).

- *Existence and uniqueness of A^* follow from the Riesz representation theorem in finite dimensions.*
- *In an orthonormal basis $\{e_k\}$, if the matrix of A is $[A]_{ij} = \langle e_i, Ae_j \rangle$, then the matrix of A^* is $[A^*]_{ij} = \langle e_i, A^*e_j \rangle = \langle Ae_i, e_j \rangle = \overline{\langle e_j, Ae_i \rangle} = \overline{[A]_{ji}}$. Thus, the matrix of A^* is the conjugate transpose (Hermitian conjugate) of the matrix of A .*
- *Adjoint properties: $(A + B)^* = A^* + B^*$, $(cA)^* = \bar{c}A^*$, $(AB)^* = B^*A^*$, $(A^*)^* = A$.*
- *An operator A is self-adjoint if and only if its matrix representation in any orthonormal basis is Hermitian ($M = M^\dagger$).*

Self-adjoint operators possess remarkable spectral properties that simplify their analysis significantly.

Theorem 4.31 (Spectral Properties of Self-Adjoint Operators). *Let $A : W \rightarrow W$ be a self-adjoint operator on a finite-dimensional complex inner product space W . Then:*

1. **Normality and Diagonalizability:** *A is normal ($AA^* = A^2 = A^*A$) and therefore unitarily diagonalizable. That is, there exists an orthonormal basis for W consisting of eigenvectors of A .*
2. **Real Eigenvalues:** *All eigenvalues of A are real numbers.*
3. **Orthogonal Eigenspaces:** *Eigenspaces corresponding to distinct eigenvalues of A are mutually orthogonal.*

Proof.

1. Normality $A = A^* \implies AA^* = A^2$ and $A^*A = A^2$, so $AA^* = A^*A$. The theorem stating that normal operators on finite-dimensional complex Hilbert spaces are unitarily diagonalizable is important (often proved by induction on dimension, showing existence of at least one eigenvector, considering the orthogonal complement which is invariant under A and A^* , and applying induction).
2. Let $\lambda \in \mathbb{C}$ be an eigenvalue of A with corresponding non-zero eigenvector v : $Av = \lambda v$. We compute $\langle Av, v \rangle$ in two ways:

$$\langle Av, v \rangle = \langle \lambda v, v \rangle = \lambda \langle v, v \rangle = \lambda \|v\|^2$$

Using self-adjointness ($A = A^*$):

$$\langle Av, v \rangle = \langle v, A^*v \rangle = \langle v, Av \rangle = \langle v, \lambda v \rangle = \bar{\lambda} \langle v, v \rangle = \bar{\lambda} \|v\|^2$$

Equating the two expressions gives $\lambda \|v\|^2 = \bar{\lambda} \|v\|^2$. Since $v \neq 0$, $\|v\|^2 \neq 0$, so we must have $\lambda = \bar{\lambda}$, which means λ is real. Thus, $\text{Spec}(A) \subset \mathbb{R}$.

3. Let λ, μ be distinct eigenvalues of A ($\lambda \neq \mu$, $\lambda, \mu \in \mathbb{R}$) with corresponding eigenvectors v, w : $Av = \lambda v$ and $Aw = \mu w$. Consider $\langle Av, w \rangle$:

$$\langle Av, w \rangle = \langle \lambda v, w \rangle = \lambda \langle v, w \rangle \quad (\text{since } \lambda \in \mathbb{R})$$

Using self-adjointness again:

$$\langle Av, w \rangle = \langle v, A^* w \rangle = \langle v, Aw \rangle = \langle v, \mu w \rangle = \mu \langle v, w \rangle \quad (\text{since } \mu \in \mathbb{R})$$

Equating the results gives $\lambda \langle v, w \rangle = \mu \langle v, w \rangle$, or $(\lambda - \mu) \langle v, w \rangle = 0$. Since $\lambda \neq \mu$, we must conclude that $\langle v, w \rangle = 0$. Thus, $v \perp w$. This extends to show that the entire eigenspaces W_λ and W_μ are orthogonal.

□

These properties lead to the finite-dimensional version of the Spectral Theorem for self-adjoint operators, which can be elegantly stated using the PVM formalism.

Theorem 4.32 (Spectral Theorem for Self-Adjoint Operators (Finite Dimensions)). *Let $A : W \rightarrow W$ be a self-adjoint operator on a finite-dimensional complex inner product space W . Then there exists a unique projection-valued measure π_A , supported on the spectrum $\text{Spec}(A) \subset \mathbb{R}$, whose values $\pi_A(E)$ are orthogonal projections, such that*

$$A = \int_{\mathbb{R}} \lambda d\pi_A(\lambda) = \sum_{\lambda_j \in \text{Spec}(A)} \lambda_j \pi_{\lambda_j}$$

where $\pi_{\lambda_j} = \pi_A(\{\lambda_j\})$ is the orthogonal projection onto the eigenspace W_{λ_j} . Furthermore, W decomposes as the orthogonal direct sum $W = \bigoplus_{\lambda_j \in \text{Spec}(A)} W_{\lambda_j}$.

Proof. Theorem 4.31 establishes that A is diagonalizable, its eigenvalues λ_j are real, and its eigenspaces W_{λ_j} are mutually orthogonal. The direct sum $W = \bigoplus W_{\lambda_j}$ is therefore an orthogonal direct sum. As discussed in Remark 4.15, when the eigenspace decomposition is orthogonal, the corresponding projections π_{λ_j} onto W_{λ_j} (along the orthogonal complement $\bigoplus_{k \neq j} W_{\lambda_k}$) are orthogonal projections ($P^2 = P, P^* = P$). The PVM $\pi_A(E) = \sum_{\lambda_j \in E \cap \text{Spec}(A)} \pi_{\lambda_j}$ thus consists of orthogonal projections. Its support is clearly $\text{Spec}(A)$, which is a subset of \mathbb{R} . The spectral decomposition $A = \sum \lambda_j \pi_{\lambda_j}$ holds as established in Proposition 4.20. Uniqueness of the PVM follows from the uniqueness of the spectral decomposition. □

4.3.3 Unitary Operators and the Spectral Theorem

Another crucial class of operators preserving the inner product structure are unitary operators, representing geometric symmetries like rotations and reflections in Hilbert space.

Definition 4.33 (Unitary Operator). A linear operator $U : W \rightarrow W$ on a finite-dimensional complex inner product space W is **unitary** if it preserves the inner product:

$$\langle Uv, Uw \rangle = \langle v, w \rangle \quad \text{for all } v, w \in W.$$

Proposition 4.34 (Equivalent Characterizations of Unitary Operators). *For a linear operator $U : W \rightarrow W$ on a finite-dimensional complex inner product space, the following are equivalent:*

1. U is unitary ($\langle Uv, Uw \rangle = \langle v, w \rangle$).
2. U preserves norms: $\|Uv\| = \|v\|$ for all $v \in W$.
3. $U^*U = \text{id}_W$.
4. U is invertible and $U^{-1} = U^*$.
5. $UU^* = \text{id}_W$.
6. The columns (or rows) of the matrix representation of U in any orthonormal basis form an orthonormal basis for \mathbb{C}^d .

Sketch.

- (1 \implies 2): Set $v = w$. $\|Uv\|^2 = \langle Uv, Uv \rangle = \langle v, v \rangle = \|v\|^2$.
- (2 \implies 1): Use the polarization identity to recover the inner product from the norm.
- (1 \iff 3): $\langle Uv, Uw \rangle = \langle v, U^*Uw \rangle$. This equals $\langle v, w \rangle$ for all v, w if and only if $U^*U = \text{id}_W$.
- (3 \implies 4): Since W is finite-dimensional, a left inverse is also a right inverse, so $U^*U = I$ implies U is invertible and $U^{-1} = U^*$.
- (4 \implies 5): $U^{-1} = U^* \implies UU^* = UU^{-1} = I$.
- (5 \implies 3): $UU^* = I \implies U$ is invertible with $U^{-1} = U^*$, so $U^*U = U^{-1}U = I$. Equivalence with (6) involves writing out the inner product or matrix conditions in an orthonormal basis.

□

Unitary operators share the property of normality with self-adjoint operators, leading to similar spectral conclusions.

Theorem 4.35 (Spectral Properties of Unitary Operators). *Let $U : W \rightarrow W$ be a unitary operator on a finite-dimensional complex inner product space W . Then:*

1. **Normality and Diagonalizability:** U is normal ($UU^* = I = U^*U$) and therefore unitarily diagonalizable. There exists an orthonormal basis for W consisting of eigenvectors of U . (Follows from the general theorem that normal operators are unitarily diagonalizable).
2. **Eigenvalues on Unit Circle:** All eigenvalues λ of U have modulus 1, i.e., $|\lambda| = 1$.
3. **Orthogonal Eigenspaces:** Eigenspaces corresponding to distinct eigenvalues of U are mutually orthogonal.

Proof.

- (1) Normality $U^*U = UU^* = I$ is part of the definition/equivalent characterizations. Diagonalizability follows from normality.
- (2) Let $Uv = \lambda v$ for a non-zero eigenvector v . Since U preserves norms:

$$\|v\|^2 = \|Uv\|^2 = \|\lambda v\|^2 = |\lambda|^2 \|v\|^2$$

As $\|v\|^2 \neq 0$, we must have $|\lambda|^2 = 1$. Thus, eigenvalues lie on the unit circle $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$.

(3) Orthogonality of eigenspaces holds for all normal operators. Let $Uv = \lambda v$ and $Uw = \mu w$ with $\lambda \neq \mu$ ($|\lambda| = |\mu| = 1$).

$$\langle v, w \rangle = \langle Uv, Uw \rangle = \langle \lambda v, \mu w \rangle = \bar{\lambda}\mu \langle v, w \rangle$$

So $(1 - \bar{\lambda}\mu)\langle v, w \rangle = 0$. Since $|\lambda| = 1$, $\bar{\lambda} = 1/\lambda$. The condition becomes $(1 - \mu/\lambda)\langle v, w \rangle = 0$. Since $\lambda \neq \mu$, $\mu/\lambda \neq 1$, so the first factor is non-zero. Therefore, $\langle v, w \rangle = 0$. \square

Again, these properties lead to a spectral theorem for unitary operators.

Theorem 4.36 (Spectral Theorem for Unitary Operators (Finite Dimensions)). *Let $U : W \rightarrow W$ be a unitary operator on a finite-dimensional complex inner product space W . Then there exists a unique projection-valued measure π_U , supported on the spectrum $\text{Spec}(U) \subset \mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$, whose values $\pi_U(E)$ are orthogonal projections, such that*

$$U = \int_{\mathbb{T}} \lambda d\pi_U(\lambda) = \sum_{\lambda_j \in \text{Spec}(U)} \lambda_j \pi_{\lambda_j}$$

where $\pi_{\lambda_j} = \pi_U(\{\lambda_j\})$ is the orthogonal projection onto the eigenspace W_{λ_j} . Furthermore, W decomposes as the orthogonal direct sum $W = \bigoplus_{\lambda_j \in \text{Spec}(U)} W_{\lambda_j}$.

Proof. Similar to the self-adjoint case: Theorem 4.35 establishes unitary diagonalizability with orthogonal eigenspaces and eigenvalues on \mathbb{T} . The PVM $\pi_U(E) = \sum_{\lambda_j \in E \cap \text{Spec}(U)} \pi_{\lambda_j}$ consists of orthogonal projections and is supported on $\text{Spec}(U) \subset \mathbb{T}$. The spectral decomposition $U = \sum \lambda_j \pi_{\lambda_j}$ holds by Proposition 4.20. \square

4.3.4 Connection via Exponentiation and Unitary Representations

The functional calculus, particularly for normal operators, provides a powerful link between self-adjoint and unitary operators via the complex exponential function. This connection is important for understanding continuous symmetries and time evolution in quantum mechanics.

Let $A : W \rightarrow W$ be a self-adjoint operator. Its spectrum $\text{Spec}(A)$ lies on the real axis \mathbb{R} . Consider the function $f_t(\lambda) = e^{it\lambda}$, where $t \in \mathbb{R}$ is a fixed real parameter. This function maps the real line \mathbb{R} to the unit circle $\mathbb{T} \subset \mathbb{C}$. Using the functional calculus for the normal (specifically, self-adjoint) operator A , we can define the operator $U_t = f_t(A)$:

$$U_t := e^{itA} = f_t(A) = \sum_{\lambda_j \in \text{Spec}(A)} f_t(\lambda_j) \pi_{\lambda_j} = \sum_{\lambda_j \in \text{Spec}(A) \subset \mathbb{R}} e^{it\lambda_j} \pi_{\lambda_j}$$

where $\pi_A = \{\pi_{\lambda_j}\}$ is the orthogonal PVM for A .

Proposition 4.37. *For any self-adjoint operator A on a finite-dimensional complex inner product space W and any $t \in \mathbb{R}$, the operator $U_t = e^{itA}$ is unitary. The family $\{U_t\}_{t \in \mathbb{R}}$ forms a one-parameter group of unitary operators: $U_0 = \text{id}_W$ and $U_{s+t} = U_s U_t$.*

Proof. We use the $*$ -homomorphism property of the functional calculus for normal operators (Theorem 4.23(6)). The function defining U_t is $f_t(\lambda) = e^{it\lambda}$. Its complex conjugate function is $\bar{f}_t(\lambda) = f_{-t}(\lambda) = e^{-it\lambda} = \overline{e^{it\lambda}}$ (since $\lambda, t \in \mathbb{R}$). The adjoint of U_t is:

$$(U_t)^* = (f_t(A))^* = \bar{f}_t(A)$$

Since $\bar{f}_t(\lambda) = e^{-it\lambda}$, we have $\bar{f}_t(A) = e^{-itA} = U_{-t}$. Thus, $(e^{itA})^* = e^{-itA}$. Now we check the unitary condition:

$$(U_t)^* U_t = (e^{itA})^* e^{itA} = e^{-itA} e^{itA}$$

Using the multiplicative property of the functional calculus (Theorem 4.23(3)) for the functions $g(\lambda) = e^{-it\lambda}$ and $h(\lambda) = e^{it\lambda}$:

$$e^{-itA} e^{itA} = g(A)h(A) = (gh)(A)$$

where $(gh)(\lambda) = e^{-it\lambda} e^{it\lambda} = e^0 = 1$. So (gh) is the constant function $\mathbf{1}$.

$$(U_t)^* U_t = \mathbf{1}(A) = \text{id}_W$$

Hence, U_t is unitary.

The group property $U_{s+t} = U_s U_t$ follows from $e^{i(s+t)\lambda} = e^{is\lambda} e^{it\lambda}$ and the multiplicative property of the functional calculus. $U_0 = e^{i0A} = e^0(A) = \mathbf{1}(A) = \text{id}_W$. \square

This demonstrates that the exponentiation map $A \mapsto e^{itA}$ takes self-adjoint operators (generators) to one-parameter unitary groups (transformations). The map $x \mapsto e^{ix}$ relates the additive group \mathbb{R} (home of spectra for self-adjoint operators) to the multiplicative group \mathbb{T} (home of spectra for unitary operators):

$$\exp(-i\cdot) : \mathbb{R} \longrightarrow \mathbb{T} \quad (4.9)$$

This relationship finds its deepest expression in Stone's theorem on one-parameter unitary groups in infinite-dimensional Hilbert spaces, which establishes a one-to-one correspondence between strongly continuous one-parameter unitary groups U_t and (possibly unbounded) self-adjoint operators A via $U_t = e^{itA}$ (or e^{-itA}). This theorem is important for describing time evolution in quantum mechanics, where the Hamiltonian H (a self-adjoint operator) generates the unitary time evolution group $U_t = e^{-itH/\hbar}$.

The ideas of simultaneous diagonalization and exponentiation combine naturally in the context of unitary representations of abelian groups.

Example 4.38 (Unitary Representations of \mathbb{R}^n). Let V be a finite-dimensional real vector space, considered as an additive abelian Lie group. A **unitary representation** of V on a finite-dimensional complex inner product space W is a continuous group homomorphism

$$A : V \longrightarrow U(W) \quad (4.10)$$

where $U(W)$ is the Lie group of unitary operators on W . Since V is abelian, the image $A(V) \subset U(W)$ is a commuting family of unitary operators. By Corollary 4.26, the operators $\{A(\xi) \mid \xi \in V\}$ are simultaneously unitarily diagonalizable. Let $W = \bigoplus_{\theta} W_{\theta}$ be the decomposition into simultaneous eigenspaces. For a vector $w \in W_{\theta}$, $A(\xi)w = \chi_{\theta}(\xi)w$ for some scalar $\chi_{\theta}(\xi)$ which depends on ξ and the joint eigenspace θ . Since $A(\xi)$ is unitary, $|\chi_{\theta}(\xi)| = 1$. Since A is a homomorphism ($A(\xi + \eta) = A(\xi)A(\eta)$), we must have $\chi_{\theta}(\xi + \eta) = \chi_{\theta}(\xi)\chi_{\theta}(\eta)$. This means $\chi_{\theta} : V \rightarrow \mathbb{T}$ is a continuous character of the additive group $V = \mathbb{R}^n$. Such characters are precisely of the form $\chi_{\theta}(\xi) = e^{i\langle \theta, \xi \rangle}$ for some unique linear functional $\theta \in V^* = \text{Hom}(V, \mathbb{R})$. Therefore, the simultaneous diagonalization involves a decomposition indexed by elements θ in the dual space V^* . The spectral decomposition for each operator $A(\xi)$ can be written using a projection-valued measure π_A supported on a finite subset of V^* :

$$A(\xi) = \int_{V^*} e^{i\langle \theta, \xi \rangle} d\pi_A(\theta) = \sum_{\theta_j \in \text{Supp}(\pi_A)} e^{i\langle \theta_j, \xi \rangle} \pi_{\theta_j} \quad (4.11)$$

where π_{θ_j} is the orthogonal projection onto the simultaneous eigenspace W_{θ_j} on which every $A(\xi)$ acts as multiplication by the character $e^{i\langle \theta_j, \xi \rangle}$.

Remark 4.39 (Generalization to Abelian Groups). *Example 4.38 generalizes to unitary representations $A : G \rightarrow U(W)$ of any locally compact abelian (LCA) group G . The spectral decomposition occurs over the Pontryagin dual group \hat{G} , which is the group of continuous characters $\chi : G \rightarrow \mathbb{T}$. The spectral theorem for such representations states that W decomposes into an orthogonal direct sum of spaces W_χ on which $A(g)$ acts as $\chi(g) \text{id}_{W_\chi}$, leading to a spectral decomposition $A(g) = \int_{\hat{G}} \chi(g) d\pi(\chi)$ for a PVM π on \hat{G} . This is important in harmonic analysis and representation theory.*

4.3.5 Generalization: Unitary Representations of Locally Compact Abelian Groups and Pontryagin Duality

The spectral decomposition of unitary representations of \mathbb{R}^n leads to a very general theory encompassing all locally compact abelian (LCA) groups. This generalization is very powerful since LCA groups provide the natural framework for symmetries encountered in many areas, including Euclidean translations and time evolution (\mathbb{R}^n), lattice symmetries (\mathbb{Z}^n), rotational symmetry (S^1), discrete time systems (\mathbb{Z}), and more exotic structures arising in number theory (\mathbb{Q}_p , adeles). The spectral theorem for LCA groups, closely related to Pontryagin duality, forms the foundations of abstract harmonic analysis, revealing that any unitary representation of such a group decomposes canonically according to its important frequencies, or characters.

We first establish the algebraic and topological setting.

Definition 4.40 (LCA Group). A topological group (G, \cdot, τ) is **locally compact abelian (LCA)** if the underlying topological space (G, τ) is Hausdorff and locally compact, and the group operation is commutative.

The combination of algebraic (abelian) and topological (locally compact Hausdorff) properties is crucial. Local compactness ensures the existence of a Haar measure (a translation-invariant measure), vital for integration and L^p spaces. Hausdorff is a standard separation axiom. Commutativity simplifies the representation theory immensely compared to the non-abelian case.

The concept dual to a group element is its “frequency component,” formalized as a character.

Definition 4.41 (Continuous Character). Let G be an LCA group. A **continuous character** of G is a continuous group homomorphism $\chi : G \rightarrow \mathbb{T}$, where $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ is the circle group under complex multiplication. That is, χ is continuous and satisfies $\chi(g_1 g_2) = \chi(g_1) \chi(g_2)$ for all $g_1, g_2 \in G$.

The target group \mathbb{T} is chosen because we are interested in unitary representations, whose eigenvalues (in the abelian case) must lie on the unit circle.

The set of all continuous characters of G forms a group itself, the Pontryagin dual.

Definition 4.42 (Pontryagin Dual Group). Let G be an LCA group. The **Pontryagin dual** of G , denoted \hat{G} , is the set of all continuous characters of G . The group operation on \hat{G} is pointwise multiplication: for $\chi_1, \chi_2 \in \hat{G}$, their product $\chi_1 \chi_2$ is defined by $(\chi_1 \chi_2)(g) = \chi_1(g) \chi_2(g)$ for all

$g \in G$. The identity element is the trivial character $\chi_0(g) = 1$ for all g , and the inverse of χ is $\chi^{-1}(g) = \overline{\chi(g)} = (\chi(g))^{-1}$. The topology on \hat{G} is the **compact-open topology**, which coincides with the topology of uniform convergence on compact subsets of G .

A important result of the theory, attributed to Pontryagin and van Kampen, is that the dual object retains the essential properties of the original.

Theorem 4.43 (Duality Theorem for LCA Groups). *If G is an LCA group, then its Pontryagin dual \hat{G} , equipped with pointwise multiplication and the compact-open topology, is also an LCA group.*

This allows for an iteration of the duality construction. The Pontryagin Duality Theorem asserts a powerful symmetry.

Theorem 4.44 (Pontryagin Duality Theorem). *Let G be an LCA group and let $\hat{\hat{G}}$ be the dual of its dual group \hat{G} . There is a canonical topological group isomorphism $\alpha : G \rightarrow \hat{\hat{G}}$ given by evaluation:*

$$\alpha(g)(\chi) = \chi(g) \quad \text{for } g \in G, \chi \in \hat{\hat{G}}.$$

Thus, G is canonically isomorphic to the dual of its dual.

This theorem establishes a perfect duality. It also implies relationships between topological properties:

Theorem 4.45. *G is compact if and only if \hat{G} is discrete; G is discrete if and only if \hat{G} is compact.*

Proof. There are four directions to prove:

G is compact $\implies \hat{G}$ is discrete: To show \hat{G} is discrete, we show $\{\chi_0\}$ is an open set. Let $K = G$. Since G is compact, K is compact. Let $V = \{z \in \mathbb{T} : \operatorname{Re}(z) > 0\}$. V is an open neighborhood of 1 in \mathbb{T} and contains no non-trivial subgroup of \mathbb{T} . Consider the basic open set $U(G, V) = \{\chi \in \hat{G} : \chi(G) \subseteq V\}$. If $\chi \in U(G, V)$, then $\chi(G)$ is a subgroup of \mathbb{T} contained in V . Thus, $\chi(G) = \{1\}$, which implies $\chi = \chi_0$. So, $U(G, V) = \{\chi_0\}$, meaning $\{\chi_0\}$ is open in \hat{G} . Hence, \hat{G} is discrete.

\hat{G} is discrete $\implies G$ is compact: If \hat{G} is discrete, then by H discrete $\implies \hat{H}$ compact, applied to the group $H = \hat{G}$, we have that $\hat{\hat{G}}$ is compact. By Pontryagin duality, G is topologically isomorphic to $\hat{\hat{G}}$. Therefore, G is compact.

G is discrete $\implies \hat{G}$ is compact: If G is discrete, any homomorphism $\chi : G \rightarrow \mathbb{T}$ is continuous. The compact-open topology on \hat{G} coincides with the topology of pointwise convergence. \hat{G} can be viewed as the set of all group homomorphisms from G to \mathbb{T} . Consider the product space $\mathbb{T}^G = \prod_{g \in G} \mathbb{T}_g$, where each $\mathbb{T}_g = \mathbb{T}$. By Tychonoff's theorem, \mathbb{T}^G is compact. The map $\Psi : \hat{G} \rightarrow \mathbb{T}^G$ given by $\chi \mapsto (\chi(g))_{g \in G}$ embeds \hat{G} into \mathbb{T}^G . The image $\Psi(\hat{G})$ is the set $\{(\phi_g)_{g \in G} \in \mathbb{T}^G : \phi_{gh} = \phi_g \phi_h \text{ for all } g, h \in G\}$. Let (χ_α) be a net in \hat{G} converging pointwise to $\phi = (\phi_g) \in \mathbb{T}^G$. So, $\chi_\alpha(g) \rightarrow \phi_g$ for all $g \in G$. Then $\phi_{gh} = \lim_\alpha \chi_\alpha(gh) = \lim_\alpha (\chi_\alpha(g) \chi_\alpha(h)) = (\lim_\alpha \chi_\alpha(g)) (\lim_\alpha \chi_\alpha(h)) = \phi_g \phi_h$. Thus, ϕ represents a homomorphism, so $\phi \in \Psi(\hat{G})$. This shows $\Psi(\hat{G})$ is a closed subset of the compact space \mathbb{T}^G . Hence, \hat{G} is compact.

\hat{G} is compact $\implies G$ is discrete: If \hat{G} is compact, then by H compact $\implies \hat{H}$ discrete, applied to the group $H = \hat{G}$, we have that $\hat{\hat{G}}$ is discrete. By Pontryagin duality, G is topologically isomorphic to $\hat{\hat{G}}$. Therefore, G is discrete. \square

Example 4.46 (Examples of Pontryagin Duality).

- $G = \mathbb{R}^n$. Characters $\chi_\theta(\xi) = e^{i\langle \theta, \xi \rangle}$ for $\theta \in \mathbb{R}^n$. The map $\theta \mapsto \chi_\theta$ establishes the topological group isomorphism $\mathbb{R}^n \cong \mathbb{R}^n$. \mathbb{R}^n is self-dual.
- $G = \mathbb{Z}^n$. Characters $\chi_z(k) = z^k = z_1^{k_1} \cdots z_n^{k_n}$ for $z = (z_1, \dots, z_n) \in \mathbb{T}^n$. Thus $\hat{\mathbb{Z}}^n \cong \mathbb{T}^n$. (Discrete group maps to compact dual).
- $G = \mathbb{T}^n$. Characters $\chi_k(z) = z^k = z_1^{k_1} \cdots z_n^{k_n}$ for $k = (k_1, \dots, k_n) \in \mathbb{Z}^n$. Thus $\hat{\mathbb{T}}^n \cong \mathbb{Z}^n$. (Compact group maps to discrete dual).
- If G is a finite abelian group, \hat{G} is also a finite abelian group (isomorphic to G , though not canonically in general).

With the dual group established, we can now state the central theorem regarding the structure of unitary representations of LCA groups.

Definition 4.47 (Continuous Unitary Representation). Let G be an LCA group and \mathcal{H} be a complex Hilbert space. A homomorphism $A : G \rightarrow U(\mathcal{H})$ is a **continuous unitary representation** if the map $G \times \mathcal{H} \rightarrow \mathcal{H}$ given by $(g, \psi) \mapsto A(g)\psi$ is continuous. For separable \mathcal{H} , this is equivalent to requiring strong operator topology continuity: for each $\psi \in \mathcal{H}$, the map $g \mapsto A(g)\psi$ is continuous from G to \mathcal{H} .

Since G is abelian, the image $A(G)$ is a commuting family of unitary operators in $U(\mathcal{H})$. This commutativity is the key to a simultaneous spectral decomposition over the dual group \hat{G} .

Theorem 4.48 (Spectral Theorem for Unitary Representations of LCA Groups). *Let $A : G \rightarrow U(\mathcal{H})$ be a continuous unitary representation of an LCA group G on a complex separable Hilbert space \mathcal{H} . Then there exists a unique **regular** projection-valued measure (PVM) π defined on the Borel σ -algebra $\mathcal{B}(\hat{G})$ of the Pontryagin dual group \hat{G} , taking values in the lattice of orthogonal projections $\text{Proj}^\perp(\mathcal{H})$ on \mathcal{H} , such that for every group element $g \in G$:*

$$A(g) = \int_{\hat{G}} \chi(g) d\pi(\chi)$$

The integral signifies that for all $\phi, \psi \in \mathcal{H}$:

$$\langle \phi, A(g)\psi \rangle = \int_{\hat{G}} \chi(g) d\mu_{\phi, \psi}(\chi)$$

where $\mu_{\phi, \psi}(E) = \langle \phi, \pi(E)\psi \rangle$ is the unique regular complex Borel measure on \hat{G} determined by the PVM and the vectors ϕ, ψ . The PVM π is supported on the spectrum of the representation.

Remark 4.49.

- **Diagonalization over Characters:** This theorem provides simultaneous diagonalization for the commuting family $\{A(g)\}_{g \in G}$. The “eigenvalues” are the characters $\chi \in \hat{G}$, and the “eigenvalue” of a specific operator $A(g)$ corresponding to the character χ is simply the scalar $\chi(g) \in \mathbb{T}$. The PVM π plays the role of projecting onto the (possibly infinitesimal or continuously distributed) “eigenspaces” corresponding to sets of characters.
- **Integral Definition:** The operator integral $A(g) = \int \chi(g) d\pi(\chi)$ is defined rigorously using the scalar measures $\mu_{\phi, \psi}$. For any bounded Borel function $f : \hat{G} \rightarrow \mathbb{C}$, one can define the bounded operator $B = \int f(\chi) d\pi(\chi)$ uniquely via the sesquilinear form $\langle \phi, B\psi \rangle = \int f(\chi) d\mu_{\phi, \psi}(\chi)$. Theorem 4.48 applies this construction to the function $f_\chi(\chi) = \chi(g)$ (which is bounded since $|\chi(g)| = 1$).
- **Regularity:** The uniqueness statement includes regularity of the PVM, which is a technical condition related to approximation properties needed to ensure well-definedness and compatibility with the topology of \hat{G} and associated measures.
- **Support and Spectrum:** The support of the PVM π can be considered the spectrum of the representation. If the support is discrete (a finite or countable set of characters $\{\chi_j\}$), then π is atomic, $\pi(E) = \sum_{\chi_j \in E} \pi_j$ where $\pi_j = \pi(\{\chi_j\})$ is the projection onto the eigenspace $W_{\chi_j} = \{w \mid A(g)w = \chi_j(g)w \text{ for all } g\}$. The Hilbert space decomposes into an orthogonal direct sum $\mathcal{H} = \bigoplus_j W_{\chi_j}$, and the integral reduces to the sum $A(g) = \sum_j \chi_j(g) \pi_j$. This recovers the finite-dimensional result. If the support of π has a continuous part, the decomposition involves direct integrals of Hilbert spaces.

The spectral theorem for LCA groups is linked to abstract harmonic analysis and generalized Fourier transforms.

- **Fourier Transform as Spectral Decomposition:** The classical Fourier transform on \mathbb{R}^n (or Fourier series on \mathbb{T}^n) arises from applying Theorem 4.48 to the (left) regular representation of $G = \mathbb{R}^n$ (or $G = \mathbb{T}^n$) on $\mathcal{H} = L^2(G)$. The representation is given by translation: $(A(g)f)(h) = f(g^{-1}h)$. The theorem decomposes $L^2(G)$ over the characters $\chi \in \hat{G}$. The unitary operator that implements this decomposition (diagonalizes the translation operators into multiplication operators by characters) is precisely the Fourier-Plancherel transform $\mathcal{F} : L^2(G) \rightarrow L^2(\hat{G})$. The spectral theorem essentially becomes the statement of the Plancherel theorem and Fourier inversion for LCA groups.
- **Decomposition of Functions/Signals:** Just as Fourier analysis decomposes a signal into its frequency components, the spectral theorem decomposes a unitary representation into components corresponding to the irreducible characters of the group (which are the important “frequencies” or “modes” associated with the group’s structure).

Overall, the spectral theorem for continuous unitary representations of LCA groups provides a powerful canonical decomposition, reducing the study of the representation $A : G \rightarrow U(\mathcal{H})$ to understanding the structure of the dual group \hat{G} and the distribution of the representation across \hat{G} as described by the spectral measure π . It is a powerful result unifying algebra (group structure), topology (LCA structure, continuity), and analysis (Hilbert spaces, measure theory, integration). It is particularly useful for the quantum mechanical description of systems with abelian symmetries.

4.4 Spectral Theory on Hilbert Space

The spectral theory developed for diagonalizable operators in finite-dimensional spaces is useful, but its full power only becomes apparent in the transition to infinite-dimensional complex Hilbert spaces \mathcal{H} . This generalization is very useful for modern physics, particularly quantum mechanics, where state spaces are typically infinite-dimensional function spaces like $L^2(\mathbb{R}^n)$, and key observables correspond to operators that are not only infinite-dimensional but also often unbounded. This transition, however, introduces many complications. We sketch this generalization, highlighting the key challenges and results. Further details, rigorous proofs, and broader context can be found in standard texts such as [28].

4.4.1 Challenges in Infinite Dimensions: Unboundedness and Spectrum

Several crucial distinctions arise when moving from finite to infinite dimensions.

In finite dimensions, every linear operator $A : W \rightarrow W$ is automatically bounded; that is, there exists a constant $C \geq 0$ such that $\|Aw\| \leq C\|w\|$ for all $w \in W$. In infinite dimensions, many physically and mathematically crucial operators fail to be bounded.

Definition 4.50 (Bounded and Unbounded Operators). Let A be a linear operator defined on a subspace $\text{Dom}(A) \subseteq \mathcal{H}$, mapping into \mathcal{H} .

- A is **bounded** if $\text{Dom}(A) = \mathcal{H}$ and there exists $C \geq 0$ such that $\|A\psi\| \leq C\|\psi\|$ for all $\psi \in \mathcal{H}$. The smallest such C is the operator norm $\|A\|$.
- A is **unbounded** if it is not bounded. This typically occurs because either A is only defined on a proper subspace $\text{Dom}(A) \neq \mathcal{H}$ (usually required to be dense for interesting theory), or $\|A\psi\|/\|\psi\|$ is not bounded as ψ ranges over non-zero vectors in $\text{Dom}(A)$.

Key operators in quantum mechanics, such as position $(Q\psi)(x) = x\psi(x)$ and momentum $(P\psi)(x) = -i\hbar \frac{d\psi}{dx}$ on $\mathcal{H} = L^2(\mathbb{R})$, are canonical examples of unbounded operators. For instance, applying P to sharply peaked wavefunctions can yield functions with arbitrarily large L^2 norms, showing P cannot be bounded. Unbounded operators require careful handling of their domains.

Definition 4.51 (Domain, Adjoint, Symmetric, Self-Adjoint). Let A be a linear operator with a domain $\text{Dom}(A)$ that is dense in \mathcal{H} .

- The **adjoint** A^* of A is defined as follows: Its domain $\text{Dom}(A^*)$ consists of all $\phi \in \mathcal{H}$ for which there exists a unique $\eta_\phi \in \mathcal{H}$ such that $\langle A\psi, \phi \rangle = \langle \psi, \eta_\phi \rangle$ for all $\psi \in \text{Dom}(A)$. For such ϕ , we define $A^*\phi = \eta_\phi$. The requirement that $\text{Dom}(A)$ is dense ensures η_ϕ (and thus $A^*\phi$) is unique if it exists. A^* is always a closed operator.
- A is **symmetric** (or Hermitian) if $\langle A\psi, \phi \rangle = \langle \psi, A\phi \rangle$ for all $\psi, \phi \in \text{Dom}(A)$. This is equivalent to $A \subseteq A^*$, meaning $\text{Dom}(A) \subseteq \text{Dom}(A^*)$ and $A^*\psi = A\psi$ for all $\psi \in \text{Dom}(A)$.
- A is **self-adjoint** if $A = A^*$. This requires both symmetry ($A \subseteq A^*$) and the domain condition $\text{Dom}(A) = \text{Dom}(A^*)$.

Self-adjointness is a significantly more restrictive condition than symmetry in infinite dimensions. A symmetric operator may have many self-adjoint extensions or none at all. Boundary conditions often play a crucial role in determining the domain of the adjoint and thus whether an operator is

self-adjoint. Self-adjointness is the condition required for an operator to be an observable in quantum mechanics and for the spectral theorem (and Stone's theorem) to hold in its strongest form.

The spectrum also requires slightly more detail.

Definition 4.52 (Resolvent and Spectrum). Let A be a (possibly unbounded) linear operator on \mathcal{H} with dense domain $\text{Dom}(A)$.

- The **resolvent set** $\rho(A)$ is the set of $\lambda \in \mathbb{C}$ such that the operator $(A - \lambda \text{id})$ has a bounded inverse $(A - \lambda \text{id})^{-1} : \mathcal{H} \rightarrow \mathcal{H}$ defined on all of \mathcal{H} .
- The **spectrum** $\sigma(A)$ is the complement of the resolvent set: $\sigma(A) = \mathbb{C} \setminus \rho(A)$.

For self-adjoint operators, the spectrum is always a closed subset of the real line, $\sigma(A) \subseteq \mathbb{R}$.

In infinite dimensions, the spectrum $\sigma(A)$ can be decomposed into different parts:

- **Point Spectrum** $\sigma_p(A)$: The set of eigenvalues of A . $\lambda \in \sigma_p(A)$ if $(A - \lambda \text{id})$ is not injective (i.e., $\ker(A - \lambda \text{id}) \neq \{0\}$).
- **Continuous Spectrum** $\sigma_c(A)$: The set of $\lambda \in \mathbb{C}$ such that $(A - \lambda \text{id})$ is injective, its range $\text{Range}(A - \lambda \text{id})$ is dense in \mathcal{H} , but the range is not equal to \mathcal{H} (implying the inverse exists but is unbounded).
- **Residual Spectrum** $\sigma_r(A)$: The set of $\lambda \in \mathbb{C}$ such that $(A - \lambda \text{id})$ is injective, but its range $\text{Range}(A - \lambda \text{id})$ is not dense in \mathcal{H} .

These three sets are disjoint and their union is $\sigma(A)$. A crucial simplification for self-adjoint operators is that their residual spectrum is empty: $\sigma_r(A) = \emptyset$. Their spectrum consists only of point and continuous spectrum (where the continuous spectrum is sometimes further decomposed into absolutely continuous and singular continuous parts, $\sigma_c(A) = \sigma_{ac}(A) \cup \sigma_{sc}(A)$).

Absence of Eigenvectors in \mathcal{H} : A striking feature of infinite dimensions is that self-adjoint operators corresponding to important physical quantities might possess purely continuous spectra, meaning they have no eigenvectors belonging to the Hilbert space \mathcal{H} itself.

Example 4.53 (Position Operator Q). Consider $A = Q$ acting as $(Q\psi)(x) = x\psi(x)$ on $\mathcal{H} = L^2(\mathbb{R})$. Its domain consists of $\psi \in L^2(\mathbb{R})$ such that $x\psi(x)$ is also in $L^2(\mathbb{R})$. Q is self-adjoint on this domain. Its spectrum is the entire real line, $\sigma(Q) = \mathbb{R}$. However, Q has no eigenvalues in $L^2(\mathbb{R})$. If $Q\psi = \lambda\psi$, then $(x - \lambda)\psi(x) = 0$ for almost every x . This forces $\psi(x) = 0$ for $x \neq \lambda$. A function non-zero only at a single point has L^2 norm zero, so the only solution in $L^2(\mathbb{R})$ is $\psi = 0$. Thus $\sigma_p(Q) = \emptyset$, and $\sigma(Q) = \sigma_c(Q) = \mathbb{R}$. (Physicists often work with non-normalizable “eigenfunctions” like Dirac delta functions $\delta(x - \lambda)$ satisfying $x\delta(x - \lambda) = \lambda\delta(x - \lambda)$, but these are distributions, not elements of $L^2(\mathbb{R})$.)

The existence of continuous spectra and the potential absence of a spanning set of eigenvectors necessitates a spectral theory founded on projection-valued measures, which can handle continuous decompositions of the identity.

4.4.2 Spectral Measures on Hilbert Space

Earlier, we presented the PVM formalism in Definition 4.27, which gives us the core object for the spectral theorem. Now, let's discuss this in more detail, starting with some properties.

Remark 4.54 (Properties of PVMs).

- *Property (3) immediately implies that $\pi(E_1)$ and $\pi(E_2)$ commute for any $E_1, E_2 \in \Sigma$.*
- *If $E_1 \cap E_2 = \emptyset$, then $\pi(E_1)\pi(E_2) = \pi(E_1 \cap E_2) = \pi(\emptyset) = 0$. Since $\pi(E_1)$ and $\pi(E_2)$ are orthogonal projections, this means their images are orthogonal subspaces: $\text{Im}(\pi(E_1)) \perp \text{Im}(\pi(E_2))$.*
- *Monotonicity: If $E_1 \subseteq E_2$, then $\pi(E_1) \leq \pi(E_2)$, meaning $\text{Im}(\pi(E_1)) \subseteq \text{Im}(\pi(E_2))$, which is equivalent to $\pi(E_1)\pi(E_2) = \pi(E_1) = \pi(E_2)\pi(E_1)$. This follows from $\pi(E_1) = \pi(E_1 \cap E_2) = \pi(E_1)\pi(E_2)$.*

Spectral measures allow the definition of associated scalar and vector-valued measures, which are essential for integration.

Proposition 4.55 (Associated Scalar Measures). *Let $\pi : \Sigma \rightarrow \text{Proj}^\perp(\mathcal{H})$ be a spectral measure.*

1. *For any $\psi \in \mathcal{H}$, the set function $\mu_\psi : \Sigma \rightarrow [0, \infty)$ defined by*

$$\mu_\psi(E) := \langle \psi, \pi(E)\psi \rangle = \|\pi(E)\psi\|^2$$

is a finite positive measure on (X, Σ) with total mass $\mu_\psi(X) = \|\psi\|^2$.

2. *For any $\phi, \psi \in \mathcal{H}$, the set function $\mu_{\phi, \psi} : \Sigma \rightarrow \mathbb{C}$ defined by*

$$\mu_{\phi, \psi}(E) := \langle \phi, \pi(E)\psi \rangle$$

is a complex measure on (X, Σ) . It can be recovered from the positive measures via polarization:

$$\mu_{\phi, \psi} = \frac{1}{4} \sum_{k=0}^3 i^k \mu_{\phi + i^k \psi}.$$

Proof. Countable additivity of μ_ψ and $\mu_{\phi, \psi}$ follows directly from the countable additivity of π (in SOT) and the linearity/continuity of the inner product. For μ_ψ : $\mu_\psi(\cup E_i) = \langle \psi, \pi(\cup E_i)\psi \rangle = \langle \psi, \sum \pi(E_i)\psi \rangle = \sum \langle \psi, \pi(E_i)\psi \rangle = \sum \mu_\psi(E_i)$. Finiteness $\mu_\psi(X) = \langle \psi, \pi(X)\psi \rangle = \langle \psi, \text{id} \psi \rangle = \|\psi\|^2$. Similarly for $\mu_{\phi, \psi}$. \square

These scalar measures pave the way for defining integrals of functions with respect to the operator-valued measure π .

4.4.3 Strongly Continuous One-Parameter Unitary Groups and Stone's Theorem

A crucial link between self-adjoint operators (potentially unbounded) and bounded unitary operators is provided by Stone's theorem on one-parameter unitary groups. This generalizes the relationship $A = A^* \leftrightarrow U_t = e^{itA}$ observed in finite dimensions (Section 4.3.4). We first need the correct notion of continuity for families of operators on \mathcal{H} .

Definition 4.56 (Strong Continuity). Let \mathcal{H} be a separable Hilbert space and $U(\mathcal{H})$ be the group of unitary operators on \mathcal{H} . A **strongly continuous one-parameter unitary group** is a group homomorphism

$$U : (\mathbb{R}, +) \longrightarrow U(\mathcal{H}) \tag{4.12}$$

(satisfying $U_0 = \text{id}$ and $U_{s+t} = U_s U_t$) such that the map $t \mapsto U_t$ is continuous with respect to the **strong operator topology (SOT)**. This means that for every fixed vector $\psi \in \mathcal{H}$, the map

$$t \mapsto U_t \psi$$

is continuous as a map from \mathbb{R} to \mathcal{H} (using the norm topology on \mathcal{H}).

Remark 4.57 (Topologies). *SOT-continuity is weaker than norm-continuity (which would require $\|U_t - U_s\| \rightarrow 0$ as $t \rightarrow s$). Norm-continuity implies that the generator must be bounded. SOT-continuity allows for unbounded generators, which is essential for physics.*

Stone's theorem establishes an important bijective correspondence.

Theorem 4.58 (Stone's Theorem on One-Parameter Unitary Groups). *Let \mathcal{H} be a complex Hilbert space. There is a one-to-one correspondence between self-adjoint operators A on \mathcal{H} and strongly continuous one-parameter unitary groups $U : \mathbb{R} \rightarrow U(\mathcal{H})$. The correspondence is given by*

$$U_t = e^{itA}$$

where e^{itA} is defined via the spectral theorem and Borel functional calculus for the self-adjoint operator A . Conversely, given the group U_t , its unique self-adjoint **infinitesimal generator** A is obtained by

$$A\psi = \lim_{h \rightarrow 0} \frac{U_h \psi - \psi}{ih}$$

where the domain of A , $\text{Dom}(A)$, consists precisely of those vectors $\psi \in \mathcal{H}$ for which the limit exists (in the norm topology of \mathcal{H}). This domain is always a dense subspace of \mathcal{H} . Equivalently,

$$A\psi = \frac{1}{i} \left. \frac{d}{dt} U_t \psi \right|_{t=0}$$

where the derivative is taken in the strong sense (limit in norm).

Proof Ideas. This is a deep theorem in functional analysis.

$(A \implies U_t)$: Given a self-adjoint A , the spectral theorem (Section 4.4.4) allows definition of $U_t = e^{itA}$ using the Borel functional calculus for the function $f(\lambda) = e^{it\lambda}$. Since A is self-adjoint and $t \in \mathbb{R}$, f is a bounded Borel function with $|f(\lambda)| = 1$. The functional calculus properties ensure U_t is unitary, $U_{s+t} = U_s U_t$, and $U_0 = I$. Proving strong continuity requires using the properties of integration against the spectral measure π_A and dominated convergence arguments for the scalar measures μ_ψ .

$(U_t \implies A)$: Given U_t , one defines A via the strong derivative as in the theorem statement. Proving that $\text{Dom}(A)$ is dense and, critically, that the resulting operator A is self-adjoint (not just symmetric) is the non-trivial part. Techniques often involve analyzing the resolvent of A or using Fourier analysis (e.g., Bochner's theorem relating positive definite functions to measures). The uniqueness of the generator is also very important. \square

Stone's theorem provides the rigorous mathematical foundation for quantum dynamics. In quantum mechanics, the state of a system evolves according to the Schrödinger equation $i\hbar \frac{d}{dt} \psi(t) = H\psi(t)$, where H is the Hamiltonian operator. For the evolution to be physically meaningful (preserving

probabilities, i.e., $\|\psi(t)\|^2 = \|\psi(0)\|^2$, the evolution operator $U(t)$ relating $\psi(t) = U(t)\psi(0)$ must be unitary. Stone's theorem guarantees that such a unitary evolution group $U(t)$ exists and satisfies the Schrödinger equation (in integrated form $U(t) = e^{-iHt/\hbar}$) if and only if the Hamiltonian H is a self-adjoint operator. The possibility of H being unbounded is essential for realistic physical systems.

4.4.4 The Spectral Theorem for Self-Adjoint Operators on Hilbert Space

This theorem is arguably the most important result in the theory of operators on Hilbert space, providing a canonical representation for any self-adjoint operator (bounded or unbounded) analogous to the diagonalization of Hermitian matrices. It asserts that every self-adjoint operator can be represented as multiplication by a real variable, possibly after a change of representation, or equivalently, can be reconstructed from its spectral measure.

The infinite-dimensional PVM formulation directly generalizes the finite-dimensional version using the PVM defined on Borel sets.

Theorem 4.59 (Spectral Theorem - PVM Form). *Let A be a self-adjoint operator on a separable complex Hilbert space \mathcal{H} . There exists a unique spectral measure (PVM) π_A , defined on the Borel σ -algebra $\mathcal{B}(\mathbb{R})$ of the real line, taking values in $\text{Proj}^\perp(\mathcal{H})$, such that:*

1. π_A is supported on the spectrum $\sigma(A) \subseteq \mathbb{R}$. That is, $\pi_A(\mathbb{R} \setminus \sigma(A)) = 0$.
2. The operator A can be represented as an integral with respect to π_A :

$$A = \int_{\mathbb{R}} \lambda d\pi_A(\lambda) = \int_{\sigma(A)} \lambda d\pi_A(\lambda). \quad (4.13)$$

The integral and the domain of A are defined precisely as:

$$\text{Dom}(A) = \left\{ \psi \in \mathcal{H} : \int_{\sigma(A)} |\lambda|^2 d\mu_\psi(\lambda) < \infty \right\}$$

where $\mu_\psi(E) = \|\pi_A(E)\psi\|^2$. For $\psi \in \text{Dom}(A)$, $A\psi$ is defined uniquely by the relation

$$\langle \phi, A\psi \rangle = \int_{\sigma(A)} \lambda d\mu_{\phi, \psi}(\lambda) \quad \text{for all } \phi \in \mathcal{H},$$

where $\mu_{\phi, \psi}(E) = \langle \phi, \pi_A(E)\psi \rangle$.

Proof. See Theorem VII.6 on page 263 of [28]. □

This formulation provides a powerful “continuous diagonalization.” The Hilbert space \mathcal{H} is decomposed infinitesimally according to the spectral values $\lambda \in \sigma(A)$, with $\pi_A(E)$ projecting onto the subspace associated with spectral values in E . A acts like multiplication by λ on the “infinitesimal eigenspace” corresponding to λ .

Multiplication Operator Formulation: This version provides a more concrete representation, stating that any self-adjoint operator looks like multiplication by the independent variable on some L^2 space, up to a change of basis (unitary transformation).

Theorem 4.60 (Spectral Theorem - Multiplication Operator Form). *Let A be a self-adjoint operator on a separable complex Hilbert space \mathcal{H} . Then there exist a measure space (X, Σ, μ) (where μ is a σ -finite measure), a real-valued measurable function $g : X \rightarrow \mathbb{R}$, and a unitary operator $U : \mathcal{H} \rightarrow L^2(X, \mu)$ such that:*

1. U maps the domain of A onto the domain of the multiplication operator M_g :

$$U(\text{Dom}(A)) = \text{Dom}(M_g) = \left\{ f \in L^2(X, \mu) : \int_X |g(x)|^2 |f(x)|^2 d\mu(x) < \infty \right\}.$$

2. A is unitarily equivalent to M_g : For any $\psi \in \text{Dom}(A)$,

$$(UA\psi)(x) = g(x)(U\psi)(x) \quad \text{or equivalently} \quad UAU^{-1} = M_g$$

where $(M_g f)(x) = g(x)f(x)$.

Furthermore, the measure space can often be chosen more specifically. For instance, one can decompose X based on spectral multiplicity, and often take X to be (a subset of) \mathbb{R} and $g(\lambda) = \lambda$.

Sources. See references cited for the PVM form, as the two are essentially equivalent. This form is often derived by constructing suitable “cyclic vectors” ψ and considering the space generated by applying functions of A to ψ , then using the Riesz-Markov theorem to represent linear functionals related to A as integration against a measure μ_ψ . A direct sum over such cyclic subspaces yields the full representation. \square

This formulation makes the analogy with finite-dimensional diagonalization explicit: U acts as the change-of-basis matrix (to the “basis” where A is diagonal), and M_g is the “diagonal” form, where multiplication by $g(x)$ replaces multiplication by eigenvalues λ_j . The spectral properties of A (point vs. continuous spectrum) are reflected in the properties of the measure μ (atoms vs. continuous part) and the function g .

Borel Functional Calculus: A crucial consequence of both formulations of the spectral theorem is the ability to define $f(A)$ for a wide class of functions f , extending the functional calculus developed for polynomials or continuous functions.

Definition 4.61 (Borel Functional Calculus). Let A be a self-adjoint operator with spectral measure π_A . For any Borel measurable function $f : \mathbb{R} \rightarrow \mathbb{C}$, the operator $f(A)$ is defined via the spectral integral:

$$f(A) := \int_{\mathbb{R}} f(\lambda) d\pi_A(\lambda) = \int_{\sigma(A)} f(\lambda) d\pi_A(\lambda).$$

Its domain is

$$\text{Dom}(f(A)) = \left\{ \psi \in \mathcal{H} : \int_{\sigma(A)} |f(\lambda)|^2 d\mu_\psi(\lambda) < \infty \right\}.$$

Equivalently, using the multiplication operator form $UAU^{-1} = M_g$, we define

$$f(A) := U^{-1}M_{f \circ g}U$$

with domain $U^{-1}(\text{Dom}(M_{f \circ g}))$.

This functional calculus is remarkably robust.

Theorem 4.62 (Properties of Borel Functional Calculus). *Let A be self-adjoint. The map $\Psi : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ (Borel functions to operators) defined by $\Psi(f) = f(A)$ has the following properties (among others):*

1. *It is an algebra homomorphism for bounded functions (linearity, multiplicativity $\Psi(fg) = \Psi(f)\Psi(g)$). For unbounded functions, care with domains is needed, but composition often works as expected, e.g., $(f \circ g)(A) = f(g(A))$ under suitable conditions.*
2. *It is a $*$ -homomorphism: $(f(A))^* = \bar{f}(A)$. This implies $f(A)$ is self-adjoint if f is real-valued, and $f(A)$ is unitary if $|f(\lambda)| = 1$ for $\lambda \in \sigma(A)$.*
3. *Norm property: If f is bounded, $\|f(A)\| = \sup_{\lambda \in \sigma(A)} |f(\lambda)|$.*
4. *Supports composition: If g is Borel and f is Borel, $(f \circ g)(A) = f(g(A))$ under appropriate domain conditions.*
5. *Spectral Mapping Theorem: For reasonable functions f (e.g., continuous, or Borel with care), $\sigma(f(A)) = \overline{f(\sigma(A))}$ (closure of the image of the spectrum under f).*

This functional calculus allows us to rigorously define operators like e^{itA} (for Stone's theorem), \sqrt{A} (for positive operators), projection operators $\mathbf{1}_E(A) = \pi_A(E)$, and propagators in physics.

4.4.5 Connection to C^* -Algebras and the Functional Calculus for Bounded Normal Operators

While the general spectral theorem, particularly in its PVM formulation, provides a comprehensive framework applicable even to unbounded self-adjoint operators, there exists an alternative and powerfully elegant approach rooted in the abstract theory of C^* -algebras for the important special case of *bounded normal operators*. This algebraic perspective not only recovers the spectral theorem for this class but also provides useful ideas for studying the representation of operator algebras as function algebras. It's very important for noncommutative geometry and abstract harmonic analysis.

Let \mathcal{H} be a complex Hilbert space. Recall that $B(\mathcal{H})$ denotes the space of all bounded linear operators $A : \mathcal{H} \rightarrow \mathcal{H}$.

Definition 4.63 (C^* -Algebra). A C^* -**algebra** is a complex Banach algebra \mathcal{A} (i.e., a complex algebra that is also a Banach space with a submultiplicative norm: $\|ab\| \leq \|a\|\|b\|$) equipped with an involution map $*$: $\mathcal{A} \rightarrow \mathcal{A}$ satisfying for all $a, b \in \mathcal{A}$ and $\lambda \in \mathbb{C}$:

1. $(a^*)^* = a$ (Involution property)
2. $(ab)^* = b^*a^*$ (Anti-multiplicativity)
3. $(\lambda a + b)^* = \bar{\lambda}a^* + b^*$ (Conjugate-linearity)
4. The C^* -**identity**: $\|a^*a\| = \|a\|^2$.

Remark 4.64 (Significance of C^* -identity). *The C^* -identity is a remarkably strong condition linking the norm and the algebraic structure ($*$ -operation). It implies, among other things, that the involution is an isometry ($\|a^*\| = \|a\|$), and it guarantees that abstract C^* -algebras can always be faithfully represented as norm-closed, self-adjoint subalgebras of $B(\mathcal{H})$ for some Hilbert space \mathcal{H} (Gelfand-Naimark-Segal construction).*

Example 4.65 (Canonical Examples of C^* -Algebras).

1. The algebra $B(\mathcal{H})$ of bounded linear operators on a Hilbert space \mathcal{H} with the operator norm and the Hilbert space adjoint operation is the prototypical example of a C^* -algebra.
2. For any compact Hausdorff space X , the algebra $C(X)$ of continuous complex-valued functions on X is a commutative C^* -algebra with pointwise addition and multiplication, complex conjugation $f^*(x) = \overline{f(x)}$ as the involution, and the supremum norm $\|f\|_\infty = \sup_{x \in X} |f(x)|$ as the norm. The C^* -identity $\|f^*f\| = \| |f|^2 \|_\infty = (\sup |f(x)|)^2 = \|f\|_\infty^2$ holds.

The structure of *commutative* unital C^* -algebras is completely characterized by the Gelfand-Naimark theorem.

Theorem 4.66 (Gelfand-Naimark Theorem for Commutative C^* -Algebras). *Let A be a commutative unital C^* -algebra. Let $X = \hat{A}$ be the **spectrum** (or **maximal ideal space**, or **character space**) of A , consisting of all non-zero algebra homomorphisms $\phi : A \rightarrow \mathbb{C}$ (characters), equipped with the weak-* topology (the coarsest topology making the evaluation maps $\hat{a} : \phi \mapsto \phi(a)$ continuous for all $a \in A$). Then X is a non-empty compact Hausdorff space. The **Gelfand transform** $\gamma : A \rightarrow C(X)$ defined by*

$$\gamma(a)(\phi) = \phi(a) \quad (\text{for } a \in A, \phi \in X)$$

*is an isometric *-isomorphism from A onto the C^* -algebra $C(X)$ of continuous functions on X .*

Remark 4.67 (Interpretation). *This important theorem asserts that every abstract commutative unital C^* -algebra can be concretely realized as the algebra of continuous functions on some compact Hausdorff space, its spectrum. The algebraic and topological structure of A is completely encoded in the topological structure of X and the function algebra $C(X)$. The norm satisfies $\|\gamma(a)\|_\infty = \sup_{\phi \in X} |\phi(a)| = \|a\|$ (isometry), and the involution corresponds to complex conjugation $\gamma(a^*) = \overline{\gamma(a)}$.*

Now, let's apply this powerful theorem to operators. Consider a *bounded normal operator* N on a Hilbert space \mathcal{H} . Recall normal means $NN^* = N^*N$.

Definition 4.68 (C^* -Algebra Generated by a Normal Operator). Let $N \in B(\mathcal{H})$ be a normal operator. The C^* -algebra generated by N and the identity I , denoted $C^*(N, I)$, is the smallest C^* -subalgebra of $B(\mathcal{H})$ that contains both N and I . It can be constructed by taking all polynomials $p(N, N^*)$ in N and its adjoint N^* , and then taking the closure with respect to the operator norm.

Proposition 4.69. *If N is a normal operator, then the C^* -algebra $C^*(N, I)$ is commutative.*

Proof. Since N is normal, N commutes with N^* . Any polynomial $p(N, N^*)$ involves sums of terms of the form $c_{kl}N^k(N^*)^l$. Since N and N^* commute, any two such polynomials $p(N, N^*)$ and $q(N, N^*)$ also commute: $p(N, N^*)q(N, N^*) = q(N, N^*)p(N, N^*)$. The set of these polynomials is dense in $C^*(N, I)$ by construction. Since multiplication is continuous with respect to the norm topology, the commutativity extends to the norm closure. Hence, $C^*(N, I)$ is a commutative C^* -algebra. \square

Since $C^*(N, I)$ is a commutative unital C^* -algebra, the Gelfand-Naimark theorem applies. There

exists an isometric $*$ -isomorphism:

$$\gamma : C^*(N, I) \xrightarrow{\cong} C(X)$$

where $X = \widehat{C^*(N, I)}$ is the spectrum of the algebra. A crucial identification connects the abstract spectrum X to the standard operator spectrum $\sigma(N)$.

Theorem 4.70 (Spectrum Identification). *Let $N \in B(\mathcal{H})$ be a normal operator. The spectrum $X = \widehat{C^*(N, I)}$ of the C^* -algebra $C^*(N, I)$ is homeomorphic to the operator spectrum $\sigma(N) \subset \mathbb{C}$ via the map $\phi \mapsto \phi(N)$.*

Idea. A character $\phi \in X$ is determined by its value $\lambda = \phi(N)$. One shows that $N - \lambda I$ must be non-invertible in $C^*(N, I)$ (because $\phi(N - \lambda I) = \phi(N) - \lambda\phi(I) = \lambda - \lambda = 0$, and characters map invertible elements to non-zero numbers), which implies $\lambda \in \sigma(N)$. Conversely, for any $\lambda \in \sigma(N)$, one can construct a character ϕ_λ such that $\phi_\lambda(N) = \lambda$. The map $\phi \mapsto \phi(N)$ is shown to be a homeomorphism. \square

Combining these results yields the Gelfand isomorphism tailored to the operator N :

$$\gamma : C^*(N, I) \xrightarrow{\cong} C(\sigma(N))$$

This isomorphism maps the operator N itself to the identity function $f(\lambda) = \lambda$ restricted to the compact set $\sigma(N) \subset \mathbb{C}$. Similarly, N^* maps to the function $f(\lambda) = \bar{\lambda}$. More generally, a polynomial $p(N, N^*)$ maps to the function $p(\lambda, \bar{\lambda})|_{\sigma(N)}$.

This isomorphism immediately provides a rigorous way to apply any *continuous* function to the operator N .

Definition 4.71 (Continuous Functional Calculus). Let $N \in B(\mathcal{H})$ be a normal operator. For any continuous function $f \in C(\sigma(N))$, the operator $f(N) \in C^*(N, I) \subset B(\mathcal{H})$ is defined via the inverse Gelfand transform:

$$f(N) := \gamma^{-1}(f)$$

This map $f \mapsto f(N)$ is called the **continuous functional calculus** for N .

Theorem 4.72 (Properties of Continuous Functional Calculus). *The continuous functional calculus $\Psi_c : C(\sigma(N)) \rightarrow C^*(N, I)$ defined by $\Psi_c(f) = f(N)$ is an isometric $*$ -isomorphism satisfying:*

1. *It extends the polynomial functional calculus: if $p(z)$ is a polynomial, $\Psi_c(p|_{\sigma(N)}) = p(N)$. More generally, if $p(z, w)$ is a polynomial in two variables, $\Psi_c(p(\lambda, \bar{\lambda})|_{\sigma(N)}) = p(N, N^*)$.*
2. *It is a $*$ -homomorphism: $(f + g)(N) = f(N) + g(N)$, $(fg)(N) = f(N)g(N)$, $\bar{f}(N) = (f(N))^*$.*
3. *It is isometric: $\|f(N)\| = \|f\|_\infty = \sup_{\lambda \in \sigma(N)} |f(\lambda)|$.*
4. *Spectral Mapping Theorem: $\sigma(f(N)) = f(\sigma(N))$ for all $f \in C(\sigma(N))$.*

Proof. These properties follow directly from the fact that γ is an isometric $*$ -isomorphism and properties of the Gelfand transform (e.g., mapping spectrum to range for functions). \square

The continuous functional calculus is powerful but limited to continuous functions f . To recover the spectral theorem involving projections (which correspond to discontinuous characteristic functions $\mathbf{1}_E$) and to integrate arbitrary measurable functions, we need to extend this calculus.

This extension relies on the Riesz-Markov-Kakutani representation theorem. For any vector $\psi \in \mathcal{H}$, the map $f \mapsto \langle \psi, f(N)\psi \rangle$ defines a positive linear functional on $C(\sigma(N))$. By Riesz-Markov, there exists a unique finite positive regular Borel measure μ_ψ on $\sigma(N)$ such that

$$\langle \psi, f(N)\psi \rangle = \int_{\sigma(N)} f(\lambda) d\mu_\psi(\lambda) \quad \text{for all } f \in C(\sigma(N)).$$

Using polarization, we can define complex measures $\mu_{\phi, \psi}$ such that $\langle \phi, f(N)\psi \rangle = \int_{\sigma(N)} f(\lambda) d\mu_{\phi, \psi}(\lambda)$.

These measures allow us to define $f(N)$ for any bounded Borel measurable function $f : \sigma(N) \rightarrow \mathbb{C}$. The operator $f(N)$ is defined uniquely via the sesquilinear form:

$$\langle \phi, f(N)\psi \rangle := \int_{\sigma(N)} f(\lambda) d\mu_{\phi, \psi}(\lambda)$$

This extended map $f \mapsto f(N)$ from the algebra $\mathcal{B}_b(\sigma(N))$ of bounded Borel functions on $\sigma(N)$ to $B(\mathcal{H})$ is still a *-homomorphism.

Now, we can define the PVM. For any Borel set $E \subseteq \sigma(N)$, let $\mathbf{1}_E$ be its characteristic function (which is a bounded Borel function). Define the projection operator $\pi_N(E)$ via the Borel functional calculus:

$$\pi_N(E) := \mathbf{1}_E(N) = \int_{\sigma(N)} \mathbf{1}_E(\lambda) d\pi_N(\lambda)$$

One can verify that $\pi_N(E)$ is indeed an orthogonal projection ($\pi_N(E)^2 = \pi_N(E)$, $\pi_N(E)^* = \pi_N(E)$) and that the map $E \mapsto \pi_N(E)$ satisfies the properties of a spectral measure (Definition 4.27).

Finally, applying the Borel functional calculus to the identity function $f(\lambda) = \lambda$, we recover the spectral theorem integral representation for the bounded normal operator N :

$$N = \text{id}_{\mathbb{C}}(N) = \int_{\sigma(N)} \lambda d\pi_N(\lambda)$$

Thus, the abstract C^* -algebra approach, via Gelfand theory and extension using Riesz-Markov, provides an alternative pathway to the spectral theorem for the important class of bounded normal operators.

This C^* -algebraic method relies inherently on the operator (and its generated algebra) being bounded (specifically, residing within a C^* -algebra). It does not directly apply to unbounded self-adjoint operators like position, momentum, or typical Hamiltonians. For these, the spectral theorem requires the more direct PVM or multiplication operator formulations developed earlier, which handle domain issues intrinsically.

Nonetheless, the C^* -algebra perspective provides a lot of information on the structure: the spectral theorem importantly diagonalizes the operator by representing the commutative algebra it generates as an algebra of functions on its spectrum. This viewpoint motivates generalizations in noncommutative geometry, where one studies noncommutative C^* -algebras as if they were function algebras on hypothetical “noncommutative spaces,” extending geometric and topological concepts beyond the classical commutative setting.

References

- [1] Ralph Abraham and Jerrold E. Marsden. *Foundations of Mechanics*. Second. Reading, Mass.: Benjamin/Cummings Pub. Co., Advanced Book Program, 1978.
- [2] Vladimir I. Arnold. *Mathematical Methods of Classical Mechanics*. Second. Vol. 60. Graduate Texts in Mathematics. New York: Springer-Verlag, 1989.
- [3] Michèle Audin. *Geometry*. Universitext. Berlin: Springer, 2003.
- [4] Sean Bates and Alan Weinstein. *Lectures on the Geometry of Quantization*. Vol. 8. Berkeley Mathematics Lecture Notes. Providence, RI: American Mathematical Society, 1997.
- [5] John B. Conway. *A Course in Functional Analysis*. Second. Vol. 96. Graduate Texts in Mathematics. New York: Springer-Verlag, 1990.
- [6] Kenneth R. Davidson. *C^* -Algebras by Example*. Vol. 6. Fields Institute Monographs. Providence, RI: American Mathematical Society, 1996.
- [7] Paul Adrien Maurice Dirac. *The Principles of Quantum Mechanics*. Fourth Edition, Revised. The International Series of Monographs on Physics 27. Oxford University Press, 1981.
- [8] Manfredo P. Do Carmo. *Riemannian Geometry*. Boston: Birkhäuser, 1992.
- [9] David S. Dummit and Richard M. Foote. *Abstract Algebra*. Third. Hoboken, NJ: John Wiley & Sons, 2004.
- [10] L. D. Faddeev and O. A. Yakubovskii. *Lectures on Quantum Mechanics for Mathematics Students*. Vol. 47. Student Mathematical Library. Providence, RI: American Mathematical Society, 2009.
- [11] Richard P. Feynman. *The Principle of Least Action*. The Feynman Lectures on Physics, Vol. II, Chapter 19. Accessed on May 23, 2025. 1964. URL: https://www.feynmanlectures.caltech.edu/II_19.html.
- [12] Gerald B. Folland. *A Course in Abstract Harmonic Analysis*. Studies in Advanced Mathematics. Boca Raton, FL: CRC Press, 1995.
- [13] Daniel S. Freed. *Quantum Theory From A Geometric Viewpoint*. URL: <https://people.math.harvard.edu/~dafr/GQT12.pdf>.
- [14] Victor Guillemin and Shlomo Sternberg. *Symplectic Techniques in Physics*. Cambridge: Cambridge University Press, 1984.
- [15] Brian C. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Second. Vol. 222. Graduate Texts in Mathematics. Cham: Springer, 2015.
- [16] Brian C. Hall. *Quantum Theory for Mathematicians*. Vol. 267. Graduate Texts in Mathematics. New York: Springer, 2013.
- [17] Kenneth Hoffman and Ray Kunze. *Linear Algebra*. Second. Englewood Cliffs, NJ: Prentice-Hall, 1971.
- [18] Tosio Kato. *Perturbation Theory for Linear Operators*. Classics in Mathematics. Reprint of the second (1976) edition. Berlin: Springer-Verlag, 1995.
- [19] John M. Lee. *Introduction to Riemannian Manifolds*. Second. Vol. 176. Graduate Texts in Mathematics. Cham: Springer, 2018.
- [20] John M. Lee. *Introduction to Smooth Manifolds*. Second. Vol. 218. Graduate Texts in Mathematics. New York: Springer, 2013. ISBN: 978-1-4419-9981-8.
- [21] Paulette Libermann and Charles-Michel Marle. *Symplectic Geometry and Analytical Mechanics*. Dordrecht: D. Reidel Publishing Company, 1987.
- [22] George W. Mackey. *The Mathematical Foundations of Quantum Mechanics: A Lecture-Note Volume*. New York-Amsterdam: W. A. Benjamin, Inc., 1963.

- [23] Jerrold E. Marsden and Tudor S. Ratiu. *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems*. Second. Vol. 17. Texts in Applied Mathematics. New York: Springer, 1999.
- [24] Dusa McDuff and Dietmar Salamon. *Introduction to Symplectic Topology*. Third. Vol. 27. Oxford Graduate Texts in Mathematics. Oxford: Oxford University Press, 2017.
- [25] Gerard J. Murphy. *C*-Algebras and Operator Theory*. Boston: Academic Press, 1990.
- [26] John von Neumann. “On Rings of Operators. III”. In: *Collected Works, Volume III: Rings of Operators*. Ed. by A. H. Taub. Pergamon Press, 1961, pp. 6–131.
- [27] Peter J. Olver. *Applications of Lie Groups to Differential Equations*. Second. Vol. 107. Graduate Texts in Mathematics. New York: Springer-Verlag, 1993.
- [28] Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics, Vol. I: Functional Analysis*. Revised and Enlarged. San Diego: Academic Press, 1980.
- [29] Walter Rudin. *Fourier Analysis on Groups*. Interscience Tracts in Pure and Applied Mathematics 12. New York: Interscience Publishers (John Wiley & Sons), 1962.
- [30] Walter Rudin. *Functional Analysis*. Second. International Series in Pure and Applied Mathematics. New York: McGraw-Hill, 1991.
- [31] I. E. Segal. “Postulates for General Quantum Mechanics”. In: *Annals of Mathematics*. Second Series 48.4 (1947), pp. 930–948.
- [32] Ana Cannas da Silva. *Lectures on Symplectic Geometry*. Vol. 1764. Lecture Notes in Mathematics. Berlin, Heidelberg: Springer-Verlag, 2001.
- [33] Jean-Marie Souriau. *Structure of Dynamical Systems: A Symplectic View of Physics*. Trans. by C. H. Cushman-de Vries. Translated from the 1970 French original *Structure des systèmes dynamiques*. Boston: Birkhäuser, 1997.
- [34] F. Strocchi. *An Introduction to the Mathematical Structure of Quantum Mechanics*. World Scientific, 2005.
- [35] John Von Neumann. *Mathematical Foundations of Quantum Mechanics*. Investigations in Physics 2. Princeton University Press, 1955.
- [36] Frank W. Warner. *Foundations of Differentiable Manifolds and Lie Groups*. Vol. 94. Graduate Texts in Mathematics. New York: Springer-Verlag, 1983.
- [37] Nicholas M. J. Woodhouse. *Geometric Quantization*. Second. Oxford Mathematical Monographs. Oxford: Clarendon Press, 1992.