मौलिक विज्ञान प्रकर्ष केन्द्र
UM-DAE-CEBS

Semester 7 Reading
Project Report

Symplectic Geometry

JOVI K

SEM 7, PHYSICS
UM-DAE-CEBS

Project Guide: Prof. Ameeya Bhagwat

Department of Physics,

University of Mumbai – Department of Atomic Energy – Centre for Excellence in Basic Sciences

Abstract

In this report we study Symplectic Geometry, an outgrowth of classical mechanics. We give a comprehensible introduction and further provide an elementary introduction to its counterpart in mathematical physics known as geometric quantization, a way of passing from a classical mechanical system to a quantum mechanical system by methods from symplectic geometry.

Acknowledgement

I give my deepest thanks to my supervisor Prof. Ameeya Bhagwat for introducing me to this beautiful subject and for being such an unlimited source of inspiration.
I am also indebted to my fellow classmates for their careful reading of an early draft of the manuscript and for their many valuable comments and suggestions that have been incorporated in the text to improve the presentation.

Table of Contents

Introduction

"Geometry is the archetype of the beauty of the world."

Johannes Kepler

The mathematical theory underlying Hamiltonian mechanics is currently called Symplectic Geometry. It essentially started off as an outgrowth of classical mechanics. Due to its importance, it is now widely used in various fields of Mathematical Physics including Classical Mechanics, Quantum Mechanics (via geometric quantization, deformation quantization), representations of Lie groups – "the Kirillov's orbit method", PDEs (via microlocal analysis), Gauge Theory, geometric invariant theory in algebraic geometry, et cetera.

There has been a lot of active study recently in the disciplines of Symplectic Geometry. Understanding the geometry of dynamical systems and the process of "quantization" as applied not only to the theory of dynamical systems, but also as a vital tool in the analysis of group representations has made significant advances in symplectic geometry. The topic is recent versions of a themes that have dominated mathematical thought for the past three centuries - the relations between the wave and the corpuscular theories of light. We shall be taking up the subject of Symplectic geometry in whole its glory, followed by a light discussion on Geometric Quantization.

The presumed structure of this article is to start with a MOTIVATION – asking the question why are we doing what we are doing, every once in a while, followed by DEFINITIONS and THEOREMS. These may further be enumerated by REMARKS, wherever required, which is usually additional comments and examples illustrating the process.

We shall assume some elementary knowledge of Differential Topology, say, Charts and Coordinate Maps, Push forwards and Pull backs, Tangent fields and bundles, Projection and Section Maps. Some topics including Differential forms and Lie Groups are covered in the next two sections for completeness.

Unless explicitly stated otherwise, the spaces we will be dealing with are assumed to be Hausdorff, second countable with C^∞ maximal atlases. Further the associated topology, if needs be, may be assumed to be the canonical topology i.e., open balls.

We won't be necessarily differentiating between C^n and C^∞ which is because of the following theorem:

THEOREM: (Whitney) Any maximal C^k-atlas, with $k > 1$, contains a C^∞-atlas. Moreover, any two maximal C^k- atlases that contain the same C^∞-atlas are identical.

1. Differential Forms

MOTIVATION: We require integration under curves/simplexes in a differential manifold to be coordinate independent (independent of choice of charts), which naturally gives rise an important class of objects that remain invariant under coordinate transformation. Furthermore, manifolds don't inherently come with any notion of their size: we shall see the basic additional extra structure that makes these possible.

DEFINITION 1.1: An **exterior form of degree k**, or a **k-form**, denoted ω^k is a C^n differentiable map of k tangent vectors to $\mathbb{R}$, which is k-linear and skew-symmetric.

To enumerate the properties, for $\xi_i \in TM$ and a k-form, ω^k,

- K-Linear:

$$\omega^k(\xi_1, \dots, \lambda_i\xi_i + \lambda_j\xi_j, \dots \xi_i) = \lambda_i\omega^2(\xi_1, \dots, \lambda_i\xi_i + \lambda_j\xi_j, \dots \xi_i) + \lambda_j\omega^2(\xi_1, \dots, \lambda_i\xi_i + \lambda_j\xi_j, \dots \xi_i)$$

- Skew symmetric under exchange:

$$\omega^k(\xi_{i_1}, \dots, \xi_{i_k}) = \mathcal{P}^k\omega^k(\xi_1, \dots, \xi_k) \quad \text{where } \mathcal{P}^k \text{ is the permutation tensor of rank } k$$

This would mean that, say for ω^2, $\omega^2(\xi_1, \xi_2) = -\omega^2(\xi_2, \xi_1)$ i.e., this is essentially the signed area of parallelogram with sides ξ_1, ξ_2. Readers may also note the similarity between this and the 2×2 determinant, $\det(\xi_i)$ which is also an alternate representation[1] of a signed area of the parallelogram.

REMARKS –

- Geometrically, k-forms represent the oriented "volume" of k-parallelotope.
- Linearity and k-linearity are NOT the same thing. The map $f: \mathbb{R}^2 \to \mathbb{R}$ given by $(x, y) \mapsto x + y$ is linear but not bilinear, while the map $(x, y) \mapsto xy$ is bilinear but not linear.

DEFINITION 1.2: Given a 1-forms $\alpha_1, \ldots, \alpha_k$ on V, we define their k[th] **exterior product** or **wedge product** to be the k-form action on $\xi_1, \cdots, \xi_k \in TM$ by:

$$(\alpha_1 \wedge \ldots \wedge \alpha_k)(\xi_1, \cdots, \xi_k) \coloneqq \omega^k(\xi_1, \cdots, \xi_k)$$

- The wedge product is essentially an Anti-symmetrized Tensor product of dual vectors.

PROPOSITION: By the property of determinants (signed volume of vectors), $\omega^k(\xi_1, \cdots, \xi_k)$ may be treated as $\det \alpha_i(\xi_j)$.

THEOREM 1.1: For an n-dimensional vector space V, the maximal k-form which is the n-form is called the **Volume form**, and every n-form on V is a scalar multiple of $x_1 \wedge \ldots \wedge x_n$ where x_i are the dual vectors from a basis of V.

REMARK –

- We call it the volume form for the obvious reason that it measures volumes in V.
- The space of k-forms forms a vector space, $\bigwedge^k V^*$. Each such k-form on V gives us a way of measuring the signed volume inside the vector space V.

We shall be interested in tangent space, $T_x M$, to some point x of some manifold M, i.e., for us $V \equiv T_x M$.[2] And so, the form will measure the infinitesimal "areas" of a set of tangent vectors, that is to say that in our context k-forms will "eat" k tangent vectors, $\xi_i \in T_x M$ and "spit out" a real number,

$$\text{For } \omega^k \in \textstyle\bigwedge^k T_x^* M, \quad \omega^k: T_x M^{(1)} \times \ldots \times T_x M^{(k)} \to \mathbb{R}$$

As an illustration we shall write a general k-form in local coordinates:

Suppose $(U \subset X, \phi)$ be a local coordinate chart of k-dimensional manifold X and denote the coordinate maps $\phi^i \equiv x_i$ (functions that map points in submanifold U to $\mathbb{R}$). And we know that the tangent vector at some point $\bar{y} \in U$ may be represented as $\left.\frac{\partial}{\partial x_i}\right|_{\bar{y}}$, where $\bar{y} \equiv (y_1, \cdots, y_k)$, and so the dual to that may be represented by $dx_i|_{\bar{y}} \equiv d_{\bar{y}} x_i$. We may build up the dual basis at $\bar{y}$ using wedge products as,

$$d_{\bar{y}} x_{i_1} \wedge \ldots \wedge d_{\bar{y}} x_{i_k} \text{ for } i_1 < \cdots < i_k$$

Thus, a general k-form α in $U \subset X$ would be (in local coordinates) at point $\bar{y} \in U$:

$$\omega^k{}_{\bar{y}} = \sum_{i_1 < \cdots < i_k} a_{i_1 \cdots i_k}(\bar{y}) d_{\bar{y}} x_{i_1} \wedge \cdots \wedge d_{\bar{y}} x_{i_k} \qquad \text{... 1-1}$$

MOTIVATION: To be able to use k-forms and do calculus with it on manifold we would atleast require differentiable assignments at every point x. Thus, we are motivated to define *differential k-forms*.

DEFINITION 1.3: A **differential k-form** (or sometimes just a k-form) α on a manifold X is a smooth assignment of a k-form $\alpha_x \in \bigwedge^k T_x^* x$ for all $x \in X$. We denote the vector space of differential k-forms on X by $\Omega^k(X)$.

---

[1] A **representation** is a very general relationship that expresses similarities (or equivalences) between mathematical objects or structures.

[2] Tangent space is a vector space.

ILLUSTRATION -
As we have already illustrated before (2-1) a general k-form α at $\bar{y} \in U$ would be (in local coordinates):

$$\omega^k{}_{\bar{y}} = \sum_{i_1 < \cdots < i_k} a_{i_1 \cdots i_k}(\bar{y}) d_{\bar{y}} x_{i_1} \wedge \cdots \wedge d_{\bar{y}} x_{i_k}$$

Just like the idea of tangent fields, the k-form ω^k is smooth precisely when $a_{i_1 \cdots i_k}(\bar{y})$ is a smooth function of $\bar{y} \in X$, in which case we simply denote it as,

$$\alpha = \sum_{i_1 < \cdots < i_k} a_{i_1 \cdots i_k} dx_{i_1} \wedge \cdots \wedge dx_{i_k}$$

- Note that all k-forms on $U \cong \mathbb{R}^n$ takes precisely this form. (Theorem 4.1)

1.1. TRANSFORMATION RULE

MOTIVATION: We shall now see that differential forms are indeed precisely defined under coordinate transformations. Unlike tangent vectors which are geometric objects, differential forms may be better thought of as maps that "eat" tangent vectors. Hence, just like the nature thing to do with tangent vectors are *push forwards*, the natural thing to do with maps, for that matter differential forms, are *pullbacks*.

We shall now define the pullback of differential forms:

DEFINITION 1.4: Suppose $\alpha \in \Omega^k(Y)$ and suppose $F: X \to Y$ is a smooth map between overlapping charts of manifold; then we define the **pullback** $F^*(\alpha) \in \Omega^k(Y)$ of α to be the k-form on X that act on vectors via:

$$F^*\alpha_x(\xi_1, \dots, \xi_k) = \alpha_{F(x)}(D_x F(\xi_1), \dots, D_x F(\xi_k))$$

ILLUSTRATION: To illustrate this in local coordinates:
Suppose we have a map, F, between two local coordinates, $F: (U, \phi) \to (U, \psi): x_i \mapsto y_i \equiv F_i\,(x_1, \dots, x_n)$ and a n-differential form –

$$\alpha_y = a(y) \mathrm{d}y_1 \wedge \cdots \wedge \mathrm{d}y_n$$

Then $F^*(\alpha)$ is given by-

$$\begin{aligned} F^*(\alpha)_x(v_1, \dots, v_n) &= a(F(x)) \mathrm{d}y_1 \wedge \cdots \wedge \mathrm{d}y_n \big(D_x F(v_1), \dots, D_x F(v_n)\big) \\ &= (a \circ F)(x) \det\left(\mathrm{d}y_i\left(D_x F(v_j)\right)\right) = (a \circ F)(x) \det\left(D_x F_i(v_j)\right) \\ &= (a \circ F)(x) \det\left(\sum_{k=1}^{n} \frac{\partial F^i}{\partial x_k} v_j^k\right) = (a \circ F)(x) \det\left(\frac{\partial F^i}{\partial x_k}\right) \det\left(v_j^k\right) \end{aligned}$$

by the multiplicative properties of determinants, we have the *transformation rule for differential forms*:

$$F^*(\alpha)_x = (a \circ F)(x) \det(D_x F) \mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_n \qquad \text{... 1-2}$$

This is way too familiar for us since this is exactly the change of variables rule for multivariable integration, using the *Jacobian determinant*.

REMARK:
- In fact, Eq 2-2 is a more general version of change of variable rule, as you might recollect that the in the original formula for Reimann Integral, the change of volume factor is actually $|\det(D_x F)|$, which is because we implicitly assume that the manifold to be *orientable*.

DEFINITION 1.5: An **orientation** of a manifold X is a choice of charts (U_i, ϕ_i) that cover X and so that the transition maps $\phi_j \circ \phi_i^{-1}$ are orientation-preserving. We say X is orientable if such an orientation exists. Further, the transition maps are orientation-preserving if $\det(D_x F) > 0$.

- Thus, the two formulas coincide for an orientation-preserving map F.

MOTIVATION: This brings us to a very important application of differential forms, namely for integration that takes advantage of this invariant property.

THEOREM 1.2: Given an n-manifold X with an orientation, and α an n-form on X with *compact support*, there is a **well-defined integral**

$$\int_X \alpha \in \mathbb{R}$$

that is linear in α and satisfies the invariance property that for any diffeomorphism $F: Y \to X$

$$\int_Y F^*(\alpha) = \int_X \alpha$$

1.2. APPLICATION IN PHYSICS

- The electromagnetic field strength F is a differential 2-form built from the electric and magnetic fields, which in turn are 1-forms.
- In classical mechanics, if Q is a smooth manifold describing the possible system configurations, then the phase space is T^*Q, which is in fact a symplectic space.

1.3. DERIVATIVES OF DIFFERENTIAL FORMS

DEFINITION 1.6: The exterior derivative of a k-form α on X is a $(k+1)$ −form $d\alpha \in \Omega^{k+1}(X)$, defined as follows: if α can be written in a local coordinate chart as,

$$\alpha = \sum_{i_1<\cdots<i_k} a_{i_1\ldots i_k}(x)\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k} \ \text{ then } \ \mathrm{d}\alpha = \sum_{i_1<\cdots<i_k} \sum_{j=1}^{n} \frac{\partial a_{i_1\ldots i_k}(x)}{\partial x_j} \mathrm{d}x_j \wedge \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k}$$

REMARK:

Derivative of exterior product is in fact the generalization of differential vector operators as seen below:

Let $X \equiv \mathbb{R}^3$ with the usual basis e_1, e_2, e_3, and recollect that the k-forms on $\mathbb{R}^3$ belong to the dual space of $(\mathbb{R}^3)^*$,

- Gradient operator:
  For a smooth function f on $\mathbb{R}^3$,
  $$\mathrm{d}f = \frac{\partial f}{\partial x_i}\mathrm{d}x_i 1 \equiv \nabla f \cdot \mathrm{d}x \in \Lambda^1(\mathbb{R}^3)^*$$
- Curl operator:
  For a 1-form α on $\mathbb{R}^3$,
  $$\mathrm{d}\alpha \equiv d(a_1\,\mathrm{d}x + a_2\,\mathrm{d}y + a_3\,\mathrm{d}z) = \left(\frac{\partial a_y}{\partial x} - \frac{\partial a_x}{\partial y}\right)\mathrm{d}x \wedge \mathrm{d}y + \cdots \in \Lambda^2(\mathbb{R}^3)^*$$
- Divergence Operator:
  For a 2-form β on $\mathbb{R}^3$,
  $$\mathrm{d}\beta \equiv d\big(a_{xy}\,\mathrm{d}x \wedge \mathrm{d}y + a_{xz}\,\mathrm{d}x \wedge \mathrm{d}z + a_{yz}\,\mathrm{d}y \wedge \mathrm{d}z\big) = \left(\frac{\partial a_{xy}}{\partial z} + \frac{\partial a_{xz}}{\partial y} + \frac{\partial a_{yz}}{\partial x}\right)\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z \in \Lambda^2(\mathbb{R}^3)^*$$

PROPOSITIONS: Let α be a k-form and β be an ℓ-form and let $F: X \to Y$ be a smooth transition map on some manifold M. Then we have,

- Pullback: $F^*(\mathrm{d}\alpha) = \mathrm{d}F^*(\alpha)$ and $F^*(\alpha \wedge \beta) = F^*(\alpha) \wedge F^*(\beta)$
  This follows because the exterior products are coordinate independent.
- Exterior Derivative of wedge product: $\mathrm{d}(\alpha \wedge \beta) = \mathrm{d}\alpha \wedge \beta + (-1)^k \alpha \wedge \mathrm{d}\beta$
- Fundamental Cohomology result:
  $$\mathrm{d}(\mathrm{d}\alpha) = 0 \qquad \text{... 1-3}$$
  It follows that div ∘ curl $= 0$ and curl ∘ grad $= 0$. The proof for this makes use of the fact that order of mixed partial derivatives for smooth functions can be interchanged (*Clairaut's theorem*).

2. Lie Algebra

DEFINITION 2.1: A **Lie group** is a manifold G along with smooth maps $m: G \times G \to G$ called multiplication and $i: G \to G$ called inversion, as well as a distinguished point $e \in G$, that together satisfy the axioms to be a group: namely, m is associative, e is a left and right identity, and every element has a (two-sided) inverse given by i.

Further, just as for usual groups we may define:

DEFINITION 2.2: A (left) **action of a Lie group** G on a manifold X is a smooth map $a: G \times X \to X$, written $(g, x) \mapsto g \cdot x$, that defines an action of the group G on the set X, that is, e acts as the identity map, and $(gh) \cdot x = g \cdot (h \cdot x)$. We say that G gives a **smooth symmetry** of X.

To illustrate the close association of the action of Lie group and symmetric consider the following:

- Notice that $\mathbb{R}^n$ is a Lie group under the addition map, $m: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n: (x, y) \mapsto x + y$ and the map $i: \mathbb{R}^n \to \mathbb{R}^n: x \mapsto -x$ with $e = 0$ (the smoothness of these maps may be easily verified). Then the action of Lie Group $G \equiv \mathbb{R}^n$ acts on $\mathbb{R}^n$ by *translation*. This captures the notion that $\mathbb{R}^n$ has a *translation-symmetry*.
- Very similarly, the group of rotations of $\mathbb{R}^n$, $\mathrm{SO}(n) \subset \mathrm{GL}_n(\mathbb{R})$ also forms a Lie group and the action of this group on $\mathbb{R}^n$ is essentially *rotation*. This captures the notion that $\mathbb{R}^n$ has a *rotational-symmetry*.

DEFINITION 2.3: The Lie algebra $\mathfrak{g}$ of a Lie group G is the tangent space $T_e G$ at the identity of G.

REMARK –

- Given any $v \in T_e G$, we can form a vector field X_v anywhere on G by using the multiplication map to transport the vector around. Hence without loss of generality we may take tangent vector at any point in G and not necessarily identity element.

MOTIVATION: It may not be evident how such a construction forms an algebra. Following we shall illustrate how such a system already has an algebraic structure (through *Lie bracket*).

If we let $L_g: G \to G$ denote the map given by left multiplication by $g \in G$, then its derivative defines a map $\mathcal{L}_e L_g: T_e G \to T_g G$. More generally, for $e \equiv x_o$ we shall consider the following define for **Lie derivative**:

2.1. LIE DERIVATIVES

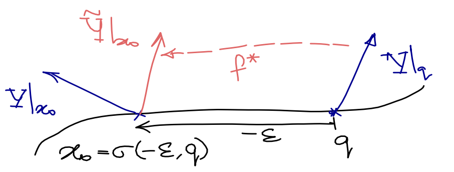
MOTIVATION: Lie derivatives are one way of comparing two things (geometric objects like tangent vectors or differential forms) at two different spaces by use of pullback.

- Lie derivative of two Vectors –

$$\mathcal{L}_X Y(x_0) = \lim_{\varepsilon \to 0} \left[\frac{\tilde{Y}\big|_{x_0} - Y\big|_{x_0}}{\varepsilon} \right] := [X, Y]$$

- Lie derivative of a scalar –

$$\mathcal{L}_X f(x_0) = X(f(x_0))$$



This is a pictorial representation of pulling back $Y|_q$ to point x_0 along the flow f.

2.2. PROPERTIES OF LIE BRACKET

- Antisymmetry: $\mathcal{L}_X Y = [X, Y] = -[Y, X] = -\mathcal{L}_Y X$
- Bilinearity: $[X, Y + Z] = [X, Y] + [X, Z]$ and $[X + Z, Y] = [X, Y] + [Z, Y]$
- Leibnitz rule: $\mathcal{L}_X YZ = Y(\mathcal{L}_X Z) + (\mathcal{L}_X Y)Z$
- Jacobi Identity: $\big[X, [Y, Z]\big] + \big[Y, [Z, X]\big] + \big[Z, [X, Y]\big] = 0$

$$[\mathcal{L}_X, \mathcal{L}_Y]Z = \big[X, [Y, Z]\big] - \big[Y, [X, Z]\big] = \big[X, [Y, Z]\big] + \big[Y, [Z, X]\big] = +\big[Z, [X, Y]\big] = \mathcal{L}_{[X,Y]} Z$$

2.3. CARTAN'S MAGIC FORMULA

THEOREM 2.1: (Cartan's Magic Formula) The Lie derivative of a differential form is given by

$$\mathcal{L}_v \alpha = \mathrm{d}\iota_v \alpha + \iota_v \mathrm{d}\alpha \qquad \text{... 2-1}$$

where ι_v is the interior product with a vector field defined above.

PROOF: Although this may be easily verified in local coordinates (as shown below), actual proof is more subtle and interested reader may refer any standard differential manifold text, say (Loomis & Sternberg, 2014).

VERIFICATION: Let $\omega \in \Omega^1(m) \coloneqq \omega_\mu \, \mathrm{d}x^\mu$ and $\mathrm{X} = X^\mu \, {}^{\partial}/_{\partial x^\mu}$, then,

$$\mathrm{d}(\iota_X \omega) = \partial_\nu \left(w_\mu X^\mu \right) \mathrm{d}X^\nu$$

And,

$$\begin{aligned} \iota_X d\omega &= \iota_X \, {}^{1}/_{2} \left[\partial_\nu \omega_\mu \, dx^\nu \wedge dx^\mu + \partial_\mu \omega_\nu \, dx^\mu \wedge dx^\nu \right] \\ &= {}^{1}/_{2} \left[X^\mu \, \partial_\nu \omega_\mu - \mathrm{X}^\mu \, \partial_\mu \omega^\nu \right] \end{aligned}$$

Hence,

$$\mathrm{d}(\iota_X \omega) + \iota_X d\omega = \mathcal{L}_X \omega \quad \text{or} \quad \boldsymbol{d\iota_X + \iota_X d = \mathcal{L}_X}$$

REMARKS –

- This is essentially a relation that relates Lie derivative, exterior derivative, and interior product.

2.4. POISSON BRACKET

DEFINITION 2.4: The Poisson bracket of two functions $f, g \in C^\infty(M; \mathbb{R})$ is

$$\{f, g\} \coloneqq \omega\left(X_f \, , X_g\right) \xrightarrow{\text{coordinates}} \sum_{i=1}^{N} \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i}$$

REMARK –

- A **Poisson algebra** is an associative algebra together with a Lie bracket that also satisfies Leibniz's law.

Symplectic Geometry

> "Everything is a Lagrangian submanifold."
>
> Alan Weinstein

3. Symplectic Space

Let V be a finite dimensional vector space over R.

DEFINITION 3.1: An *antisymmetric, nondegenerate bilinear form* (non-degenerate 2-form) on V is called a **symplectic form.**

To state the properties explicitly, a symplectic bilinear form is a mapping $\omega: V \times V \to R$ that is

- Bilinear: linear in each argument separately,
- Alternating: $\omega(v, v) = 0$ holds for all $v \in V$, and
- Nondegenerate: $\omega(u, v) = 0$ for all $v \in V$ implies that u is zero. ($\mathbf{\dim V = 2n}$)

DEFINITION 3.2: A vector space possessing a given symplectic form is called a **symplectic vector space** (V, Ω), or is said to have a **symplectic structure**.

DEFINITION 3.3: Space possessing a given symplectic form is called a **symplectic vector space** (V, Ω), or is said to have a **symplectic structure**.

DEFINITION 3.4: A linear map $A: \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ that gives a symplectomorphism of the canonical symplectic form J on $\mathbb{R}^{2n}$, that is:

$$A^T J A = J$$

is called a **symplectic transformation.** We denote the group of such transformations by $\mathrm{Sp}(2n)$ and call it the **symplectic group.**

REMARK –

- A Lie group G acting on X a symplectic manifold acts symplectically if every map $\Phi_g: X \to X$ given by $x \mapsto g \cdot x$ is a symplectomorphism.

THEOREM 3.1: (Multilinear Algebra): Let Ω be a **skew-symmetric bilinear map** on V. Then there is a basis $u_1, \dots, u_k, \dots, e_1, \dots, e_n, f_1, \dots, f_n$ of V such that

$$\begin{aligned} \Omega(u_i, v) &= 0, && \text{for all } i \text{ and all } v \in V, \\ \Omega(e_i, e_j) &= 0 = \Omega(f_i, f_j), && \text{for all } i, j, \text{ and} \\ \Omega(e_i, f_j) &= \delta_{ij}, && \text{for all } i, j. \end{aligned}$$

REMARK –

- This basis decomposition is not unique.
- $k + 2n = \dim V$; n is invariant of (V, Ω), $2n$ is **rank of Ω**

COROLLARY: From the requirement of nondegeneracy for Symplectic space, it follows that every symplectic space is even-dimensional.

A symplectic vector space (V, Ω) has a basis $(e_1, \dots, e_n, f_1, \dots, f_n)$ satisfying $\Omega(e_i, f_j) = \delta_{ij}$ and $\Omega(e_i, e_j) = 0 = \Omega(f_i, f_j)$ Such a basis is called a **symplectic basis** of (V, Ω).

$$\text{Hence, } \Omega(\mathrm{u}, \mathrm{v}) = [-\mathrm{u}-] \begin{bmatrix} 0 & \mathrm{I_n} \\ -\mathrm{I_n} & 0 \end{bmatrix} \begin{bmatrix} | \\ \mathrm{v} \\ | \end{bmatrix} 1 \equiv u^T J v \quad \text{where } \mathrm{u}, \mathrm{v} \in \mathrm{V} \times \mathrm{V} \qquad \dots 3\text{-}1$$

4. Symplectic Manifold

DEFINITION 4.1: A **symplectomorphism** or **canonical transformation** ϕ between symplectic vector spaces (V, Ω) and (V', Ω') is a linear isomorphism $\phi : V \xrightarrow{\sim} V$ such that $\phi^*\Omega' = \Omega$.

REMARKS –

- In the context of manifolds, various authors tend to include diffeomorphism in the definition.
- By definition of pullback, $(\varphi^*\Omega')(u, v)\Omega'\big(\phi(u), \phi(v)\big)$
- If a symplectomorphism exists, (V, Ω) and (V', Ω') are said to be symplectomorphism.
- Symplectomorphism form an Equivalence Relation.

DEFINITION 4.2: A manifold M is symplectic if it contains an addition structure of *closed de-Rham 2-form,* ω on M and denote a **Symplectic Manifold** as the pair (M, ω).

DEFINITION 4.3: A **de-Rham 2-form** / exterior 2-form is a map $\omega_p : T_P M \times T_p M \to \mathbb{R}$ such that for each $p \in M$, ω_p is *skew-symmetric bilinear* on the **tangent space to M at p**, and ω_p varies smoothly in p.

Example – For $M = \mathbb{R}^{2n}$ with linear coordinates $x_1, \dots, x_n, \dots, y_1, \dots, y_n$ called **canonical basis** we have the symplectic form,

$$\omega = \sum_{i=1}^{n} dx_i \wedge dy_i$$

called the **Canonical Symplectic Form.**

MOTIVATION: Symplectic manifold is a peculiar space that allows us to do Hamiltonian Mechanics. More specifically, it turns out that the Cotangent Bundle $(\mathbb{R}^{2n})$ of Position space $(\mathbb{R}^n)$ is in fact the Phase Space with an inherent 2-form, the analogue of Poisson Bracket (Theorem 4.1).

4.1. DARBOUX'S THEOREM

THEOREM 4.1: (Darboux, 1882) Suppose (X, ω) is a 2n-dimensional symplectic manifold; then for every $x \in X$ there is a coordinate chart (U, ϕ) with $x \in U$ such that ϕ gives a symplectomorphism between (U, ω) and an open subset of $\mathbb{R}^{2n} \equiv T^*\mathbb{R}^n$ called **Darboux Chart**, with the canonical symplectic form,

$$\omega = \sum_{i=1}^{n} \mathrm{d}x_i \wedge \mathrm{d}y_i$$

PROOF: The proof is very involved and beyond the scope of this manuscript. Interested readers may refer to any Symplectic Geometry text, including (Cannas da Silva, 2008).

REMARKS –

- This essentially means that there is no local symplectic geometry: all symplectic manifolds locally look the same! This is in stark contrast to Riemannian geometry, where there is a whole suite of different local invariants given by various kinds of curvatures.

4.2. COTANGENT BUNDLE AS A SYMPLECTIC MANIFOLD

Let X be any n-dimensional manifold with coordinate charts $(U, x_1, \dots, x_n)$ for $x \in U$ with $x_i : U \to R$ and $M = T^*X$ its cotangent bundle. It follows that the differentials $(dx_1)_p, \dots, (dx_n)_p$ form a basis of the cotangent space at p, T_p^*X and so the Transition functions between charts are contravariant transformations.

Hence, we have the induced map,

$$T^*U \to \mathbb{R}^{2n} : (x, \zeta) \mapsto (x_1, \dots, x_n, \zeta_1, \dots, \zeta_n)$$

where $x_1, \dots, x_n, \zeta_1, \dots, \zeta_n$ are the cotangent coordinates.

THEOREM 4.2: Every cotangent bundle is a symplectic manifold with the canonical symplectic form. Using the coordinate chart, $T^*U \to \mathbb{R}^{2n}: (x,\zeta) \mapsto (x_1, \dots, x_n, \zeta_1, \dots, \zeta_n)$ define the tautological 1-form α on T^*U by

$$\alpha = \sum_{i=1}^{n} \zeta_i \, dx_i$$

Further, define the canonical 2-form as,

$$\omega = -d\alpha = \sum_{i=1}^{n} dx_i \wedge d\zeta_i \qquad \dots\ 4\text{-}1$$

PROOF: The existence of such coordinates are guaranteed since Cotangent Bundle is a manifold. In coordinates, it is intuitively evident that the basic vectors are very similar to the Darboux Coordinates. Now, the choice of coordinates are independent since any choice would look similar to the formal coordinate chart and so the various forms should follow.

REMARKS –

- Since α is coordinate independent, ω too is!
- $d\omega = -\,d^2\alpha = 0$, i.e., closed hence a symplectic form.
- Note that the underlying Symplectic form need not be unique.

4.3. PROPERTY OF TAUTOLOGICAL 1-FORM

THEOREM 4.3: For the tautological 1-form on T^*X and a section $s_\mu: X \to T^*X: x \mapsto (x, \mu_x)$. We have,

$$s_\mu^* \alpha = \mu \qquad \dots\ 4\text{-}2$$

PROOF: Let $V \in T_x X$ be arbitrary, then (write $\mu_x \in T_x^* X$)

$$\left(s_\mu^* \alpha\right)_x (V) = \alpha_{(x,\mu_x)}\left(s_{\mu *} V\right) = \pi^*_{(x,\mu_x)} \mu_x \left(s_{\mu *} V\right)$$
$$= \mu_x\left(\pi_{(x,\mu_x)*} s_{\mu *} V\right) = \mu_x\left((\pi \circ s_\mu)_* V\right) = \mu_x(V)$$

$$T^*X \xrightarrow{\pi} X \qquad TT^*X \xleftarrow{\pi_*} TX \qquad T^*T^*X \xleftarrow{\pi^*} T^*X$$

5. Complex Vector Space

MOTIVATION: It turns out that a symplectic manifold naturally contains an additional structure of almost complex Manifold. We shall capture this property briefly below.

DEFINITION 5.1: A complex structure on a vector space, V is a linear map $J: V \to V$ such that $J^2 = -I$ and (V, J) is a **complex vector space**.

Recollect the Matrix notation of Symplectic form from Equation 3-1.

$$\omega(v, \omega) = -v^T J w \quad \text{where} \quad J = \begin{pmatrix} 0 & -\mathrm{I} \\ \mathrm{I} & 0 \end{pmatrix}$$

It may be verified that $J^2 = -I$.

DEFINITION 5.2: For a symplectic vector space (V, Ω), a complex structure J on V is **compactible** (Ω-compactible) if,

$$G_J(u, v) \coloneqq \Omega(u, Jv) \ \ \forall u, v \in V \text{ is a positive inner product on } V$$

Further, if J is a symplectic transformation i.e., $\Omega(Ju, Jv) = \Omega(u, v)$ is called a **Kählerian vector space**.

THEOREM 5.1: For a symplectic vector space (V, Ω), there exist a compatible complex structure J on V.

5.1. ALMOST COMPLEX STRUCTURE

DEFINITION 5.3: An almost complex structure on a manifold M is a smooth field of complex structures on TM,

$$X \in M \mapsto J_x: T_X M \to T_X M \ \text{ linear } \ \& \ \ J^2 = -I$$

(M, J) is an **almost complex manifold**.

DEFINITION 5.4: Let (M, ω) be a symplectic manifold. An almost complex structure J on M is called compatible (with ω or ω-compatible) if the assignment,

$$x \mapsto g_x: T_xM \times T_xM \to \mathbb{R}; \quad g_x(u, v) := \omega_x(u, J_x v)$$

is a Riemannian metric on M.

THEOREM 5.2: Let (M, ω) be a symplectic manifold, and g a Riemannian metric on M. Then there exists a canonical almost complex structure J on M which is compatible.

REMARKS –

- An almost complex manifold is NOT the same as a Complex Manifold, complex manifold has a global property of Complex Integrability on the manifold (readers familiar with Complex Analysis may recollect that complex integration has to do with the Holomorphicity of the map).

6. Lagrangian Submanifolds

MOTIVATION: A very important subspace of symplectic manifolds is the submanifold where the symplectic form vanishes. Under quantization of a symplectic manifold, they correspond to quantum states and as a result display many distinctly quantum properties. Hence these provide a link between Classical and quantum world. Among other applications include their vigorous use in optimization problems.

Let (M, ω) be a 2n-dimensional symplectic manifold.

DEFINITION 6.1: For the inclusion map, $i: L \hookrightarrow M$, L is Lagrangian Submanifold if and only if $i^*\omega = 0$ and $\dim L = 1/2 \dim M$.

Equivalently,

A submanifold L of M is a Lagrangian Submanifold if, at each $p \in L$, T_pL is a Lagrangian subspace of T_pM, i.e., $\omega_p\big|_{T_pL} = 0$ and $\dim T_pL = 1/2 \dim T_p M$ (maximally coisotropic).

REMARKS

- Since Lagrangian submanifold is essentially an embedded submanifold[3], in R^{2n} with Darboux's coordinates $(q_1, \dots, q_n, p_1, \dots, p_n)$, as seen later, the spaces $\mathbb{R}^n \times \{p\}$ and $\{q\} \times \mathbb{R}^n$ (p, q are fixed respectively) are all Lagrangian. In quantum mechanics, you can think of these as the eigenstates of the position and momentum operators.
- Lagrangian has nothing to do with the Lagrangian function of Classical Mechanics, but rather with the fact that ω was originally called the Lagrange bracket, and historically the space where the bracket vanished came to be subsequently called Lagrange Subspaces.

6.1. LAGRANGIAN SUBSPACE ON COTANGENT BUNDLES

- For any smooth manifold Q, the *cotangent fibres* $T_q^*Q \subseteq T^*Q$ are Lagrangian submanifolds of the cotangent bundle with the canonical symplectic form.

THEOREM 6.1: The graph of α is a Lagrangian submanifold of T^*Q with the standard symplectic structure exactly when $d\alpha = 0$. Moreover, the graph is an exact Lagrangian precisely when $\alpha = df$ for $f: Q \to \mathbb{R}$ a smooth function, where f is called the **generating function**. In particular, when $\alpha = 0$ we call this Lagrangian the zero section.

PROOF: Using Theorem 4.2 and noting that:

$$s_\alpha^*\omega = -ds_\alpha^*\lambda = -d\alpha = 0 \text{ in which case } \alpha := df \text{ for some } f: Q \to \mathbb{R}$$

ILLSUTRATION: To get an intuitive idea, we move to the local coordinates:
Using the coordinate chart, $T^*U \to \mathbb{R}^{2n}: (x, \zeta) \mapsto (x_1, \dots, x_n, \zeta_1, \dots, \zeta_n)$ for $U \subseteq X$we have the canonical 2-form,

---

[3] An **(embedded) submanifold** L of smooth manifold X is a subset $L \subseteq X$ such that around every point $p \in L$ there exists some smooth chart (U, ϕ) in which $L \cap U$ is given by the hyperplane $x_1 = x_2 = \dots = x_n = 0$ inside $\mathbb{R}^n$.

$$\omega = -d\alpha = \sum_{i=1}^{n} dx_i \wedge d\zeta_i$$

The zero section of T^*X is,

$$X_0 \coloneqq \{(x,\zeta) \in T^*X | \zeta = 0 \text{ in } T^*X\}$$

$\alpha = \sum_{i=1}^{n} \zeta_i \, dx_i$ vanishes on $X_0 \cap T^*U$ since $i^*\alpha = 0$ for the inclusion $i: X_0 \hookrightarrow T^*X$, and so does ω. Hence, the **zero section is a Lagrangian submanifold**.

REMARKS –

- There is a one-to-one correspondence between the set of Lagrangian submanifolds of T^*X and the set of closed 1-forms on X.

6.2. LAGRANGIAN SUBSPACE AND SYMPLECTOMORPHISM

MOTIVATION: Theorem 7.1 has a very important consequence as seen below.

Let (X_1, ω_1) and (X_2, ω_2) be two 2n-dimensional symplectic manifolds with a diffeomorphism $\varphi: X_1 \xrightarrow{\sim} X_2$. Define the **twisted product space** $(X_1 \times \overline{X_2}, \widetilde{\omega})$, such that

$$\widetilde{\omega} \coloneqq \pi_1^*\omega_1 - \pi_2^*\omega_2 \text{ and } X_1 \times \overline{X_2} \in \Gamma_\varphi \coloneqq \text{Graph } \varphi = \{(p, \varphi(p)) | p \in X_1\}$$

THEOREM 6.2: A diffeomorphism φ is a symplectomorphism if and only if the graph, Γ_φ of the diffeomorphism $\varphi: X_1 \to X_2$ is a *Lagrangian Submanifold*[4] of $(X_1 \times \overline{X_2}, \widetilde{\omega})$.

PROOF: If $\Gamma_\varphi: X_1 \to X_1 \times X_2$ represents the graph of φ, then,

$$\Gamma_\varphi{}^*((\omega_1, -\omega_2)) = \omega_1 - \varphi^*\omega_2$$

and this is zero if and only if φ is a symplectomorphism.

[4] A subspace that is also a manifold in its own rights.

Classical Mechanics

> "If I have seen further than others, it is by standing upon the shoulders of giants."
>
> Sir Isaac Newton

As was mentioned earlier, Symplectic Manifold is the space of Hamiltonian Mechanics. We shall now see how Hamiltonian Dynamics arises naturally out of Symplectic Differentiable Manifolds.

7. Hamiltonian vector spaces

Let M be a differentiable manifold, and $\rho: M \times \mathbb{R} \to M$ a map, where we set $\rho_t(p) \coloneqq \rho(p,t)$.

THEOREM 7.1: (Picard) In the neighborhood of any point p and for sufficiently small time t, there is a one-parameter family of local diffeomorphisms ρ_t called **isotopy** satisfying,

$$\frac{\partial \rho_t}{dt} = v_t \circ \rho_t \quad \text{and} \quad \rho_0 = id_M$$

- One parameter family of diffeomorphism: $\sigma\big(t, \sigma(s+x)\big)1 = \sigma(t+s, x)$

DEFINITION 7.2: When $v_t = v$ is independent of t, the associated isotopy is called the *exponential map* or the **flow** and denoted $\exp tv$.

Let (M, ω) be a symplectic manifold and let $H: M \to R$ be a smooth function. Its differential dH is a 1-form. By nondegeneracy, there is a unique vector field X_H on M such that $\iota_{X_H}\omega = dH$.

DEFINITION 7.3: A vector field X_H, that satisfies $\iota_{X_H}\omega = dH$ (where ι_{X_H} is *inner product*, i.e., $\iota_{X_H}\omega \equiv \omega(X_H, \_)$) is called the **Hamiltonian Vector Field** corresponding to the integral curve H called *Hamiltonian Function*.

- X is Hamiltonian $\Leftrightarrow \iota_X\omega$ is exact i.e., $d(\iota_X\omega) = d^2H = 0$, from Eqn. 1-3

REMARK on Sign Conventions –

- Many authors disagree on this sign convention and instead write $\iota_{X_H}\omega = -dH$. For our sign convention we shall be following (McDuff & Salamon, 2017).

Consider Euclidean space $\mathbb{R}^{2n}$ with coordinates $(q_1, \dots, q_n, p_1, \dots, p_n)$ and $\omega_0 = \Sigma dq_j \wedge dp_j$. The curve $\rho_t = (q(t), p(t))$ is an integral curve for X_H exactly if

$$\frac{dq_i}{dt}(t) = \frac{\partial H}{\partial p_i}\,;\, \frac{dp_i}{dt}(t) = -\frac{\partial H}{\partial q_i} \qquad \dots 7\text{-}1$$

Indeed for, $X_H = \sum_{i=1}^{n}\left(\frac{\partial H}{\partial p_i}\frac{\partial}{\partial q_i} - \frac{\partial H}{\partial q_i}\frac{\partial}{\partial p_i}\right)$ we have,

$$\iota_{X_H}\omega = \sum_{j=1}^{n} \iota_{X_H}(dq_j \wedge dp_j) = \sum_{j=1}^{n}\left[(\iota_{X_H}dq_j) \wedge dp_j - dq_j \wedge (\iota_{X_H}dp_j)\right]$$

$$= \sum_{j=1}^{n}\left(\frac{\partial H}{\partial p_j}dp_j + \frac{\partial H}{\partial q_j}dq_j\right) = dH$$

A Hamiltonian system is a triple (M, ω, H), where (M, ω) is a symplectic manifold and $H \in C^\infty(M, \mathbb{R})$ is a function, called the Hamiltonian Function.

DEFINITION 7.4: Suppose Q is an n-dimensional smooth manifold; let $x = T^*Q$ be the cotangent bundle and $\pi: T^*Q \to Q$ be the projection. We define the **tautological 1-form** λ on X to be:

$$\lambda_{(q,\alpha)}(v) = \alpha\left(D_{(q,\alpha)}\pi(v)\right) \quad \text{where} \quad (q,\alpha) \in Q \times T^*Q \quad \text{and} \quad v \in T_{(q,\alpha)}X$$

DEFINITION 7.5: On a cotangent bundle $X = T^*Q$ the **canonical 2-form** ω is defined to be $-d\lambda$.

DEFINITION 7.6: In general, we say a symplectic manifold is **exact** ($d\omega = 0$) if the symplectic form ω is the exterior derivative of some 1-form λ, called the Liouville form.

- For our purpose this is essentially true for Tautological 1-form (Definition 8.4).

ILLUSTRATION - We shall now illustrate the above definitions in a local chart:
Let $Q = \mathbb{R}^n$ with coordinates $(q_1, \dots, q_n)$ and $T^*Q \cong \mathbb{R}^{2n}$ with $(q_1, \dots, q_n, p_1, \dots, p_n)$, while the projection map is $\pi: T^*Q \to Q: (q_1, \dots, q_n, p_1, \dots, p_n) \mapsto (q_1, \dots, q_n)$. Then, by definition, we have $D_\pi: \mathbb{R}^{2n} \to \mathbb{R}^n$ given simply by projection to the first n coordinates, so that

$$\lambda_{(q,p)}(v_1, \dots, v_n, v_{n+1}, \dots, v_{2n}) = \sum_{i=1}^{n} p_i v_i \quad \text{where} \quad \lambda = \sum_{i=1}^{n} p_i \mathrm{d}q_i$$

Further, the canonical 2-form would be,

$$\omega \coloneqq -\mathrm{d}\lambda = -\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial p_i}{\partial p_j} \mathrm{d}p_j \wedge \mathrm{d}q_i = \sum_{i=1}^{n} \mathrm{d}q_i \wedge \mathrm{d}p_i$$

- We call this the **canonical symplectic form** on $\mathbb{R}^{2n}$. The reason is apparent from **Darboux's theorem.** It may be easily verified that $d\omega = 0$.

Having fixed the basis say $\{e_i\} \equiv (q_1, \dots, q_n, p_1, \dots, p_n)$, this may be rewritten in Matrix notation as

$$\omega(e_i, e_j) = \sum_{k=1}^{n} \mathrm{d}q_k \wedge \mathrm{d}p_k(e_i, e_j) = \sum_{k=1}^{n} \det\begin{pmatrix} q_k(e_i) & q_k(e_j) \\ p_k(e_i) & p_k(e_j) \end{pmatrix}$$

$$J = \begin{pmatrix} 0 & -\mathrm{I} \\ \mathrm{I} & 0 \end{pmatrix} \quad \Rightarrow \quad \omega(v, \omega) = -v^T J w$$

- J is invertible matrix, since we are requiring ω to be non-degenerate.

Following Theorem 4.1, we may now calculate the Hamiltonian flow associated to a smooth function H on the cotangent bundle T^*Q in the above local coordinates:

$$\mathrm{d}H(v) = (\nabla H)^T v = -X_H^T J v \quad \Rightarrow \quad -X_H^T J = (\nabla H)^T$$

Hence taking the transpose we have,

$$X_H = -J \nabla H$$

This was precisely the formula we had for the **symplectic gradient of the function H**! Therefore, the flow of the Hamiltonian vector field of H on the cotangent bundle gives exactly the Hamiltonian dynamics associated to the function H!

DEFINITION 7.7: The Poisson bracket of two functions $H, f: X \to \mathbb{R}$ on a symplectic manifold is another smooth function $\{H, f\}$ on X defined by $-\{H, f\} = \omega(X_H, X_f)$

THEOREM 7.2: We have $\{f, H\} = 0$ if and only if f is constant along integral curves of X_H, in which case f is conserved under the Hamiltonian flow and is called an **integral of motion** (or a first integral or a constant of motion).

PROOF – Let ρ_t be the flow of X_H. Then,

$$\frac{d}{dt}(f \circ \rho_t) = \rho_t^* \mathcal{L}_{X_H} f = \rho_t^* \iota_{X_H} df = \rho_t^* \iota_{X_H} \iota_{X_f} \omega$$
$$= \rho_t^* \omega(X_f, X_H) = \rho_t^* \{f, H\}$$

PROPOSITION: **Hamiltonian Isotopy** is a symplectomorphism, since

$$\frac{d}{dt} \rho_t^* \omega = \rho_t^* \mathcal{L}_{X_H} \omega = \rho_t^* (d \underbrace{\iota_{X_H} \omega}_{dH} + \iota_{X_H} \underbrace{d\omega}_{0}) = 0.$$

- In essence Theorem 7.2, implies that for any smooth function $f: X \to \mathbb{R}$, we have $\dot{f} = \{f, H\}$ along the trajectories of H. Thus, *Poisson brackets control how quantities evolve under Hamiltonian dynamics.*

7.1. LIE ALGEBRA IN SYMPLECTIC MANIFOLD

PROPOSITION: The collection of all smooth functions on a symplectic manifold X forms a Lie algebra with respect to the Poisson bracket: that is, the Poisson bracket satisfies:

- (Bilinearity) $\{\lambda f + \mu g, h\} = \lambda\{f, h\} + \mu\{g, h\}$, for $\mu, \lambda \in \mathbb{R}$
- (Antisymmetry) $\{f, g\} = -\{g, f\}$
- (Jacobi Identity) $\{f, \{g, h\}\} + \{h, \{f, g\}\} + \{g, \{h, f\}\} = 0$

REMARKS

- Poisson bracket behaves a lot like the commutator brackets from quantum mechanics. In fact, the analogy runs deeper,

PROPOSITION: Suppose $\mathbb{R}^{2n}$ has the canonical symplectic form $\omega = \sum_{i=1}^{n} dq_i \wedge dp_i$; then the Poisson brackets are given by:

$$\{q_i, p_j\} = \delta_{ij}$$

PROOF: In local chart, we have,

$$\{f, g\} = \omega(X_f, X_g) = -(X_f)^T J(X_g) = -(-J\nabla f)^T J(-J\nabla g) = -(\nabla f)^T J^T J^2 (\nabla g) = -(\nabla f)^T J(\nabla g)$$

Since $\nabla q_i, \nabla p_j$ are exactly the standard basis vectors we have that $\{q_i, p_j\}$ are exactly the entries in the lower left-hand Id entry of J, that is, $\{q_i, p_j\} = \delta_{ij}$.

REMARKS:

- Even in classical mechanics, position and momentum coordinates don't commute. (This should not be confused with their corresponding flows, which indeed does commute)
- This analogy motivates the subject of **geometric quantization** - to find a means of associating to any symplectic manifold a Hilbert space with observables obeying the same commutation relations (to order $\hbar$).

DEFINITION 7.8: Let X be a manifold and V a vector field on X; the **Lie derivative** $\mathcal{L}_V(\alpha)$ of a differential k-form α on X is defined to be the differential k-form given at a point $x \in X$ by:

$$(\mathcal{L}_V \alpha)_x = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} (\Phi_t^* \alpha)_x$$

where Φ_t is the time-t flow of V and the derivative with respect to t is taken inside the vector space $\wedge^k T_x^* X$. Similarly, we can define the **Lie bracket** of two vector fields V, W to be the vector field $[V, W]$ given at a point $x \in X$ by:

$$[V, W]_x = -(\mathcal{L}_V W)_x = -\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \left(D_{\Phi_t(x)} \Phi_t^{-1} W\right)_x$$

where the derivative is taken inside the vector space $T_x X$.

PROPOSITION: If X is a symplectic manifold, then the symplectic gradient gives a homomorphism of Lie algebras from the smooth functions on X to the Lie algebra of vector fields on X: in other words,

$$X_{\{f,g\}} = [X_f, X_g]$$

for any two smooth functions f, g on X.

PROOF: Since differentiation is a linear operation, we have: $\omega\left(\mathcal{L}_{X_f} X_g, \_\right) = \mathcal{L}_{X_f}\left(\omega(X_g, \_)\right)$ where $\omega(X_g, \_)$ is a 1-fom. Then it follows that:

$$\mathcal{L}_{X_f}\left(\omega(X_g, \_)\right) = \mathcal{L}_{X_f}(\mathrm{d}g) = \mathrm{d}\left(\mathcal{L}_{X_f} g\right) \Rightarrow \mathrm{d}\left(\mathcal{L}_{X_f} g\right) = \mathrm{d}\left(\mathrm{d}g(X_f)\right) = \mathrm{d}\{f, g\}$$

where we have exchanged the derivatives. But this exactly means, $X_{\{f,g\}} = [X_f, X_g]$.

REMARK:

- Had we chosen opposite signs in either Hamilton's equation or the definition of the Lie bracket, this would have been a **Lie algebra anti-homomorphism**.

THEOREM 7.3: Suppose (X, ω) is a symplectic manifold: then a diffeomorphism $F: X \to X$ is a symplectomorphism if and only if F preserves Poisson brackets, in the sense that $\{f, g\} \circ F = \{f \circ F, g \circ F\}$.

DEFINITION 7.9: For a differentiable manifold M, the **infinitesimal action** $\phi: \mathfrak{g} \to \mathrm{Vect}(M)$ is a homomorphism of Lie algebras: we call this an infinitesimal symmetry of M.

7.2. LIOUVILLE'S THEOREM

THEOREM 7.4: The flow of every Hamiltonian gives a symplectomorphism.

PROOF: Let $\Phi_t: X \to X$ be the diffeomorphism of X given by the time-t flow of the vector field X_H of the Hamiltonian $H: X \to \mathbb{R}$. We are to show,

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi_t^* \omega = 0 \quad \text{i.e.,} \quad \mathcal{L}_{X_H}\omega = 0$$

Using Cartan's Magic Formula,

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

But, $\mathrm{d}(\iota_{X_H}\omega) = d(dH) = 0$, Property of Hamiltonian field, and $\mathrm{d}\omega = 0$ as we have required ω to be closed.

- From a different point of view, this is perhaps a belated justification of why we wanted ω to be closed.

THEOREM 7.5: (Liouville's theorem) The flow of a Hamiltonian vector field preserves volumes of subsets of X.

- Since we already have a 2-form defined, taking n wedge products of that will give the volume form on a 2n-dimensional symplectic manifold.

PROOF: Given a volume form at some point, the volume form at some other point along the Hamiltonian flow may be found using the invariance of pullback along the flow lines-

$$\int_{\Phi_t(U)} \omega^{\wedge n} = \int_U \Phi_t^*(\omega^{\wedge n})$$

Then we take the derivative and use the linearity of the integral:

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \int_U \Phi_t^*(\omega^{\wedge n}) = \int_U \mathcal{L}_{X_H}\omega^{\wedge n} = 0 \quad \text{since} \quad \mathcal{L}_{X_H}\omega^{\wedge n} = 0 \text{ from Theorem 8.4}$$

- In fact, all symplectomorphism preserve volume since by definition, a symplectomorphism doesn't change the forms (invariant under transformations).

7.3. LAGRANGIAN SUBMANIFOLD AND LAW OF LEAST ACTION

MOTIVATION: We had promised that the reader than Lagrangian submanifold is in many ways more important than the symplectic manifold itself and we shall see this now.

Given a symplectic manifold X with an exact symplectic form ω = dα for some 1-form α, along with a Hamiltonian H, we can define the symplectic action or Hamiltonian action of a path $\omega_2\gamma: [0,1] \to X$:

$$S(\gamma) = \int_0^1 \alpha(\dot{\gamma})\mathrm{d}t - \int_0^1 H(\gamma(t))\mathrm{d}t$$

To completely determine this system, we specify the following boundary conditions: two submanifolds L_1, L_2 of X so that we require $\gamma(0) \in L_0$ and $\gamma(1) \in L_1$. Examples might include $L_0 = T_{q0}^*Q, L_1 = T_{q1}^*Q$ for the usual Dirichlet boundary conditions, or L_0, L_1 are the zero-section for von Neumann boundary conditions.

For the one parameter family of curves, $\gamma_s: [0,1] \to X$ with $s \in \mathbb{R}$ and $\gamma_0 = \gamma$. If γ is a critical value of the symplectic action, we demand,

$$\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} S(\gamma_s) = 0$$

Now,

$$\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} S(\gamma_s) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \alpha(\dot{\gamma}_s)\mathrm{d}t - \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} H(\gamma_s(t))\mathrm{d}t$$
$$= \int_0^1 \alpha\left(\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \dot{\gamma}_s\right)\mathrm{d}t - \int_0^1 \mathrm{d}H(V(t))\mathrm{d}t = 0$$

where V (t) is the vector field along γ(t) given by:

$$V(t) = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \gamma_s(t)$$

We may change the order of s and t derivative to get,

$$\int_0^1 \alpha_{\gamma(t)}\left(\frac{\mathrm{d}}{\mathrm{d}t}V(t)\right)\mathrm{d}t - \int_0^1 \mathrm{d}H(V(t))\mathrm{d}t$$
$$= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\left(\alpha_{\gamma(t)}(V(t))\right)\mathrm{d}t + \omega(\dot{\gamma}, V(t))\mathrm{d}t - \int_0^1 \mathrm{d}H(V(t))\mathrm{d}t = 0$$

And integrating by parts gives,

$$\alpha(V(1)) - \alpha(V(0)) - \int_0^1 [\omega(\dot{\gamma}, V(t)) - \mathrm{d}H(V(t))]\mathrm{d}t = 0$$

- Assuming the first two terms are zero (justification below), the integral has to be essentially zero which implies $\omega(\dot{\gamma}, \_) = dH$. In other words, *minima of the action must necessarily occur when γ is a flow of the Hamiltonian vector field X_H.*
- We require $\alpha(V(1)) - \alpha(V(0)) = 0$ where $V(0) \in T_{\gamma(0)}L_0$ and $V(1) \in T_{\gamma(1)}L_1$. The only way this is possible for every such γ_s is if α is zero on the tangent spaces to L_1 and L_0, which essentially means $\omega|_{L_0} = \omega|_{L_1} = 0$. Since we want our spaces of possible boundary conditions to be maximal, this means we should have L_1, L_0 be **Lagrangian submanifolds** of X.

REMARKS:

- To conclude, Lagrangian are the natural place to put boundary conditions. Moreover, solutions to equations in physics are given by intersections of Lagrangian submanifolds, for example under quantization, this Lagrangian intersection corresponds to taking the correlator $\langle L_0|e^{tH}|L_1\rangle$.

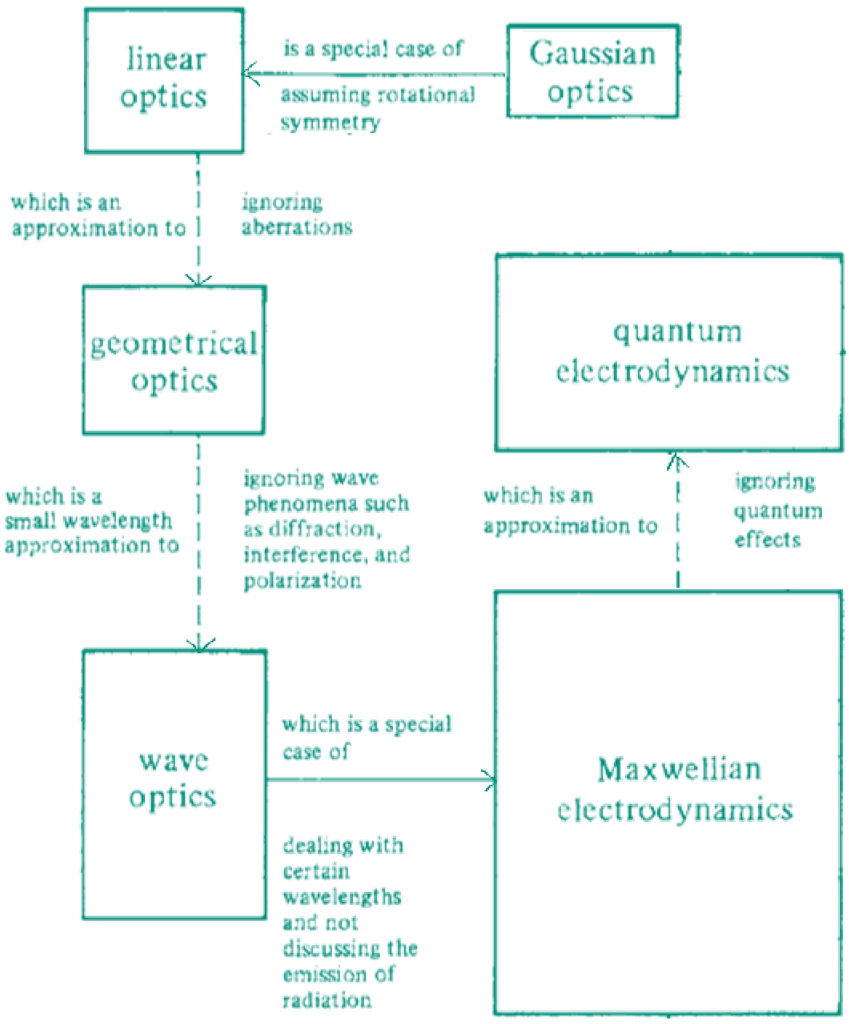
Introduction to Geometric Quantization

> "When you change the way you look at things, the things you look at change."
>
> Max Plank

8. Introduction

The development of various formulations for Optics and the associated Wave-Corpuscular conflicts have been enthusiastically portraited throughout the history of Theoretical Physics. The very heart of Quantum Mechanics and de-Broglie Wave Particle Duality was heavily inspired by their analogues in Optical Mechanics formulations. Every once in a while, when you doubt Quantum Theories its worth falling back to the good old Optics. The Modern formulation of Optics is the closest classical ancestor that draws an indistinguishable likeness between Particle and Wave.

The main purpose of this chapter is to discuss the relation between linear optics, geometric optics, and wave optics, stressing Hamilton's point of view and the corresponding relations between classical and quantum mechanics. We shall use the different formulation of Optics to draw some striking conclusions. This section will rarely use the actual machinery of Symplectic Geometry, hence making it accessible to Physics enthusiasts as well.



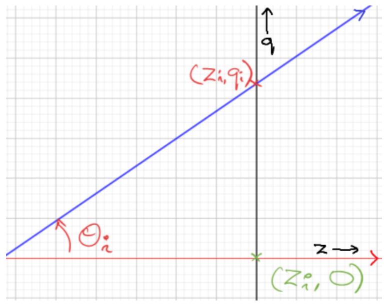
Concise Schematic of Hierarchy in Theory of Optics. Extracted from (Guillemin & Sternberg, 1990).

9. Some Formulations of Classical Optics

9.1. GAUSSIAN OPTICS

In **Gaussian optics** we are interested in tracing the trajectory of a light ray as it passes through the various refracting surfaces of the optical system (or is reflected by reflecting surfaces). By rotational symmetry, it is clearly sufficient to restrict attention to rays lying in one fixed plane and hence we study the rays which are coplanar (since due to rotational symmetry, we can always choose a plane to make rays coplanar).

For reasons that will become apparent soon, the coordinates for denoting a light ray are $\boldsymbol{q_i}$ perpendicular to z-axis and $\boldsymbol{p_i = n \sin\theta_i} \approx n\theta_i$ where θ_i is the angle between ray and the z-axis (Anticlockwise direction is positive by convention).



Schematic representation of choice of Coordinates

MOTIVATION: The problem is to relate the ray emerging from the system (can be some combination of Mirrors and Lenses) to the ray that was incident. Basically, to determine $\binom{q_2}{p_2}$ as a function of $\binom{q_1}{p_1}$. Ignoring all terms quadratic or higher (linear Approx.), we may write,

$$\binom{q_2}{p_2} = M_{21}\binom{q_1}{p_1}$$

REMARKS –

- It is worth taking note of the properties of M_{21}, namely, *orientability* ($\det M_{21} = 1$) and *area preserving* (the physical system merrily changes the direction of the incident under Linear Optics Approx.). Careful readers may have already noticed the similarity of this Determinant and the Symplectic Form.

ILLUSTRATION –

As an illustration of the procedure, let us take the opportunity to define focal Planes:

Take reference plane z_1 to lie a distance s_1 to the left of a thin lens, while z_2 lies a distance s_2 to the right of the lens. Between these planes, the matrix is,

$$\begin{pmatrix}1 & s_2\\0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\-1/f & 1\end{pmatrix}\begin{pmatrix}1 & s_1\\0 & 1\end{pmatrix} = \begin{pmatrix}1-s_2/f & s_2+s_1-s_1s_2/f\\-1/f & 1-s_1/f\end{pmatrix}$$

where s_1 and s_2 are positive in the direction of light ray.

i. When $s_1 = s_2 = f$, i.e., when the reference planes are *focal planes* of the system:
$\Rightarrow \begin{pmatrix}1 & f\\0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\-1/f & 1\end{pmatrix}\begin{pmatrix}1 & f\\0 & 1\end{pmatrix} = \begin{pmatrix}0 & f\\-1/f & 0\end{pmatrix}$ i.e. to say $\binom{q_2}{p_2} = \begin{pmatrix}0 & f\\-1/f & 0\end{pmatrix}\binom{0}{p_1}$
If a ray, incident on the lens, passes through this plane at $q_1 = 0$ with slope p_1, the outgoing ray has zero slope and so is parallel to the axis, as expected.

ii. The planes are **conjugate** if the upper-right entry of this matrix is 0 (i.e., if $s_2 + s_1 - s_1s_2/f = 0$). Thus, we obtain, $1/s_1 + 1/s_2 = 1/f$, the well-known *thin-lens equation*.

THEOREM 9.2: Any 2×2 matrix with determinant 1 can arise as the matrix of some optical system. i.e., there is an isomorphism between $Sl(2,\mathbb{R})$ and Gaussian optics.

PROOF: The proof involves two parts – the mathematical part which is essentially decomposing any such matrix into product of matrices, and the second part is mapping each of these components to the physical system.

Since M is not singular, at most both element corresponding to any diagonal can be simultaneously zero.
CASE I – If $C \neq 0$, then for any 2×2 matrix,

$$\begin{pmatrix}A & B\\C & D\end{pmatrix} = \begin{pmatrix}1 & -s\\0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\C & 1\end{pmatrix}\begin{pmatrix}1 & -t\\0 & 1\end{pmatrix} \quad \text{where} \quad \begin{cases}A + sC = 1\\ t = -(B+sD)\end{cases}$$

Refraction Operator: Corresponds to Bending of the incident light ray due to a refracting surface of power, $C = (n_2 - n_1)k$.

Translation Operator: Ray continues to travel in a straight line between two reference planes (z_1 and z_2 separated by $-t/n$) lying in the same medium of refractive index n.

The steps that lead to this decomposition is LU decomposition followed by some algebraic manipulations which we haven't shown here.

Hence this decomposition does correspond to Physical system that consists of a translation of $-t/n_1$ followed by a refraction with associated power of $(n_2 - n_1)k$ followed by another translation of $-s/n_1'$.

CASE II - If $C = 0$,

$$\Rightarrow \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ -1/f_2 & 1 \end{pmatrix} \begin{pmatrix} 1 & f_1 + f_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1/f_1 & 1 \end{pmatrix} = \begin{pmatrix} 1 - \dfrac{f_1 + f_2}{f_1} & f_1 + f_2 \\ 0 & 1 - \dfrac{f_1 + f_2}{f_2} \end{pmatrix}$$

Note that this is the only possible matrix where $C = 0$, since there is an added constraint of $\det M_{21} = 0$.

Astronomical Telescope: This is the case an astronomical telescope which consists of an objective lens of large positive focal length f_1 and an eyepiece of small positive focal length f_2 separated by a distance $f_1 + f_2$.

REMARK –

- More generally, it is not $Sl(2, \mathbb{R})$ that is related to gaussian optics, but $Sp(2, \mathbb{R})$. For $n = 2$ case it so happens that $Sl(2, \mathbb{R}) = Sp(2, \mathbb{R})$, where Sp is the symplectic transformations. This is the primary reason why we had emphasised the Orientability and Area Preserving character of these Operators earlier.

9.2. LINEAR OPTICS

Linear Optics does NOT assume rotation symmetry, but demands that all angles involved be small. Hence, its also called Paraxial Approximation.

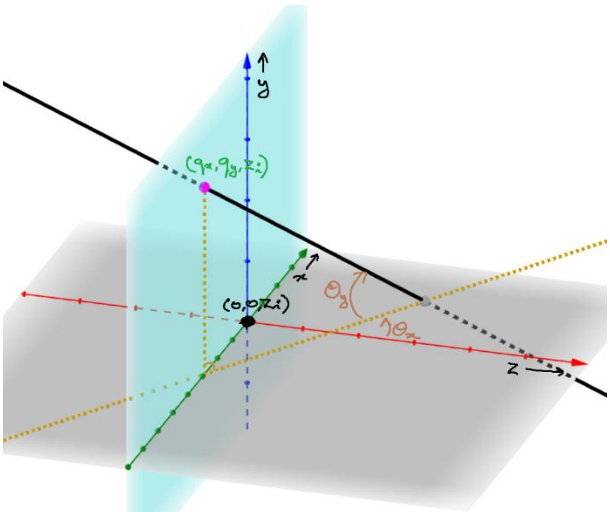
To specify four coordinates for a ray in 3D: -

$$q_x, q_y \,; \theta_x, \theta_y$$

where q_i, p_i takes the same form as in previous case. And so, for a reference plane z_1, the ray will correspond to vector,

$$u_1 = \begin{pmatrix} q_{x_1} \\ q_{y_1} \\ p_{x_1} \\ p_{y_1} \end{pmatrix} \equiv \begin{pmatrix} \bar{q}_1 \\ \bar{p}_1 \end{pmatrix}$$

which is same as the gaussian approximation except now, $\bar{q}, \bar{p}$ are 2 component vectors.



Schematic representation of choice of Coordinates

MOTIVATION: As before the natural question to ask is the relation between incident and emergent rays and what kind of 4 x 4 matrices can actually arise in linear optics?

REMARKS –

- Like the previous case, we expect these operators also to have orientability and Area Preserving characters. Further, from the representation of rays as a 4-component vector, we could draw the analogy that these operators should belong to $Sp(4, \mathbb{R})$.

THEOREM 9.3: Linear optics is isomorphic to the study of the group $Sp(4, \mathbb{R})$.

PROOF: As before, the proof involves two parts – the mathematical part which is essentially decomposing any such matrix into product of matrices, and the second part is mapping each of these components to the physical system. Here, we shall merely jot down the key ideas without getting involved in a vigorous proof.

i. Physical part – refraction at surface represented by $\begin{pmatrix} I & 0 \\ -P & I \end{pmatrix}$ where $P = P^t$ and

– motion in a medium of constant index of refraction by $\begin{pmatrix} I & dI \\ 0 & I \end{pmatrix}$ where d is the optical distance along the axis

ii. Mathematical part – every symplectic matrix can be written as a product of matrices of the above types.

REMARK –

- It is worth pointing out that Gaussian Optics is a special case of Linear Optics when the System has rotational symmetry, i.e., to say $Sp(2,\mathbb{R}) \hookrightarrow Sp(4,\mathbb{R})$.
- Alternatively, it turns out that a matrix M can arise as the transformation matrix of a linear optical system if and only if, $\omega(Mu, Mu') = \omega(u, u')$, i.e., M is a **Symplectic Transformation** and we shall formalize this more precisely (Theorem 10.3).

9.3. NOTE ON GEOMETRIC OPTICS

Geometric Optics formalism is very similar to the case of Linear optics, except for the fact that an additional analysis of aberrations is also considered. Since we are more interested in the Symplectic Techniques involved, we shall skip ahead to the next relevant topic.

10. Hamilton Formalism

MOTIVATION: We shall now focus on highlighting the Symplectic Picture through Fermat's Principle. Further, we shall be following Hamilton's Formalism because of its elegance and close association with Hamiltonian Mechanics.

HISTORIC NOTE: (Hamilton, 1828) This theory was published by Hamilton at the age of only 23 and was in fact was the guiding light for the Classical Hamiltonian Mechanics. Interested readers may find the paper here.

For the system,

$$\begin{pmatrix} \bar{q}_2 \\ \bar{p}_2 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \bar{q}_1 \\ \bar{p}_1 \end{pmatrix} \Rightarrow \begin{matrix} \bar{p}_1 = (1/B)(\bar{q}_2 - A\bar{q}_1) \\ \bar{p}_2 = (1/B)(D\bar{q}_2 - \bar{q}_1) \end{matrix} \text{ for } B \neq 0 \qquad \text{... 10-1}$$

REMARK –

- When $B = 0$ (Conjugate planes), $\bar{p}_1, \bar{p}_2$ can't be uniquely determined by $\bar{q}_i$ since there are an infinite number of light rays joining $\bar{q}_1$ and $\bar{q}_2$. Hence the system in indeterminate and we shall exclude the case for time being.

DEFINITION 10.1: The **Point characteristic**[5] / **Eikonal** of the system is defined to be,

$$W = W(\bar{q}_1, \bar{q}_2) = \left({}^1\!/_2\right)\left(B^{-1}A\,\bar{q}_1 \cdot \bar{q}_1 + B^{-1}D\,\bar{q}_2 \cdot \bar{q}_2 - 2(B^t)^{-1}\,\bar{q}_1 \cdot \bar{q}_2\right) + K \text{ where } K \text{ is a const.}$$

REMARK –

- Although such a definition may look deceptive, comparing with Eq. 10-1 notice that,

$$\bar{p}_1 = -(\partial W / \partial \bar{q}_1) \text{ and } \bar{p}_2 = \partial W / \partial \bar{q}_2$$

Curiously enough, this is way too similar to Hamiltonian equations in Classical Mechanics.

THEOREM 10.1: By an appropriate choice of the constant K, we can arrange that $W(\bar{q}_1, \bar{q}_2)$ be the "optical length"[6] of the light ray joining the path[7] $\bar{q}_1$ to $\bar{q}_2$ and is,

$$L(\gamma) \equiv W(\bar{q}_1, \bar{q}_2) = L_{\text{axis}} + {}^1\!/_2\,(p_2 q_2 - p_1 q_1) \qquad \text{... 10-2}$$

where L_{axis} denotes the optical length from z_1 to z_2 along the optical axis.

REMARK:

- This is essential the same form of Characteristic function we had defined,

$$L_{\text{axis}} + {}^1\!/_2\,(p_2 q_2 - p_1 q_1) \xrightarrow{\text{Eq 10-1}} L(\gamma) = L_{\text{axis}} + \left({}^1\!/_{2B}\right)(Aq_1^2 + Dq_2^2 - 2\bar{q}_1\bar{q}_2)$$

---

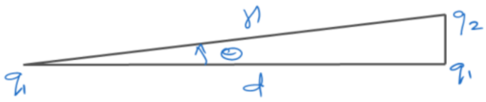
[5] To be precise, there are three characteristic functions depending on whether we want the independent variables to be only $\bar{p}$'s or only $\bar{q}$'s or a mixture of both.

[6] For a line segment of length l in a medium of constant index of refraction n, the **optical length** is nl.

[7] A **path** γ is defined as a broken line segment, l_i, where each component segment lies in a medium of constant index of refraction, n_i.

PROOF: Since vectors are involved, a general proof is very involved and the interested reader is hereby referred to the original article (Hamilton, 1828). Nevertheless, for our satisfaction we may verify that this is indeed true in the limit of Gaussian optics (since no vectors are involved a verification of the same is very straight forward).

VERIFICATION: We shall illustrate the process for one case,



i. Case when $n =$ constant: Hence $p_1 = p_2 \approx n\theta$.
Further, the optical length may be calculated using the Pythagoras Theorem as,

$$L(\gamma) = n(d^2 + (q_2 - q_1)^2)^{1/2} \approx nd + \frac{1}{2}\frac{n}{d}(q_2 - q_1)^2 = nd + \frac{1}{2}\left[\frac{n}{d}(q_2 - q_1)\right](q_2 - q_1)$$
$$= nd + \frac{1}{2}p(q_2 - q_1) \quad \text{using Eq. 10-1}$$

which is the expression 10-2 in Theorem 10.1.

ii. Similarly, the case two different refractive index on either side may be verified.

10.1. FERMAT'S PRINCIPLE

Let us fix $\bar{q}_1$ and $\bar{q}_2$ and consider the set of all paths joining $\bar{q}_1$ to $\bar{q}_2$ that consist of two segments - from $\bar{q}_1$ to $\bar{q}$ and from $\bar{q}$ to $\bar{q}_2$.

THEOREM 10.2: The actual light ray can be characterized as that path for which the optical length L takes on an extreme value, that is, for which

$$\frac{\partial L}{\partial \bar{q}} = 0$$

REMARKS -

- In the case of Gaussian Approximation, we have

$$L = n_1 d_1 + n_2 d_2 + {}^1\!/_2\{(n_1 / d_1)(q - q_1)^2 + (n_2 / d_2)(q_2 - q)^2 - (n_2 - n_1)q^2\}$$

  The extremum is:

  i. minimum if $\quad (n_1 / d_1) + (n_2 / d_2) - (n_2 - n_1) > 0$

  If $(n_2 - n_1) > 0$, minimum for small values of d_1 and d_2 while maximum for large values of d_1 and d_2

  ii. maximum if $\quad (n_1 / d_1) + (n_2 / d_2) - (n_2 - n_1) < 0$

  iii. indeterminant if $\quad (n_1 / d_1) + (n_2 / d_2) = (n_2 - n_1)$

  which is the condition that the planes are conjugate

- The fact that L is minimized only up to the first conjugate point is true in a more general setting and is known as the **Morse index theorem.** (Milnor, 1963)

10.2. FERMAT'S PRINCIPLE AND HAMILTON'S PRINCIPLE

MOTIVATON: We shall now see why we expect the transformation from one set of ray parameters to another to be symplectic, a consequence of Fermat's principle.

Given a Characteristic function $\boldsymbol{L}$, and a trajectory $\hat{\gamma}(z) = \begin{pmatrix} x \\ y \\ \dot{x} \\ \dot{y} \\ z \end{pmatrix}$, consider,

$$\int_{\hat{\gamma}} \boldsymbol{L}(x, y, \dot{x}, \dot{y}, z)dz = \int_{\hat{\gamma}} n(x, y, z)(1 + \dot{x}^2 + \dot{y}^2)^{1/2} dz$$

Also, define the mapping $\phi: \mathbb{R}^5 \to \mathbb{R}^5$,

$$q_x = x, \qquad q_y = y, \qquad z = z$$
$$p_x = \frac{\partial \boldsymbol{L}}{\partial \dot{x}} = \frac{n\dot{x}}{(1 + \dot{x}^2 + \dot{y}^2)^{1/2}}, \qquad p_y = \frac{\partial \boldsymbol{L}}{\partial \dot{y}} = \frac{n\dot{y}}{(1 + \dot{x}^2 + \dot{y}^2)^{1/2}}$$

And the Legendre Transform of $\boldsymbol{L}$ as the Function,

$$H\left(q_x, q_y, p_x, p_y, z\right) = p_x\dot{x} + p_y\dot{y} - \boldsymbol{L} \quad \text{... 10-3}$$

THEOREM 10.3: A curve ν in $\mathbb{R}^5$ given by $\nu(z) = \begin{pmatrix} q_x(z) \\ q_y(z) \\ p_x(z) \\ p_y(z) \\ z \end{pmatrix}$ will be everywhere tangent to the lines defined by $i(\eta)d\theta = 0$, for some vector η, IFF –

$$\frac{dq_x}{dz} = \frac{\partial H}{\partial p_x} \Rightarrow \frac{dx \circ \phi^{-1}\nu}{dz} = \dot{x}, \qquad \frac{dq_y}{dz} = \frac{\partial H}{\partial p_y} \Rightarrow \frac{dy \circ \phi^{-1}\nu}{dz} = \dot{y}$$
$$\frac{dp_x}{dz} = -\frac{\partial H}{\partial q_x}, \qquad \frac{dp_y}{dz} = -\frac{\partial H}{\partial q_y} \quad \text{... 10-4}$$

PROOF –

$$\int \boldsymbol{L}(x, y, \dot{x}, \dot{y}, z)dz \overset{\phi}{\rightarrow} \int \boldsymbol{L}(\nu(z))dz$$

But from Eq. 10-3,

$$\boldsymbol{L}(\nu(z))dz = \phi^*\{p_x\dot{x} + p_y\dot{y} - H\left(q_x, q_y, p_x, p_y, z\right)\}dz$$
$$\Rightarrow \boldsymbol{L}(\nu(z))dz = \phi^*\left(p_x dq_x + p_y dq_y - Hdz\right)$$

$\hat{\gamma}^* dq_x = \dot{x}dz$ and $\hat{\gamma}^* dq_y = \dot{y}dz$

- Recollect that this is actually the transformation rule for 1-forms, as in Equation 1-2.

Thus,

$$\int \boldsymbol{L}dz = \int_\nu \theta \quad \text{where} \quad \theta = p_x dq_x + p_y dq_y - Hdz$$

- It therefore defines at each point of $\mathbb{R}^5$ a one-dimensional subspace spanned by those vectors η of rank 1, that satisfy $i(\eta)d\theta = 0$.

The most general form for $i(\eta) = \partial/\partial z + A(\partial/\partial q_x) + B\left(\partial/\partial q_y\right) + C(\partial/\partial p_x) + D\left(\partial/\partial p_y\right)$

Hence,

$$i(\eta)d\theta = 0$$
$$\Rightarrow \left(C + \frac{\partial H}{\partial q_x}\right)dq_x + \left(D + \frac{\partial H}{\partial q_y}\right)dq_y - \left(A - \frac{\partial H}{\partial p_x}\right)dp_z - \left(B - \frac{\partial H}{\partial p_y}\right)dp_y$$
$$-\left(+A\frac{\partial H}{\partial q_x} + B\frac{\partial H}{\partial q_y} + C\frac{\partial H}{\partial p_x} + D\frac{\partial H}{\partial p_y}\right)dz = 0$$

Equating the coefficients of which results in the expressions.

REMARK –

- This is essentially the Hamilton's version of Euler-Lagrange equations in Optics.
- The additional constraint of Eq.10-4 means that the manifold is now 4 dimensional and,

$$\omega \coloneqq \mathrm{d}\theta = dp_x \wedge dq_x + dp_y \wedge dq_y$$

- Theorem 10.3 implies a Lagrangian Manifold hence guaranteeing the fact that these transformations are indeed symplectomorphism as claimed.

THEOREM 10.4: (Loomis & Sternberg, 2014) Calculus of Variations: $\nu(z)$ is an extremal for $\boldsymbol{L}$ if and only if $\nu(z)$ satisfies $i(\eta)d\theta = 0$, which is the Hamiltonian variation of Euler-Lagrange Equations.

PROOF: This is a very standard theorem, used alike in Variational Calculus as well a Classical Mechanics. Proof for the same may be found in any standard texts including (Loomis & Sternberg, 2014) or (Goldstein, 1980).

REMARK –

- Hence, we conclude that the trajectories predicted by Hamiltonian formalism is essentially an alternate version of Fermat's Principle.

- Theorem 10.3 and 10.4 together forms a more generalized nature of Lagrangian Subspace, which captures the wide use of Lagrangian Subspaces in Optimization problems.

11. Wave Optics

MOTIVATION: Ray Optics is unable to explain the phenomena of Interference and Diffraction which are associated with Waves. For our purpose we shall consider an "intermediate" wave formalism, where Interference and Diffraction are taken care of, but we ignore Geometric Aberrations, which are associated with non-linearity of the system.

11.1. POSTULATES

Each point of a light ray can be associated with a complex wavefunction, c were,

- $|c|$ represents the amplitude and decreases due to **attenuation** of light, also Intensity, $I \propto |c|^2$
- change in phase of c is given by $\exp 2\pi i l \,/\, \lambda$

To explain the wave phenomena, Fresnel suggested the integrality condition, where light appears to have increased intensity:

$$\Delta\phi \equiv \exp 2\pi i L(\gamma_A) \,/\, \lambda = 1$$

11.2. WAVE EXPRESSION

MOTIVATION: As always, the aim is to find $c_2(q_2),\ \forall q_2 \in z_2$ given $c_1(q_1),\ \forall q_1 \in z_1$

PROCEDURE –

i. Find $c_2(q_2)$ for a point q_2 where the wavefunction is the sum of contributions from all q_1. i.e.,

$$c_2(q_2) = \int a(q_1, q_2, \lambda) \exp\left[2\pi i \big(L(q_1, q_2)\big)/\lambda\right] c_1(q_1)\, dq_1$$

were exp factor represents the change in phase, while a is the attenuation factor together with **absolute phase** of the wave.

ii. Do this for all q_2.

11.3. ATTENUATION COEFFICIENT

For our purpose, we shall treat the system to be within Gaussian limits.

Magnitude of Attenuation coefficient –

Using the fact that total intensity of the light on the z_2 plane is the same as the total intensity of the light on the z_1 plane. i.e.,

$$\int |c_2(q_2)|^2 dq_2 = \int |c_1(q_1)|^2 dq_1$$

We get,

$$|a| = \lambda^{-1} |\det B|^{-1/2}$$

Absolute Phase of Attenuation Coefficient – $a = u|a|$

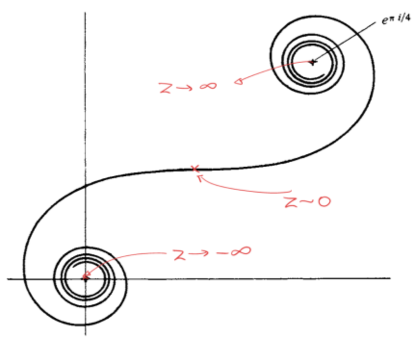
PRINCIPLE: At a distance of many wavelengths (i.e., multiples of λ) vertically distant from any stops or diaphragms in our apparatus, the light should behave approximately as if these stops are not present, i.e., say

$$c(q_1) = \begin{cases} c & \text{for } q_1 \geq 0 \\ 0 & \text{for } q_1 < 0 \end{cases}.$$

Together with the optical matrix for translation operator $\begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix}$ we have,

$$c(q_2) = uc(d\lambda)^{-1/2} \int_0^\infty \exp\left[\pi i (q_2 - q_1)^2/\lambda d\right] dq_1$$

with the famous **Fresnel Integral**[8],



A double-end Euler / Cornu spiral. The curve continues to converge to the points marked, as w tends to positive or negative infinity.

[8] Fresnel Integrals are transcendental functions whose analysis is similar to that of Gaussian Integrals and the Error Function. They commonly arise in the description of near-field Fresnel diffraction phenomena.

$$\phi(w) = \int_0^\infty \exp\,[\pi i(w-r)^2]dr$$

$$\lim_{w\to+\infty} \phi(w) = e^{\pi i/4}\ ; \quad \lim_{w\to-\infty} \phi(w) = 0$$

And hence we have the formula,

$$c(q_2) = \exp(\mp\pi i/4)\,|B\lambda|^{-\frac{1}{2}} \int c(q_1) \exp\left[\frac{2\pi i L(q_1,q_2)}{\lambda}\right] dq_1 \quad \text{for} \quad B = \pm|B| \qquad \text{... 11-1}$$

11.4. FRESNEL TRANSFORMATION VS MATRIC TRANSFORMATION

To summarize,

Gaussian Optics	Fresnel Optics
$\begin{pmatrix} q_1 \\ p_1 \end{pmatrix} \to M \begin{pmatrix} q_1 \\ p_1 \end{pmatrix} = \begin{pmatrix} Aq_1 + Bp_1 \\ Cq_1 + Dp_1 \end{pmatrix}$ with the optical matric $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$	$c_1 \to (Fc_1)(q_2) =$ $\exp\,(\mp\pi i/4)\|B\lambda\|^{-1/2} \int c_1(q_1) \times \exp[2\pi i L(q_1,q_2)/\lambda]\, dq_1$ where $L(q_1,q_2) = \frac{1}{2B}(Aq_1^2 + Dq_2^2 - 2q_1q_2) + d$ for $B \neq 0$
- works only for single ray	- based on the superposition principle.

TABLE I: Mapping Gaussian Optics to Fresnel Optics

MOTIVATION: It may be tempting to take the correspondence between Gaussian and Fresnel operators further. Before proceeding ahead, there is an important observation to make.

Notice that unlike the Gaussian Operators where $M_{13} = M_{12} \circ M_{23}$, $F_{13} \neq F_{12} \circ F_{23}$ precisely because of the $e^{\pm i\pi/4}$ factor introduced in the previous section. To rectify this, either,

i. Ignore the absolute phase factor, or
ii. Define the operators of Fresnel optics (in the Gaussian approximation) to form a **double cover**[9] of the group of 2×2 matrices of determinant 1. (Since this is more elegant, we shall briefly expand upon this).

The correspondence between Matrix Operators and Fresnel Operators can be easily matched using the relevant characteristic Function. Consider a general n-dimensional case of $Sp(2n,\mathbb{R})$, while keeping the sign of $|B|$ fixed, we have the generators,

Gaussian Optics	Fresnel Optics in Gaussian Approx.
- Matrix Operators	- Integral Operators / Generators
$\begin{pmatrix} I & dI \\ 0 & I \end{pmatrix}, d > 0$	$(U_d c)(x) = \exp\,(-\pi in/4) d^{-n/2}(2\pi)^{-n/2} \int \exp\,[i(x-y)^2/2d]c(y)dy$ [10]
$\begin{pmatrix} I & 0 \\ -P & I \end{pmatrix}, P = P^t$	$(V_{\mathrm{p}}c)(x) = \exp\,(-iPx\cdot x/2)c(x)$

TABLE I: Gaussian Operators and their Wave Operator counterparts.

That is, for the map ρ,

$$\rho(U_d) = \begin{pmatrix} I & d \\ 0 & I \end{pmatrix} \text{ and } \rho(V_{\mathrm{P}}) = \begin{pmatrix} I & 0 \\ -P & I \end{pmatrix}$$

If $G = \text{span}\,\{U_d, V_d\}$ effectively we have a *Homomorphism between G and $Sp(2n,\mathbb{R})$*, since we have seen that $Sp(2n,\mathbb{R}) \equiv \text{span}\,\{\rho(U_d), \rho(V_{\mathrm{P}})\}$.

MOTIVATION: Adding a correction factor when the ray passes through a conjugate point ($B = 0$) and generalizing, we have the interesting Theorem.

[9] Naively speaking, saying that for example the metaplectic group Mp_{2n} is a double cover of the symplectic group Sp_{2n} means that there are always two elements in the metaplectic group representing one element in the symplectic group.

[10] It turns out that this is an incomplete description. Not only does the phase of light get multiplied by $e^{id/\lambda}$ as we move along a ray. We get an extra factor of $-i = e^{-i\pi/4} \cdot e^{-i\pi/4}$ as the light passes through a conjugate point. Although this may be derived within our analysis, we have skipped it for the lack of space. Experimental evidence (Gouy, 1890)

THEOREM 11.2: If $X \in G$ is such that $\rho(X) = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ with B nonsingular, then the operator X is given by,

$$(Xc)(x) = i^{\#} e^{-\pi i n/4} |\det\, B|^{-1/2} (2\pi)^{-n/2} \int e^{iW(x,y)} c(y) dy$$

where, $W(x,y) = \frac{1}{2}[DB^{-1}x \cdot x + B^{-1}Ay \cdot y - 2B^{t-1}y \cdot x]$ and $\# = \begin{cases} \text{even if } \det B > 0 \\ \text{odd if } \det B < 0 \end{cases}$

REMARKS –

- Since $\boldsymbol{\rho}$ **is 2 to 1** everywhere; that is, that $G \equiv \mathrm{Mp}(2n, \mathbb{R})$ is a double covering of the symplectic group, $Sp(2n, \mathbb{R})$and hence called **Metaplectic group.**
- Fresnel optics is equivalent to the study of the metaplectic representation of $\mathrm{Mp}(4, \mathbb{R})$.

11.5. FRESNEL GENERATOR AND QUANTUM MECHANICAL OPERATORS

MOTIVATION - Just as there was a wave optics that was a more accurate physical theory than Hamilton's geometrical optics, there should be a wave mechanics standing in the same relation to classical mechanics. We shall show that for linear mechanical systems, a precise mathematical analogy may be drawn which turns out to be the relation between the metaplectic representation and the symplectic group $Sp(n, \mathbb{R})$.

DISCLAIMER – This session just aims to provide a glimpse of Geometric Quantization. That means to say what follows is not mathematically vigorous and many underlying theorems are oversimplified and used without proper justification. A standard treatment is well beyond the scope of this manuscript.

Any family of operators $U(t) \in Mp(2n, \mathbb{R})$ that depend continuously on t and satisfy $U(t_1 + t_2) = U(t_1) + U(t_2)$ is called a *one-parameter group of operators.* To such group there will correspond one-parameter group of matrices $M(t)$, which are close to unity for small $|t|$. For such small $|t|$ we can uniquely recover $U(t)$ from $M(t)$, since only one of the two possible choices of the Fresnel operator corresponding to $M(t)$ will be close to the identity. And we have $U(t) = U^n(t/n)$

PART I: Since $M(t)$ is a one-parameter group of symplectic matrices, it can be shown that M depends differentiable on t.

$$\text{Take,} \quad M'(0) = K \quad \Rightarrow \quad M(t) = e^{tK}$$

PART II: Also, we can have,

$$U'(0)c = \lim_{t \to 0} \frac{U(t)c - c}{t}$$

- Here this limit need not exist for all c, but we do expect it to exist for well-behaved c

It is a standard theorem that $U'(0)$ is a skew adjoint operator,

$$\Rightarrow \quad \frac{d\boldsymbol{U}}{d\boldsymbol{t}} = -\boldsymbol{iHU} \quad - \quad \textbf{Schrodinger equation}$$

PART I		PART II	
K	$M(t)$	H	$[U(t)f]$
$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$	$-\frac{1}{2}\frac{d^2}{dx^2}$	$\exp(\pm\pi i/4)\|t\|^{-\frac{1}{2}} \int e^{i(x-y)^2/2t} f(y) dy$
$\begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ -t & 1 \end{pmatrix}$	$\frac{1}{2}x^2$	$e^{-itx^2/2} f(x)$
$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$	$\begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}$	$\frac{1}{2}\left(-\frac{d^2}{dx^2} + x^2\right)$	$(-i)^{\#} e^{-\frac{\pi i}{4}} (\|\sin t\| 2\pi)^{-\frac{1}{2}} \int \exp\{i[\cos t\,(x^2 + y^2) - 2xy]/2 \sin t\} f(y) dy$

TABLE II: The relation between part I and part II are a correspondence between the matric operators, $M(t)$ and Fresnel operators, $U(t)$. For example - the 1st entry represents a constant change in q with t which corresponds to the Quantum Hamiltonian of free particle and the last entry corresponds to that of Harmonic Oscillator. Further, the last entry is effectively a sum of the first two entries.

REMARKS -

- Notice the advantage that $U(t)$ is already the solution of H for $m = 1$ and $\hbar = 1$ or replace $t \to t\hbar / m$.
- Notice that the mapping between K and H is linear (on any common domain of definition) – 3rd row entry is sum of first two rows.
- In quantum correspondence we talk of probability density instead of Intensity of square integrable functions ($L^2(\mathbb{R})$)

The general ($t \to t\hbar / m$) Harmonic oscillator solution for phase space is called **Mehler's Formula**. For the case of the harmonic oscillator the one-parameter group $U(t)$ is a representation of a compact group, the circle, and, hence, on general principles (the **Peter-Weyl theorem**), the space $L^2(\mathbb{R})$ decomposes into a direct sum of irreducible that must be one-dimensional, since the circle is abelian. Thus, $\frac{1}{2}(-d^2/dx^2 + x^2)$ has a discrete spectrum.

Bibliography

Cannas da Silva, A. (2008). *Lectures on Symplectic Geometry* (1 ed., Vol. 1). Berlin, Heidelberg: Springer. doi:10.1007/978-3-540-45330-7

Darboux, G. (1882). Sur le problème de Pfaff. *Bibliothèque nationale de France, département Sciences et techniques, 8-V. Pièce-4068*, 42.

Goldstein, H. (1980). *Classical Mechanics.* Boston: Addison-Wesley.

Gouy, L. G. (1890). *Sur une propriété nouvelle des ondes lumineuses.* Paris: Gauthier-Villars.

Guillemin, V., & Sternberg, S. (1990). *Symplectic Techniques in Physics.* UK: Cambridge University Press.

Hamilton, W. R. (1828). Theory Of Systems Of Rays. (D. R. Wilkins, Ed.) *Transactions of the Royal Irish Academy*, 69–174. Retrieved from https://www.maths.tcd.ie/pub/HistMath/People/Hamilton/Rays/PtFst.pdf

Loomis, L. H., & Sternberg, S. (2014). *Advanced Calculus* (Revised ed.). Tokyo: World Scientific. doi:10.1142/9095

McDuff, D., & Salamon, D. (2017). *Introduction to Symplectic Topology* (Vol. 3rd). UK: Oxford University Press. doi:10.1093/oso/9780198794899.001.0001

Milnor, J. (1963). Morse Theory. *Annals of Mathematics Studies*, 160. Retrieved from https://www.maths.ed.ac.uk/˜v1ranick/papers/milnmors.pdf

Index