Classical and Quantum Reduction of the Hydrogen Atom

J.V. Gaiter
University of Colorado Boulder
Department of Mathematics
April 4th, 2022

Thesis Advisor:
Dr. Markus Pflaum, Department of Mathematics

Committee Members:
Dr. Magdalena Czubak, Department of Mathematics,
Dr. Markus Pflaum, Department of Mathematics,
Dr. Steven Pollock, Department of Physics

Abstract

In this thesis we lay out an overview of the mathematics required to understand the reduction of symmetry of the classical Kepler problem and its quantum counterpart the hydrogen atom within the framework of symplectic geometry and deformation quantization respectively. In order to do so, we must cover the mathematical generalization of Hamiltonian mechanics, symplectic geometry. We will show how conservation of energy and Liouville's theorem manifest themselves within this generalization using the tools of differential geometry. After introducing differential and symplectic geometry, we give brief introduction to Lie Groups, their actions on smooth manifolds, and moment maps. Lie groups allow us to formalize the very physical ideas behind continuous symmetry, and they play a principal role in the reduction of symmetry, along with moment maps, as we will see in the latter half of this thesis. After having developed the prerequisite theory, we then tackle the reduction of the Kepler problem, otherwise known as the two body problem. This section makes mathematically rigorous the method by which the equations of motion for the two body problem are obtained in a typical undergraduate analytical mechanics course. After the classical reduction is finished we motivate the idea of quantization as P.A.M. Dirac does in his 1925 paper [2]. We then cite the Gronewold van Hove no go theorems [3] to show that such a quantization scheme can not be strictly satisfied even on $\mathbb{R}^{2n}$ and use this to motivate the definitions underlying deformation quantization. After defining deformation quantizations we give a few examples of such structures on $\mathbb{R}^{2n}$ and reduce one such example, the Weyl-Moyal product using the classical reduction tools we developed for the Kepler problem.

1 Introduction

Each approach to classical mechanics has particular strengths, making some more suited for certain theoretical or practical applications than others. Newtonian mechanics is applicable to the widest array of problems, but is unwieldy for certain systems; Lagrangian mechanics allows one to find the equations of motion for systems with relative ease, but can hide important symmetries between position and momentum. When a classical system has a large amount of symmetry, Hamiltonian mechanics allows one to easily spot these symmetries via the use of conserved quantities. One can even use these symmetries to reduce the complexity of a problem and to transform a set of opaque equations into something more tractable.

Another factor in Hamiltonian mechanics' utility is its emphasis on observables as fundamental components of the theory. Where Lagrangian mechanics begins with the Lagrangian, a convenient theoretical tool which is not commonly measured or studied in of itself, Hamiltonian mechanics begins with the Hamiltonian, a quantity which in many cases represents the real energy of the system. One can compute the Poisson bracket to determine how arbitrary observables evolve in time as well as to detect symmetries in the system. It is this emphasis on observables that makes the transition between classical mechanics and quantum mechanics easiest from the Hamiltonian point of view.

To move from the classical realm to the quantum one begins with a classical system. In most cases it suffices to think of our configuration space as $\mathbb{R}^n$, this is the set of all positions our system is allowed to take. The phase space of such a system is described by $\mathbb{R}^n \times \mathbb{R}^n$ where each point $(\mathbf{q}, \mathbf{p})$ describes both the position and momentum of the system. We are given a Hamiltonian $H\colon \mathbb{R}^{2n} \to \mathbb{R}$ where $H(\mathbf{q}, \mathbf{p})$ describes the energy of the system with position $\mathbf{q}$ and momentum $\mathbf{p}$. We can describe the time evolution of an observable $f\colon \mathbb{R}^{2n} \to \mathbb{R}$ described by the equation

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \{f, H\}.$$

we then "quantize" our observables by replacing them with operators on Hilbert space, which is the configuration space of a quantum mechanical system. In doing so we replace the Poisson bracket $\{f, H\}$ by the commutator $\frac{i}{\hbar}\left[\hat{H}, \hat{f}\right]$ to get a non-commutative set of observables that are inspired by the classical observables and their Poisson bracket relations. The idea of replacing the commutative algebra of observables with a non-commutative one with Poisson brackets replaced by commutators is originally due to P.A.M. Dirac [2].

This transition, up to a certain order, preserves the symmetries of the system. A quantity g is conserved by the dynamics generated by H if and only if $\{g, H\} = 0$, so when we replace Poisson brackets with commutators, we see that $\left[\hat{H}, \hat{g}\right] = 0$ and $\frac{\mathrm{d}\hat{g}}{\mathrm{d}t} = 0$ so our quantum observable does not evolve in time.

As with many things relating to quantum mechanics, things are not as we might like them to be. The Groenewold van Hove no go theorems [3] tell us that we cannot quantize all observables on the classical phase space $\mathbb{R}^{2n}$ in a way that replaces Poisson brackets with commutators exactly.

These no go theorems are what inspired Bayen Flato, Frønsdal, Lichnerowicz, and Sternheimer to lay the foundations for deformation quantization in their 1978 paper [1]. The key idea of deformation quantization is to take the algebra of observables on a classical system, i.e. the Poisson alegbra $(C^\infty(P), \cdot, \{\cdot, \cdot\})$ where P is a Poisson manifold, and deform it into a non-commutative algebra $C^\infty(P)[[\hbar]]$ (the space of formal power series of $\hbar$ with coefficients in $C^\infty(P)$ in a way that satisfies the Dirac quantization rules up to first order in $\hbar$. Within this theory, one does not consider $\hbar$ to be a real physical constant, but rather a formal parameter. When we take the limit $\hbar \to 0$ one recovers the classical algebra of observables.

Throughout this thesis, we will develop the ideas behind symplectic geometry, the natural generalization of classical mechanics to smooth manifolds, Lie groups, a convenient way to describe the symmetries of both classical and quantum systems, and deformation quantization, a formal way of producing non-commutative quantum observables from the classical Poisson bracket.

2 Differential Geometry

Smooth manifolds provide a natural setting to study classical and quantum mechanics. The space of classical configurations of systems can often be represented by smooth manifolds, e.g. the position of a particle on a ring can be described by a point in S^1, the circle. The most important method in classical mechanics is the generation of a differential equation which describes the time evolution of a system, and smooth manifolds give one a coordinate independent way of generating and describing these differential equations.

2.1 Smooth Manifolds

Before we define a smooth manifold, we need to define the concept of a smooth atlas. Let M be a topological space. A smooth atlas is much like its cartographic namesake. It is a collection of open sets U_α, which cover M, i.e. $\bigcup_\alpha U_\alpha = M$. Continuing the atlas analogy, the covering condition means that every point $p \in M$ is contained in a "page" U_β of the atlas. Within a smooth atlas each U_α comes with a map $\varphi_\alpha \colon U_\alpha \to \mathbb{R}^m$, called a chart, such that $\varphi_\alpha \colon U_\alpha \to \varphi(U_\alpha)$ is a homeomorphism and $\varphi_\alpha \circ \varphi_\beta^{-1} \colon \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta)$, which is a map between open sets of $\mathbb{R}^m$, is C^∞, i.e. all partial derivatives of all orders of $\varphi_\alpha \circ \varphi_\beta^{-1}$ exist and are continuous. Formally elements of an atlas are the pairs $(U_\alpha, \varphi_\alpha)$. The atlas analogy begins to break down with our second condition, but it essentially tells us that locally a space with a smooth atlas looks like $\mathbb{R}^m$ for some integer m. The integer m is called the dimension of M.

Definition 2.1. Let M be a topological space. M is said to be a *smooth manifold* if it is second countable, Hausdorff, and has a smooth atlas $\mathscr{A}$.

The smoothness condition is best illustrated in terms of coordinate representations of functions. Given $p \in M, p \in U_\alpha$ one calls the coordinates of p in the chart U_α, $\varphi_\alpha(p) = (\varphi_\alpha^1(p), \varphi_\alpha^2, \ldots, \varphi_\alpha^m(p))$. Then given a function $f \colon M \to \mathbb{R}$ and a chart φ_α we can form a *representative function* $\tilde{f} = f \circ \varphi_\alpha^{-1}$. The function $\tilde{f}$ allows us to think of f as a function from $\varphi_\alpha(U_\alpha)$ to $\mathbb{R}$ where we know how to apply conventional multivariable calculus. In order for our definition of a smooth function to coincide with that on $\mathbb{R}^m$, we give the following definition.

Definition 2.2. Let M be a smooth manifold with atlas $\mathscr{A} = (U_\alpha, \varphi_\alpha)_{\alpha \in A}$. A function $f \colon M \to \mathbb{R}$ is said to be C^k for $k \in \mathbb{N} \cup \{\infty, \omega\}$ if for each $p \in M$ there is a $\beta \in A$ with $p \in U_\beta$ such that $f \circ \varphi_\beta^{-1} \colon \varphi_\beta(U_\beta) \to \mathbb{R}$ is C^k. A function is said to be *smooth* if it is C^∞.

This effectively means that a function f is smooth if for every point p there is a chart U_β containing p such that the representative function is smooth. Because we assume that $\varphi_\alpha \circ \varphi_\beta^{-1}$ is smooth, we have $f \circ \varphi_\alpha^{-1}\big|_{\varphi_\alpha(U_\beta \cap U_\alpha)} = f \circ \varphi_\beta^{-1} \circ \varphi_\beta \circ \varphi_\alpha^{-1}$ which is smooth by the chain rule. So a function with a smooth representative in one chart has a smooth representative in another chart in their overlap.

Definition 2.3. Let M, N be smooth manifolds with smooth atlases $\mathscr{A}$ and $\mathscr{B}$ respectively. A map $f \colon M \to N$ is said to be smooth if for each $p \in M$ there are charts $(U, \varphi) \in \mathscr{A}$ and $(V, \psi) \in \mathscr{B}$ with $p \in U$ and $f(p) \in V$ such that $\psi \circ f \circ \varphi^{-1} \colon \varphi(f^{-1}(V) \cap U) \to \psi(V)$ is a smooth function.

The preceding definition generalizes the notions of the one before it. We create a representative function $\tilde{f}$ now by pre and post composing by chart functions and declare a function to be smooth if it has a smooth representative at every point.

Example 1. The simplest example of a smooth manifold is $\mathbb{R}^n$. This is second countable and Hausdorff when equipped with the standard topology. We can give it a single chart $\mathrm{Id}_{\mathbb{R}^n} \colon \mathbb{R}^n \to \mathbb{R}^n$, the identity. By our definition, $C^\infty(M)$ coincides with the typical multivariable calculus definitions. One can use the identity chart to make any open subset of $\mathbb{R}^n$ into a smooth manifold.

2.2 The Tangent Bundle

Now that we have defined our object of study, we can define what will be one of the most important tools for doing classical mechanics on manifolds and for studying smooth manifolds in general, the tangent vectors.

There are several independent ways to define the tangent space to a smooth manifold. Here we will give the definition most natural when generalizing classical mechanics.

Definition 2.4. Let M be a smooth manifold with atlas $\mathscr{A}$. Given $p \in M$, let C_p be the set of smooth functions $\gamma\colon (-\epsilon, \epsilon) \to M$ with $\gamma(0) = p$. We define two curves $\gamma, \eta \in C_p$ to be equivalent if there is a chart $(U, \varphi) \in \mathscr{A}$ with $p \in U$ and

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\varphi \circ \gamma)(t) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\varphi \circ \eta)(t).$$

We then define the *tangent space* of M at p to be $T_pM := C_p/\sim$

Before proceeding, we must show that the equivalence relation defined above is well defined.

Proposition 2.5. *The equivalence relation $\sim$ defined on C_p is well defined.*

Proof. Reflexivity and symmetry both follow from the reflexivity and symmetry of equality so we need only show that the relation is transitive. Let $\gamma, \eta, \rho \in C_p$ with $\gamma \sim \eta$ and $\eta \sim \rho$. These conditions mean that

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\varphi \circ \gamma)(t) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\varphi \circ \eta)(t) \quad \text{and} \quad \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\psi \circ \eta)(t) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\psi \circ \rho)(t)$$

for some charts φ and ψ. We then see that

$$\begin{aligned}\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\psi \circ \gamma)(t) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\psi \circ \varphi^{-1} \circ \varphi \circ \gamma)(t) \\ &= D(\psi \circ \varphi^{-1})(\varphi(p)) \left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\varphi \circ \gamma)(t) \right)\end{aligned}$$

after applying the chain rule. Then substituting for η and applying the chain rule we have:

$$\begin{aligned}&= D(\psi \circ \varphi^{-1})(\varphi(p)) \left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\varphi \circ \eta)(t) \right) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\psi \circ \varphi^{-1} \circ \varphi \circ \eta)(t) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\psi \circ \eta) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\psi \circ \rho)\end{aligned}$$

and $\gamma \sim \rho$. $\square$

This proves that $\sim$ is well defined and $C_p/\sim$ is a well defined set. We will now see how to give T_pM a vector space structure.

Proposition 2.6. *Let M be a smooth manifold and $\mathscr{A}$ an atlas. The map $d_p\varphi\colon T_pM \to \mathbb{R}^m$ by $[\gamma] \mapsto \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0}(\varphi\circ\gamma)$ is a bijection for $(U,\varphi)\in\mathscr{A}$ and $p\in U$.*

Proof. Since we specified a map on T_pM by choosing a representative, we first need to show that $d_p\varphi$ is well defined. Given $\gamma,\gamma'\in[\gamma]\in T_pM$ we know that for some chart ψ containing p,

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(\psi\circ\gamma) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(\psi\circ\gamma').$$

We then have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(\varphi\circ\gamma) &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(\varphi\circ\psi^{-1}\circ\psi\circ\gamma)\\ &= D(\varphi\circ\psi^{-1})(\psi(p))\left(\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(\psi\circ\gamma)\right)\\ &= D(\varphi\circ\psi^{-1})(\psi(p))\left(\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(\psi\circ\gamma')\right)\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(\varphi\circ\psi^{-1}\circ\psi\circ\gamma')\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(\varphi\circ\gamma')\end{aligned}$$

So $d_p\varphi$ is well defined.

We now show that $d_p\varphi$ is surjective. Given $v\in\mathbb{R}^m$ we can create a curve $\eta_v(t)=\varphi^{-1}(\varphi(p)+tv)$. We know that $\eta_v\in C_p$ since $\eta_v(0)=\varphi^{-1}(\varphi(p)+0)=p$ and $(\varphi\circ\eta_v)(t)=\varphi(p)+tv$ which is a smooth function. We then observe

$$d_p\varphi([\eta_v]) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(\varphi(\varphi^{-1}(\varphi(p)+tv))) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(\varphi(p)+tv) = v$$

meaning that $d_p\varphi$ is surjective.

We see that $d_p\varphi$ is injective as follows. Take $[\gamma],[\rho]$ with $d_p\varphi[\gamma]=d_p\varphi[\rho]$. This implies that

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(\varphi\circ\gamma) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(\varphi\circ\rho)$$

and hence $[\gamma]=[\rho]$. ☐

Since T_pM is in bijective correspondence with a vector space, we can make it inherit the vector space operations from $\mathbb{R}^m$, i.e.

$$\alpha\cdot[\gamma] := (d_p\varphi)^{-1}(\alpha\cdot d_p\varphi[\gamma])$$

and

$$[\gamma] + [\eta] := (d_p\varphi)^{-1}(d_p\varphi[\gamma] + d_p\varphi[\eta]).$$

for all $\alpha \in \mathbb{R}, [\gamma], [\eta] \in T_pM$. But, since this definition relies on a choice of chart φ, for this definition to be meaningful we need to show that it does not infact depend on the choice of φ.

Proposition 2.7. *Let M be a smooth manifold for which ψ and φ are charts containing p. Then the multiplications and additions defined on T_pM using $d_p\varphi$ and $d_p\psi$ coincide. That is:*

$$(d_p\varphi)^{-1}(\alpha \cdot d_p\varphi[\gamma]) = (d_p\psi)^{-1}(\alpha \cdot d_p\psi[\gamma])$$

and

$$(d_p\varphi)^{-1}(d_p\varphi[\gamma] + d_p\varphi[\eta]) = (d_p\psi)^{-1}(d_p\psi[\gamma] + d_p\psi[\eta])$$

for all $\alpha \in \mathbb{R}, [\gamma], [\eta] \in T_pM$.

Proof. We begin with the scalar multiplication formula. We see that

$$(d_p\psi)^{-1}(\alpha \cdot d_p\psi[\gamma]) = (d_p\psi)^{-1}(\alpha \cdot d_p\psi(d_p\varphi)^{-1}d_p\varphi[\gamma])$$

Now, we will prove that $d_p\psi(d_p\varphi)^{-1} = D(\psi \circ \varphi^{-1})(\varphi(p))$. We have $(d_p\varphi)^{-1}(v) = [\varphi^{-1}(\varphi(p) + tv)]$ so

$$d_p\psi(d_p\varphi)^{-1}(v) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(\psi(\varphi^{-1}(\varphi(p) + tv)) = D(\psi \circ \varphi)(\varphi(p))\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(\varphi(p) + tv))$$

proving our claim. This then gives

$$\begin{aligned}\alpha d_p\psi(d_p\varphi)^{-1}d_p\varphi[\gamma] &= \alpha D(\psi \circ \varphi^{-1})(\varphi(p))d_p\varphi[\gamma] \\ &= D(\psi \circ \varphi^{-1})(\varphi(p))\alpha \cdot d_p\varphi[\gamma] \\ &= d_p\psi(d_p\varphi)^{-1}\alpha \cdot d_p\varphi[\gamma]\end{aligned}$$

and

$$\begin{aligned}(d_p\psi)^{-1}(\alpha \cdot d_p\psi[\gamma]) &= (d_p\psi)^{-1}(d_p\psi(d_p\varphi)^{-1}(\alpha d_p\varphi[\gamma]) \\ &= (d_p\varphi)^{-1}(\alpha d_p\varphi[\gamma])\end{aligned}$$

Similarly:

$$\begin{aligned}(d_p\psi)^{-1}(d_p\varphi[\gamma] + d_p\varphi[\eta]) &= (d_p\varphi)^{-1}(d_p\varphi(d_p\psi)^{-1}d_p\psi[\gamma] + d_p\varphi(d_p\psi)^{-1}d_p\psi[\eta]) \\ &= (d_p\psi)^{-1}(D(\psi \circ \varphi^{-1})(\varphi(p))d_p\varphi[\gamma] + D(\psi \circ \varphi^{-1})(\varphi(p))d_p\varphi[\eta]) \\ &= (d_p\psi)^{-1}(D(\psi \circ \varphi^{-1})(\varphi(p))(d_p\varphi[\gamma] + d_p\varphi[\eta]) \\ &= (d_p\psi)^{-1}(d_p\psi(d_p\varphi)^{-1}(d_p\varphi[\gamma] + d_p\varphi[\eta]) \\ &= (d_p\varphi)^{-1}(d_p\varphi[\gamma] + d_p\varphi[\eta])\end{aligned}$$

So the vector space structure on T_pM does not depend on the choice of chart. $\square$

Given $\gamma\colon (a,b) \to M$ a smooth curve, we define $\dot\gamma(t) := [\tilde\gamma_t]$ where $\tilde\gamma_t(s) = \gamma(s-t)$. Given a smooth map $f\colon M \to N$ we can define the derivative map $T_pf\colon T_pM \to T_{f(p)}N$ by $T_pf[\gamma] = [f \circ \gamma]$.

Since we can associate to each point $p \in M$ the vector space T_pM, we might want to look at the space of all tangent spaces, to do this we define $TM := \bigsqcup_{p\in M} T_pM$, the *tangent bundle* of M. As we do not want to leave the realm of smooth manifolds, we want to give TM a smooth manifold structure. Here we have not given TM the disjoint union topology, it will inherit one from the smooth structure we give it.

Assume M is a smooth manifold with atlas $\mathscr{A}$. To each chart (ψ, V) we associate a new chart $(d\psi, TV)$ where $TV := \bigsqcup_{p\in V} T_pM$ and $d\psi\colon TV \to \psi(V)\times\mathbb{R}^m$ is defined as $d\psi[\gamma] := (\psi(\gamma(0)), d_{\gamma(0)}\psi[\gamma])$. This is clearly bijective, since we have the two sided inverse $(x, v) \mapsto (d_{\psi^{-1}(x)}\psi)^{-1}(v)$. We can then define a topology on TM by taking $U \subset TM$ open if $d\psi(U \cap V)$ is open for all charts $(\psi, V) \in \mathscr{A}$.

To show that TM is a manifold with this given topology, we must first start by showing that it is second countable and Hausdorff.

Proposition 2.8. *Let M be a smooth manifold with smooth atlas $\mathscr{A}$. The tangent bundle of M, TM is second countable and Hausdorff.*

Proof. We start by proving that TM is second countable. Since M is a manifold it is second countable and hence can be covered in a countable collection of charts $(U_i, \varphi_i)_{i\in\mathbb{Z}}$. Since TU_i is homeomorphic to $\varphi_i(U_i) \times \mathbb{R}^n$ it has a countable base B_i. Then $\bigcup_{i\in\mathbb{Z}} B_i$ defines a base for TM which is countable.

We know show that TM is Hausdorff. Take $[\gamma], [\rho] \in TM$ with $[\gamma] \neq [\rho]$ and $\gamma(0) \neq \rho(0)$. If there is no $V \in \mathscr{A}$ with $TV \ni [\gamma], [\rho]$ then any two charts TV and TU containing $[\gamma]$ and $[\rho]$ respectively will be separating neighborhoods.

Assume that $[\gamma], [\rho] \in TV$ for some chart (V, ψ) on M but with $\gamma(0) \neq \rho(0)$. Since M is Hausdorff, take $W, X \subset V$ with $\gamma(0) \in W, \rho(0) \in X$. We then have $TW \cap TX = \emptyset$ and TW, TX are open giving us separating neighborhoods of $[\gamma]$ and $[\rho]$.

Now assume that $\gamma(0) = \rho(0) \in V$. We can take separating open neighborhoods of $d\psi[\gamma]$ and $d\psi[\rho]$ in $\psi(V) \times \mathbb{R}^m$ since $\mathbb{R}^{2m}$ is Hausdorff. Call these neighborhoods U and W. Then $d\psi^{-1}(U)$ and $d\psi^{-1}(W)$ are open neighborhoods of $[\gamma]$ and $[\rho]$ respectively which do not intersect. Hence TM is Hausdorff. $\square$

As we have defined the charts, we only need to check that the transitions $d\psi \circ (d\varphi)^{-1}$ are smooth in order to conclude that TM is a smooth manifold.

We have

$$\begin{aligned} d\psi \circ (d\varphi)^{-1}(x,v) &= d\psi(\varphi^{-1}(x), (d_{\varphi^{-1}(x)}\varphi)^{-1}v) \\ &= ((\psi \circ \varphi^{-1})(x), (d_{\varphi^{-1}(x)}\psi(d_{\varphi^{-1}(x)}\varphi)(v)) \\ &= ((\psi \circ \varphi^{-1})(x), D(\psi \circ \varphi^{-1})(\varphi^{-1}(x))(v)) \end{aligned}$$

Which is smooth by the compatibility of ψ and φ.

Because TM is constructed directly from M we get a natural function $\pi_{TM} \colon TM \to M$ which takes $[\gamma]$ to $\gamma(0)$. We see that π_{TM} is a smooth map as follows. Let (ψ, V) be a chart for M and $(d\psi, TV)$ the corresponding chart of TM. Then $\psi \circ \pi_{TM} \circ d\psi^{-1}(x,v) = \psi(\pi_{TM}([\psi^{-1}(x+tv)])) = \psi(\psi^{-1}(x)) = x$. Since the representative function is smooth π_{TM} is smooth. TM is an example of what is known as a vector bundle, a space which locally looks like the product of a vector space and a smooth manifold.

Now that we have defined the tangent bundle, we can define the space of vector fields on M.

Definition 2.9. Let M be a smooth manifold. The space of vector fields on M is the set of sections of the bundle $\pi \colon TM \to M$, i.e. smooth maps $\sigma \colon M \to TM$ such that $\pi_{TM} \circ \sigma = \mathrm{Id}_M$. The set of all vector fields is denoted $\mathfrak{X}(M)$ or $\Gamma(M, TM)$.

One can write out a vector field locally as $X(p) = \sum X^i \frac{\partial}{\partial x^i}(p)$ where $\frac{\partial}{\partial x^i}(p) = (d_p\varphi)^{-1}(e^i)$.

Definition 2.10. Let $f \colon M \to N$ be a diffeomorphism between smooth manifolds M and N, that is f is a smooth bijection with smooth inverse f^{-1}. We can define the pushforward of $f_* \colon \mathfrak{X}(M) \to \mathfrak{X}(N)$ which takes a vector field X on M and maps it to $(p \mapsto T_{f^{-1}(p)} f X(f^{-1}(p))$.

One can view vector fields as the space of maps $\delta \colon C^\infty(M) \to C^\infty(M)$ such that $\delta(fg) = g \cdot \delta f + f \cdot \delta g$. This allows us to define the Lie bracket of vector fields $[X,Y] := XY - YX$ which one easily verifies to give a derivation when X and Y are derivations.

Definition 2.11. Let M be a smooth manifold and $X \in \mathfrak{X}(M)$. An integral curve of X with initial condition p is a curve $\gamma \colon (a,b) \to M$ such that $\gamma(0) = p$ and $\dot\gamma(t) = X(\gamma(t))$ for all $t \in (a,b)$.

To see how integral curves relate to more familiar notions of ordinary differential equations, we will look at their form in local coordinates.

Let X be a vector field on M and φ a chart on M. The equation $\dot\gamma(t) = X(\gamma(t))$ corresponds to $d\varphi(\dot\gamma(t)) = d\varphi(X(\gamma(t)))$ or $(\gamma(t), \frac{\mathrm{d}}{\mathrm{d}t}(\varphi \circ \gamma)(t)) = (\gamma(t), d_{\gamma(t)}\varphi X(\gamma(t)))$.

The condition on the first coordinate is satisfied trivially, so we can view this as a ordinary differential equation $\frac{\mathrm{d}}{\mathrm{d}t}(\varphi \circ \gamma)(t) = \tilde{X}(\gamma(t))$ for $\tilde{X}(p) = d_p\varphi X(p)$.

By the existence and uniqueness theorems for ordinary differential equations, for each point p there is some interval for which an integral curve with initial condition p exists and these depend smoothly on the choice of point p. This allows us to generate dynamics on all of M at once, using the following definition.

Definition 2.12. Let M be a smooth manifold and X a vector field on M. The flow of X is a map $\varphi_t\colon (a,b) \times M \to M$ such that $\varphi_t(p)$ is the integral curve of X with initial condition p.

The existence and uniqueness for ordinary differential equations guarantees the existence of a flow map and that this map will be a local diffeomorphism from M to itself.

Mechanically, one can think of a vector field as defining the velocity of some flow at every point in the manifold in a smooth manner. This will be the view point that will be the most helpful as we continue.

2.3 Differential Forms

Now that we have developed some of the ideas behind vector fields and their flows, we can begin with the theory of differential forms. Differential forms allow us to generalize oriented areas and volumes to the manifold setting and provide the natural setting for the generalized Stokes' theorem.

Definition 2.13. Let V be a finite dimensional $\mathbb{R}$-vector space. We denote $V^* = \{\lambda\colon V \to \mathbb{R} | \lambda \text{ is a linear map}\}$. This is called the dual space of V. V^* caries a natural vector space structure given by $(\lambda + \tau)(v) = \lambda(v) + \tau(v)$ and $(\alpha\lambda)(v) = \alpha(\lambda(v))$ for all $\lambda, \tau \in V^*$, $\alpha \in \mathbb{R}$, $v \in V$. One writes $\lambda(v) = \langle \lambda, v \rangle$ for any $\lambda \in V^*, v \in V$. We define the space of alternating k-linear maps on V, $\bigwedge^k V^*$ to be the space of k-linear maps $\mu\colon \prod_{i=1}^k V \to \mathbb{R}$ such that

$$\mu(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) = \operatorname{sgn}(\sigma)\mu(v_1, \ldots, v_k)$$

for all $v_1, \ldots, v_k \in V$ and $\sigma \in S_k$. We also define a $\bigwedge^0 V := \mathbb{R}$.

We see that $\bigwedge^k V^*$ carries an $\mathbb{R}$-vector space structure as a subspace of $\operatorname{Hom}(V^k, \mathbb{R})$, owing to the fact that

$$\begin{aligned}(\mu + \lambda)(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) &= \mu(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) + \lambda(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) \\ &= \operatorname{sgn}(\sigma)\mu(v_1, \ldots, v_k) + \operatorname{sgn}(\sigma)\lambda(v_1, \ldots, v_k)\end{aligned}$$

If we want to talk about all alternating forms, we will refer to $\bigwedge^\bullet V^* := \bigoplus_{j \in \mathbb{N}} \bigwedge^j V^*$.

Theorem 2.14. *Let V be a finite dimensional $\mathbb{R}$-vector space. There is a bilinear map $\wedge\colon \bigwedge^k V^* \times \bigwedge^l V^* \to \bigwedge^{k+l} V^*$ such that $\alpha \wedge \beta = (-1)^{kl}\beta \wedge \alpha$ and $(\alpha\wedge\beta)\wedge\gamma = \alpha \wedge (\beta\wedge\gamma)$. That is $\wedge$ makes $\bigwedge^\bullet V^*$ into a graded commutative algebra over $\mathbb{R}$.*

Proof. We begin by explicitly defining the wedge product. Given $\mu \in \bigwedge^k V^*$ and $\lambda \in \bigwedge^l V^*$ define

$$(\mu\wedge\lambda)(v_1,\dots,v_{k+l}) := \sum_{\sigma\in S_{k+l}} \frac{(k+l)!}{k!l!}\operatorname{sgn}(\sigma)\mu(v_{\sigma(1)},\dots,v_{\sigma(k)})\lambda(v_{\sigma(k+1)},\dots,v_{\sigma(k+l)})$$

We must first show that this results in a an element of $\bigwedge^{k+l} V^*$. We have

$$\begin{aligned}(\mu\wedge\lambda)(v_{\tau(1)},\dots,v_{\tau(k+l)}) :&= \sum_{\sigma\in S_{k+l}} \frac{1}{k!l!}\operatorname{sgn}(\sigma)\mu(v_{(\sigma(\tau(1))},\dots,v_{\sigma(\tau(k))})\lambda(v_{\sigma(\tau(k+1))},\dots v_{\sigma(\tau(k+l))})\\ &= \sum_{\sigma\in S_{k+l}} \frac{1}{k!l!}\operatorname{sgn}(\tau)\operatorname{sgn}(\sigma\tau)\mu(v_{(\sigma(\tau(1))},\dots,v_{\sigma(\tau(k))})\lambda(v_{\sigma(\tau(k+1))},\dots v_{\sigma(\tau(k+l))})\\ &= \operatorname{sgn}(\tau)\sum_{\eta\in S_{k+l}} \frac{1}{k!l!}\operatorname{sgn}(\eta)\mu(v_{\eta(1)},\dots,v_{\eta(k)})\lambda(v_{\eta(k+1)},\dots,v_{\eta(k+l)})\\ &= \operatorname{sgn}(\tau)(\mu\wedge\lambda)(v_1,\dots,v_{k+l})\end{aligned}$$

and we see that $\mu\wedge\lambda \in \bigwedge^{k+l} V^*$. To approach associativity, we note that the wedge product can be written as

$$\mu\wedge\lambda = \frac{1}{k!l!}A_{k+l}(\mu\otimes\lambda)$$

where $A_k(\kappa)(v_1,\dots v_k) = \sum_{\sigma\in S_{k+l}}\operatorname{sgn}(\sigma)\kappa(v_{\sigma(1)},\dots,v_{\sigma(k)})$. For $\alpha\in\bigwedge^a V^*, \beta\in\bigwedge^b V^*, \gamma\in\bigwedge^c V^*$ We then have

$$(\alpha\wedge\beta)\wedge\gamma = \frac{1}{a!b!}\frac{1}{(a+b)!c!}A_{a+b+c}(A_{b+c}(\alpha\otimes\beta)\otimes\gamma).$$

We see that

$$\begin{aligned}A_{a+b+c}(A_{b+c}(\alpha\otimes\beta)\otimes\gamma)(v_1,\dots,v_{a+b+c}) = &\sum_{\substack{\sigma\in S_{a+b}\\ \tau\in S_{a+b+c}}} \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)\alpha(v_{\tau(\sigma(1))},\dots v_{\tau(\sigma(a))})\\ &\beta(v_{\tau(\sigma(a+1))},\dots v_{\tau(\sigma(a+b))})\gamma(v_{\tau(a+b+1)},\dots,v_{\tau(a+b+c)})\end{aligned}$$

In permuting each term by $\tau\sigma^{-1}\tau^{-1}$ we have

$$\begin{aligned}A_{a+b+c}(A_{b+c}(\alpha\otimes\beta)\otimes\gamma)(v_1,\dots,v_{a+b+c}) = (a+b)! &\sum_{\tau\in S_{a+b+c}} \operatorname{sgn}(\tau)\alpha(v_{\tau(1)},\dots v_{\tau(a)})\\ &\beta(v_{\tau(a+1)},\dots,v_{\tau(a+b)})\gamma(v_{\tau(a+b+1)},\dots,v_{\tau(a+b+c)})\\ =\ &(a+b)!A_{a+b+c}(\alpha\otimes\beta\otimes\gamma)(v_1,\dots,v_{a+b+c})\end{aligned}$$

Then $(\alpha\wedge\beta)\wedge\gamma = \frac{1}{a!b!c!}A_{a+b+c}(\alpha\otimes\beta\otimes\gamma)$ and similarly $A_{a+b+c}(\alpha\otimes(A_{c+b}(\beta\otimes\gamma)) = (c+b)!A_{a+b+c}(\alpha\otimes\beta\otimes\gamma)$ and $(\alpha\wedge\beta)\wedge\gamma = \alpha\wedge(\beta\wedge\gamma)$. $\square$

Through a simple argument, one finds that $\bigwedge^j(\mathbb{R}^n)^*$ is given by $\mathrm{d}x^{i_1}\wedge\ldots\mathrm{d}x^{i_j}$ where $i_1 < i_2 < \ldots < i_j$ and the $\mathrm{d}x^k$ are defined by

$$\mathrm{d}x^k\,(e_j) := \delta_k^j$$

Example 2. To better understand what these represent, we consider an example when $V = \mathbb{R}^3$. Take $u, v, w \in \mathbb{R}^3$ with

$$u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, w = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}.$$

Since $\mathbb{R}^3$ has a basis (e_1, e_2, e_3) we get a dual basis $\mathrm{d}x^1\,, \mathrm{d}x^2\,, \mathrm{d}x^3$ with $\mathrm{d}x^k\,(e_j) = \delta_j^k$. If we take $\omega = \mathrm{d}x^1\wedge\mathrm{d}x^2\wedge\mathrm{d}x^3$ we see that

$$\begin{aligned}
\omega(u,v,w) &= A_3(\mathrm{d}x^1\otimes\mathrm{d}x^2\otimes\mathrm{d}x^3)(u,v,w) \\
&= (\mathrm{d}x^1\otimes\mathrm{d}x^2\otimes\mathrm{d}x^3 - \mathrm{d}x^1\otimes\mathrm{d}x^3\otimes\mathrm{d}x^2 \\
&+ \mathrm{d}x^2\otimes\mathrm{d}x^3\otimes\mathrm{d}x^1 - \mathrm{d}x^2\otimes\mathrm{d}x^1\otimes\mathrm{d}x^3 \\
&+ \mathrm{d}x^3\otimes\mathrm{d}x^1\otimes\mathrm{d}x^2 - \mathrm{d}x^3\otimes\mathrm{d}x^2\otimes\mathrm{d}x^1)(u,v,w) \\
&= u_1(v_2w_3 - v_3w_2) - u_2(v_1w_3 - v_3w_1) + u_3(v_1w_2 - v_2w_1)
\end{aligned}$$

Which is is the oriented volume of the parallelpiped spanned by u, v and w. This hints at the role of these k-forms as measuring oriented areas and volumes.

Definition 2.15. Let V and W be $\mathbb{R}$ vector spaces. Given a linear map $T\colon V\to W$ we can induce a map $T^*\colon \bigwedge^k W^* \to \bigwedge^k V^*$ by the formula

$$(T^*\mu)(w_1,\ldots,w_k) := \mu(T(w_1), T(w_2),\ldots,T(w_k))$$

Remark. The map $T\mapsto T^*$ is a generalization of the matrix transpose in the following sense. Given $v\in\mathbb{R}^n$ and an element $\lambda\in\mathbb{R}^{n*}$ we have $\langle\lambda, v\rangle = u^t v$ for a unique element $u\in\mathbb{R}^n$.

Given a map $T\colon\mathbb{R}^m\to\mathbb{R}^n$ we have

$$\begin{aligned}
\langle\lambda, Tw\rangle &= u^t(Tw) \\
&= u^tTw \\
&= (T^tu)^tw
\end{aligned}$$

If $\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$ is the isomorphism taking λ to u then

$$\begin{aligned}\langle T^*\lambda, w\rangle &= \langle \lambda, Tw\rangle \\ &= (\varphi\lambda)^t(Tw) \\ &= (T^t\varphi\lambda)^t w\end{aligned}$$

In the identification between $\mathbb{R}^n$ and $\mathbb{R}^{n*}$, we see that T^* has the same effect as T^t.

Using the k-form vector spaces and the dual maps of the isomorphisms define between T_pM and $\mathbb{R}^m$ we can create a bundle version of the k-forms. For $k \in \mathbb{N}$ we begin by defining the kth exterior bundle as the disjoint union of the exterior spaces, i.e. $\bigwedge^k T^*M := \bigsqcup_{p\in M} \bigwedge^k T_p^*M$. Using the dual map we can construct natural charts on $\bigwedge^k T^*M$. Given a chart $\varphi\colon U \to \mathbb{R}^m$ define the chart $d_k^*\varphi\colon \bigwedge^k T^*U \to \varphi(U) \times \bigwedge^k \mathbb{R}^m$ by $d_k^*\varphi(\mu_p) := (\varphi(p), ((d_p\varphi)^{-1})^*\mu_p)$. The arguments for a topology and smoothness coincide with the tangent bundle case so we easily get a vector bundle structure on $\bigwedge^k T^*M$.

Definition 2.16. We define $\Omega^k(M)$ the space of k-forms on M as the sections of the vector bundle $\pi\colon \bigwedge^k T^*M \to M$, i.e. smooth functions $\omega\colon M \to \bigwedge^k T^*M$ such that $\pi \circ \omega = \mathrm{Id}_M$.

Given a smooth map $f\colon M \to N$ we can induce the pullback of forms using the dual map.

Definition 2.17. Given $f\colon M \to N$ a smooth map between manifolds M and N, define the map $f^*\colon \Omega^k(N) \to \Omega^k(M)$ by $(f^*\omega)(p) = T_pf^*\omega$.

This allows us to move forms around using smooth maps.

Definition 2.18. We may define the wedge product $\wedge\colon \Omega^k(M)\times\Omega^l(M) \to \Omega^{k+l}(M)$ by $(\omega\wedge\eta)(p) = (\omega(p))\wedge(\eta(p))$, since each fiber of $\bigwedge^\bullet(T^*M)$ carries a wedge product as defined before. This makes $\Omega^\bullet(M) := \bigoplus_i \Omega^i(M)$ into a graded commutative algebra over $\mathbb{R}$ or a graded commutative ring over $C^\infty(M)$.

Definition 2.19. Given a form $\omega \in \Omega^k(M)$ and a vector field $X \in \mathfrak{X}(M)$ we can define the interior product $\iota_X\omega$ by

$$\iota_X\omega(X_1, \dots X_{k-1}) = \omega(X, X_1, \dots X_{k-1}),$$

for $X_1, \dots X_{k-1} \in \mathfrak{X}(M)$. The map ι_X is still a form since permutations of the remaining slots give a sign that agrees with the definition.

In coordinates, one can write any form ω as

$$\sum_{i_1<\ldots<i_k} \omega_{i_1\ldots i_k}(p)\,\mathrm{d}x^{i_1}\wedge\ldots\wedge\mathrm{d}x^{i_k}\,.$$

Given a function $f\colon M\to\mathbb{R}$ we can define a one form $\mathrm{d}f\in\Omega^1(M)$ by $\mathrm{d}f\,(p)([\gamma]) := \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} f\circ\gamma$. The form $\mathrm{d}f\,(p)$ essentially measures how quickly f changes from the point of view of a curve at p.

Proposition 2.20. *The map $f\mapsto \mathrm{d}f$ is well defined, linear and a derivation.*

Proof. We start by showing well definedness. Take $\gamma,\gamma'\in[\gamma]$ and (U,φ) a chart containing p. Then

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} f\circ\gamma &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} f\circ\varphi^{-1}(\varphi\circ\gamma)\\ &= D(f\circ\varphi^{-1})(\varphi(p))\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(\varphi\circ\gamma)\\ &= D(f\circ\varphi^{-1})(\varphi(p))\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(\varphi\circ\gamma')\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} f\circ\gamma'\end{aligned}$$

We now prove linearity. For $[\gamma],[\rho]\in T_pM$ we have

$$\begin{aligned}\mathrm{d}f\,([\gamma]+[\rho]) &= \mathrm{d}f((d_p\varphi)^{-1}(d_p\varphi[\gamma]+d_p\varphi[\rho])\\ &= \mathrm{d}f\left(\varphi^{-1}\left(\varphi(p)+t\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\varphi\circ\gamma(t)+t\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\varphi\circ\rho(t)\right)\right)\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} f\left(\varphi^{-1}\left(\varphi(p)+t\left(\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\varphi\circ\gamma+\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\varphi\circ\rho\right)\right)\right)\\ &= D(f\circ\varphi^{-1})(\varphi(p))\left(\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\varphi\circ\gamma+\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\varphi\circ\rho\right)\\ &= D(f\circ\varphi^{-1})(\varphi(p))\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\varphi\circ\gamma+D(f\circ\varphi^{-1})(\varphi(p))\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\varphi\circ\rho\\ &= \mathrm{d}f\,[\gamma]+\mathrm{d}f\,[\rho]\end{aligned}$$

with scalar multiplication following verbatim.

We see that $\mathrm{d}fg = g\,\mathrm{d}f + f\,\mathrm{d}g$ as follows. Given $[\gamma]\in T_pM$ we have

$$\begin{aligned}\mathrm{d}fg\,[\gamma] &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(fg)\circ\gamma\\ &= f(p)\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(g\circ\gamma)+g(p)\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(f\circ\gamma)\\ &= f(p)\,\mathrm{d}g\,(p)[\gamma]+g(p)\,\mathrm{d}f\,(p)[\gamma]\end{aligned}$$

□

Definition 2.21. Let $\mathrm{d}\colon \Omega^k(M) \to \Omega^{k+1}(M)$ be the map that in local coordinates has

$$d(\sum_{i_1 \dots i_k} \omega_{i_1 \dots i_k} \, \mathrm{d}x^{i_1} \wedge \dots \mathrm{d}x^{i_k}) := \sum_{i_1 \dots i_k} \mathrm{d}\omega_{i_1,\dots,i_k} \wedge \mathrm{d}x^{i_1} \dots \wedge \mathrm{d}x^{i_k}$$

A form ω is said to be closed if $\mathrm{d}\omega = 0$ and is said to be exact if $\omega = d\alpha$ for some form α. The map d is called the *exterior derivative.*

Proposition 2.22. $\mathrm{d}\colon \Omega^\bullet(M) \to \Omega^\bullet(M)$ *is a graded derivation with* $\mathrm{d}^2 = 0$*, that is:*

$$\mathrm{d}(\omega \wedge \eta) = \mathrm{d}\omega \wedge \eta + (-1)^k \omega \wedge \mathrm{d}\eta$$

with ω *and* η *taken to be* k *and* l*-forms respectively, and* $d(d\omega) = 0$ *for* $\omega \in \Omega^\bullet(M)$*.*

Proof. We start by proving the graded derivation property:

$$\begin{aligned}
\mathrm{d}\omega \wedge \eta &= \mathrm{d}\left(\sum_{\substack{i_1 < \dots < i_k \\ j_1 < \dots < j_l}} \omega_{i_1 \dots i_k} \, \mathrm{d}x^{i_1} \wedge \dots \wedge \mathrm{d}x^{i_k} \wedge \eta_{j_1 \dots j_l} \, \mathrm{d}x^{j_1} \wedge \dots \wedge \mathrm{d}x^{j_l} \right) \\
&= \mathrm{d}\left(\sum_{\substack{i_1 < \dots < i_k \\ j_1 < \dots < j_l}} \omega_{i_1 \dots i_k} \eta_{j_1 \dots j_l} \, \mathrm{d}x^{i_1} \wedge \dots \wedge \mathrm{d}x^{i_k} \wedge \mathrm{d}x^{j_1} \wedge \dots \wedge \mathrm{d}x^{j_l} \right) \\
&= \sum_{\substack{i_1 < \dots < i_k \\ j_1 < \dots < j_l}} \eta_{j_1 \dots j_l} \, \mathrm{d}\omega_{i_1 \dots i_k} \, \mathrm{d}x^{i_1} \wedge \dots \wedge \mathrm{d}x^{i_k} \wedge \mathrm{d}x^{j_1} \wedge \dots \wedge \mathrm{d}x^{j_l} \\
&+ \sum_{\substack{i_1 < \dots < i_k \\ j_1 < \dots < j_l}} \mathrm{d}\eta_{j_1 \dots j_l} \, \omega_{i_1 \dots i_k} \, \mathrm{d}x^{i_1} \wedge \dots \wedge \mathrm{d}x^{i_k} \wedge \mathrm{d}x^{j_1} \wedge \dots \wedge \mathrm{d}x^{j_l} \\
&= \sum_{\substack{i_1 < \dots < i_k \\ j_1 < \dots < j_l}} \mathrm{d}\omega_{i_1 \dots i_k} \, \mathrm{d}x^{i_1} \wedge \dots \wedge \mathrm{d}x^{i_k} \wedge \eta_{j_1 \dots j_l} \, \mathrm{d}x^{j_1} \wedge \dots \wedge \mathrm{d}x^{j_l} \\
&+ (-1)^k \sum_{\substack{i_1 < \dots < i_k \\ j_1 < \dots < j_l}} \omega_{i_1 \dots i_k} \, \mathrm{d}x^{i_1} \wedge \dots \wedge \mathrm{d}x^{i_k} \wedge \mathrm{d}\eta_{j_1 \dots j_l} \wedge \mathrm{d}x^{j_1} \wedge \dots \wedge \mathrm{d}x^{j_l} \\
&= \mathrm{d}\omega \wedge \eta + (-1)^k \omega \wedge \mathrm{d}\eta
\end{aligned}$$

To show $\mathrm{d}^2 = 0$ we start by seeing that

$$\begin{aligned}
\mathrm{d}\,\mathrm{d}\omega &= \mathrm{d} \sum_{i_1 < \dots < i_k} \mathrm{d}\omega_{i_1 \dots i_k} \, \mathrm{d}x^{i_1} \wedge \dots \wedge \mathrm{d}x^{i_k} \\
&= \sum_{i_1 < \dots < i_k} d^2 \omega_{i_1 \dots i_k} \, \mathrm{d}x^{i_1} \wedge \dots \wedge \mathrm{d}x^{i_k} + \sum \mathrm{d}\omega_{i_1 \dots i_k} \, \mathrm{d}(\mathrm{d}x^{i_1} \wedge \dots \wedge \mathrm{d}x^{i_k})
\end{aligned}$$

So it suffices to show that each of these terms vanishes. We can write $\mathrm{d}\big(\mathrm{d}x^{i_1} \wedge \ldots \mathrm{d}x^{i_k}\big) = \mathrm{d}(1) \wedge \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_k}$. We see that $\mathrm{d}(1) = 0$ since $\mathrm{d}(1)\,([\gamma]) = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} 1 \circ \gamma = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} 1 = 0$ for all $[\gamma] \in T_pM$. From this we conclude that our second term vanishes.

For the first term to vanish, it suffices to show that $\mathrm{d}^2 f = 0$ for any $f \in C^\infty(M)$. Our previous proof leads us to the equality $\mathrm{d}f\,[\gamma] = D(f \circ \varphi^{-1})(\varphi(p)) \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \varphi(\gamma(t))$ for a chart φ. We see that $D(f \circ \varphi^{-1})$ is the matrix of partial derivatives of $f \circ \varphi^{-1}$ and $\frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \varphi(\gamma(t)) = d_p\varphi[\gamma]$. If we write $[\gamma] = \sum v_i \frac{\partial}{\partial x^i}$ then the v_i are the coefficients of $d_p\varphi$ and $\mathrm{d}f\,([\gamma]) = \sum \frac{\partial f \circ \varphi}{\partial x^i} v_i$ meaning that $\mathrm{d}f\,(p) = \sum \frac{\partial f \circ \varphi^{-1}}{\partial x^i}\,\mathrm{d}x^i$. This means that

$$\begin{aligned}
\mathrm{d}^2 f &= \sum_{i,j} \frac{\partial^2 (f \circ \varphi^{-1})}{\partial x^j \partial x^i}\,\mathrm{d}x^j \wedge x^i \\
&= \sum_{j<i} \Big(\frac{\partial^2 (f \circ \varphi^{-1})}{\partial x^j \partial x^i} - \frac{\partial^2 (f \circ \varphi^{-1})}{\partial x^i \partial x^j}\Big)\,\mathrm{d}x^j \wedge \mathrm{d}x^i \\
&= 0
\end{aligned}$$

□

Given a smooth map $f\colon M \to N$ we have an induced map $T_pf\colon T_pM \to T_{f(p)}N$ meaning that there is a dual map $f_p^*\colon \bigwedge^\bullet T_{f(p)}N \to \bigwedge^\bullet T_pM$ This allows us to define a map $f^*\colon \Omega^\bullet(N) \to \Omega^\bullet(M)$ by $(f^*\mu)(p) = f_p^*\mu(f(p))$ for $\mu \in \Omega^k(N)$.

Given a vector field X with flow φ_t, we can pull a form ω back to another form $\varphi_t^*\omega$ on M. We can then see how ω changes along the flow of X, motivating the definition:

Definition 2.23. Given $\omega \in \Omega^k(M)$ and φ_t the flow of a vector field X, define the *Lie derivative* of ω with respect to X to be

$$\mathcal{L}_X\omega := \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \varphi_t^*\omega$$

This definition gives no indication as to how to calculate a Lie derivative, so the following proposition is quite helpful.

Proposition 2.24. *(Cartan's Magic Formula) Let X be a vector field on a smooth manifold M with φ_t the corresponding flow. Given $\omega \in \Omega^k(M)$ we have*

$$\mathcal{L}_X\omega = \mathrm{d}\iota_X\omega + \iota_X\,\mathrm{d}\omega$$

Proof. We start by calculating the pullback of a general map $f\colon M \to N$ in coordinates. Take $\omega \in \Omega^k(M)$ and φ, ψ charts containing $p \in M$ and $n \in N$ respectively.

We have $f^*\omega = f^* \sum \omega_{i_1 \dots i_k} \, \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_k} = \sum (\omega_{i_1 \dots i_k} \circ f) \, \mathrm{d}f_1^i \wedge \ldots \wedge \mathrm{d}f^{i_k}$. In the case where $f = \varphi_t$ we then have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \varphi_t^* \omega &= \left(\sum \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} (\omega_{i_1 \dots i_k} \circ \varphi_t) \right) \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_k} \\
&\quad + \sum \omega_{i_1 \dots i_k} \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} (\mathrm{d}f^{i_1} \wedge \ldots \mathrm{d}f^{i_k}) \\
&= \sum \iota_X \, \mathrm{d}\omega_{i_1 \dots i_k} \, \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_k} \\
&\quad + \sum \omega_{i_1 \dots i_k} \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} (\mathrm{d}f^{i_1} \wedge \ldots \mathrm{d}f^{i_k})
\end{aligned}$$

Seeing that $\frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \mathrm{d}f^{i_1} = \mathrm{d}X^i$ we then have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \varphi_t^* \omega &= \iota_X \, \mathrm{d}\omega + \sum_{i_1 \dots i_k} \omega_{i_1 \dots i_k} (-1)^j \, \mathrm{d}X^j \, \mathrm{d}x_1^i \wedge \ldots \wedge \mathrm{d}\hat{x}^j \wedge \ldots \wedge \mathrm{d}x^{i_k} \\
&= \iota_X \, \mathrm{d}\omega + \mathrm{d}\iota_X \omega
\end{aligned}$$

□

3 Symplectic Geometry

Definition 3.1. Let $\beta \in \Omega^2(M)$. β is said to be *non-degenerate* if given $X \in T_pM$, $\beta(X, Y) = 0$ for all $Y \in T_pM$ implies that $X = 0$.

This means that β establishes a musical isomorphism $\flat \colon T_pM \to T_pM^*$ given by $\flat(X) = \iota_X \beta$. This key property will allow us to relate a function on M, typically the energy of the system, to some vector field which will generate the time evolution of the system through the exterior derivative d.

Definition 3.2. A *symplectic manifold* (M, ω) is a manifold M, equipped with a closed, non-degenerate two-form ω, the so called symplectic form. Later, we will see that taking $\mathrm{d}\omega = 0$ will ensure that the Poisson bracket induced by ω satisfies that Jacobi identity.

One can think of ω as a way to measure oriented parallelograms in T_pM. This is computed by taking $\omega(u, v)$ where u, v are the vectors describing the vertices of the parallelogram. The nondegeneracy of ω means that $\omega^n = \omega \wedge \omega \wedge \cdots \wedge \omega$ is a non-zero $2n$ form, meaning that we have a canonical way to measure volumes in phase space. This allows one to prove one of the non-trivial facts of Hamiltonian mechanics, the Liouville theorem.

Definition 3.3. Let (M, ω) be a symplectic manifold and $H \colon M \to \mathbb{R}$ a smooth function on M. The Hamiltonian vector field $X_H \in \Gamma(TM)$ of H is the vector field such that $\iota_{X_H}\omega = \mathrm{d}H$. In other words $X_H = \sharp(\mathrm{d}H)$, where $\sharp \colon \Omega^1(M) \to \Gamma(TM)$ is the inverse of $\flat$ defined previously. In coordinates, we must have $\sharp(\flat(X)) = \omega^{jk}(\omega_{ij}X^i) = X^i$. In other words, $\omega_{ij}\omega^{jk} = \delta_i^k$.

Given a Hamiltionian H, the dynamics of a the system corresponding to this Hamiltionian are given by flowing along the Hamiltonian vector field X_H. In order to find the state of a system after some time t, we apply the diffeomorphism φ_t, the flow of the vector field X_H, to our initial condition $p \in M$.

To see the comparisons between symplectic mechanics and classical mechanics we use the following theorem

Theorem 3.4. *(Darboux) Let (M, ω) be a symplectic manifold. Given $p \in M$ there exists a coordinate neighborhood U of p such that in local coordinates $\omega = \sum_{i \le n} \mathrm{d}x^i \wedge \mathrm{d}x^{i+n}$ or if we take coordinates $(q^1, q^n, \dots, q^n, p^1, p^2, \dots p^n)$ then $\omega = \sum_{i \le n} \mathrm{d}q^i \wedge \mathrm{d}p^i$.*

This extremely deep theorem tells us that locally, all symplectic manifolds of the same dimension look the same.

We will now calculate what the equations of motion given by a symplectic form are in a chart of the form above. Given $H \colon M \to \mathbb{R}$ we have $\mathrm{d}H = \sum_{i \le n} \frac{\partial H}{\partial q^i}\,\mathrm{d}q^i + \frac{\partial H}{\partial p^i}\,\mathrm{d}p^i$. If we assume our vector field $X_H = \sum_{i \le n} X_q^i \frac{\partial}{\partial q^i} + X_p^i \frac{\partial}{\partial p^i}$ then

$$\iota_{X_H}\omega = \sum_{i \le n} X_q^i\,\mathrm{d}p^i - X_p^i\,\mathrm{d}q^i = \sum_{i \le n} \frac{\partial H}{\partial q^i}\,\mathrm{d}q^i + \frac{\partial H}{\partial p^i}\,\mathrm{d}p^i$$

so

$$X_q^i = \frac{\partial H}{\partial p^i} \quad \text{and} \quad X_p^i = -\frac{\partial H}{\partial q^j}$$

Since the flow is described by $\dot{q}^i = X_q^i$ and $\dot{p}^i = X_p^i$ we then have the equations

$$\dot{q}^i = \frac{\partial H}{\partial p^i} \quad \text{and} \quad \dot{p}^i = -\frac{\partial H}{\partial q^j}$$

Meaning that locally, the dynamics of any symplectic manifold satisfy Hamilton's equations.

Proposition 3.5. *Let (M, ω) be a symplectic manifold and $H \colon M \to \mathbb{R}$ a smooth function on M. The dynamics generated by X_H, the Hamiltonian vector field of H preserve ω, that is:*

$$\varphi_t^*\omega = \omega, \tag{3.0.1}$$

where φ_t is the flow of X_H.

Proof. To show this, it suffices to demonstrate that

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0}(\varphi_t^*\omega) = \mathcal{L}_{X_H}\omega = 0.$$

By Cartan's magic formula, we have:

$$\begin{aligned}\mathcal{L}_{X_H}(\omega) &= \mathrm{d}\iota_{X_H}\omega + \iota_{X_H}\mathrm{d}\omega \\ &= \mathrm{d}(\mathrm{d}H) + 0 \quad \text{(Definition of } X_H\text{)} \\ &= 0\end{aligned}$$

□

A diffeomorphism $\varphi\colon M \to N$, where M and N are symplectic, for which

$$\varphi^*\omega_N = \omega_M$$

is said to be symplectic or canonical. So we have shown that the flow of the Hamiltonian vector field is canonical.

Theorem 3.6. *(Liouville) Let N be a finite volume, open submanifold of the symplectic $2n$-manifold (M, ω), i.e. $\int_N \omega^n = C < \infty$. Let H be a smooth function on M and $\varphi_t\colon M \to M$ be the Hamiltonian flow corresponding to H. The volume of N is preserved along the flow φ_t, that is*

$$\int_N \omega^n = \int_{\varphi_t(N)} \omega^n$$

Proof. From the previously proved properties, we immediately see:

$$\int_{\varphi_t(N)} \omega^n = \int_N \varphi_t^*\omega^n = \int_N \omega^n \quad \square$$

Definition 3.7. Let M be a symplectic manifold with symplectic form ω. The Poisson bracket $\{\cdot,\cdot\}\colon C^\infty(M) \to C^\infty(M)$ is defined as follows: given $f, g \in C^\infty(M)$,

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

In coordinates, we see that

$$\{f, g\} = \sum_{i,j,k,l} \omega_{ij}\omega^{ki}\frac{\partial f}{\partial x^k}\omega^{lj}\frac{\partial g}{\partial x^l} = \sum_{j,k,l}\delta_j^k\omega^{lj}\frac{\partial f}{\partial x^k}\frac{\partial g}{\partial x^l} = \sum_{l,j}\omega^{lj}\frac{\partial f}{\partial x^j}\frac{\partial g}{\partial x^l}$$

One recognizes that this is $\mathcal{L}_{X_g}[f]$ So:

$$\{f, g\} = \mathcal{L}_{X_g}[f]$$

Proposition 3.8. *The Poisson bracket is a bilinear, antisymmetric, derivation which satisfies the Jacobi identity. i.e.: for all $f, g, h \in C^\infty(M)$ and all $\lambda, \eta \in \mathbb{R}$*

$$\{\lambda f + \eta g, h\} = \lambda\{f,h\} + \eta\{g,h\} \tag{\textbf{Lin.}}$$
$$\{f,g\} = -\{g,f\} \tag{\textbf{Antisymm.}}$$
$$\{fg,h\} = f\{g,h\} + g\{f,h\} \tag{\textbf{Deriv.}}$$
$$\{\{f,g\},h\} + \{\{h,f\},g\} + \{\{g,h\},f\} = 0. \tag{\textbf{Jacobi.}}$$

These properties make $C^\infty(M)$ into a Lie algebra under the Poisson bracket.

Proof. The first two properties follow directly from the fact that $X_{\lambda f + \eta g} = \sharp(\lambda\, \mathrm{d}f + \eta\, \mathrm{d}g) = \lambda \cdot \sharp(\mathrm{d}f) + \eta \cdot \sharp(\mathrm{d}g)$ and ω is linear and skew-symmetric.

For the derivation property, one sees that

$$X_{fg} = \sharp(\mathrm{d}(fg)) = \sharp(g\,\mathrm{d}f + f\,\mathrm{d}g) = g\sharp(\mathrm{d}f) + f\sharp(\mathrm{d}g)$$

by the Leibniz rule for the exterior derivative and linearity of $\sharp$. Substituting into our definition, one sees that

$$\begin{aligned}\{fg,h\} &= \omega(gX_f + fX_g, X_h)\\ &= g\omega(X_f, X_h) + f\omega(X_g, X_h)\\ &= g\{f,h\} + f\{g,h\}\end{aligned}$$

In proving the Jacobi Identity, we start by rewriting the Poisson brackets:

$$\begin{aligned}\{\{f,g\},h\} + \{\{h,f\},g\} + \{\{g,h\},f\} &=\\ &= \mathcal{L}_{X_h}[\{f,g\}] + \mathcal{L}_{X_g}[\{h,f\}] + \mathcal{L}_{X_f}[\{g,h\}]\\ &= \mathcal{L}_{X_h}\iota_{X_g}\iota_{X_f}\omega + \mathcal{L}_{X_g}\iota_{X_f}\iota_{X_h}\omega + \mathcal{L}_{X_f}\iota_{X_h}\iota_{X_g}\omega\\ &= \iota_{X_g}\iota_{X_f}\mathcal{L}_{X_h}\omega + \iota_{X_f}\iota_{X_h}\mathcal{L}_{X_g}\omega + \iota_{X_h}\iota_{X_g}\mathcal{L}_{X_f}\omega.\end{aligned}$$

Applying Cartan's magic formula we see that

$$\iota_{X_g}\iota_{X_f}\mathcal{L}_{X_h}\omega = \iota_{X_g}\iota_{X_f}(\mathrm{d}\iota_{X_h}\omega + \iota_{X_h}\mathrm{d}\omega) = \iota_{X_g}\iota_{X_f}\mathrm{d}\iota_{X_h}\omega$$

Then our previous formula simplifies to the following:

$$\iota_{X_g}\iota_{X_f}\mathrm{d}\iota_{X_h}\omega + \iota_{X_f}\iota_{X_h}\mathrm{d}\iota_{X_g}\omega + \iota_{X_h}\iota_{X_g}\mathrm{d}\iota_{X_f}\omega$$

Linearity of ι and d reduce this to:

$$\iota_{X_f+X_h+X_g}\iota_{X_f+X_h+X_g}\mathrm{d}\iota_{X_f+X_h+X_g}\omega = 0$$

Since $\iota_Y\iota_Y = 0$. □

The most important property of the Hamiltonian vector field is that its flow preserves the Hamiltonian and hence, energy is conserved in the dynamics generated.

Proposition 3.9. *Let (M, ω) be a symplectic manifold and $H : M \to \mathbb{R}$ a smooth function. Then*

$$\varphi_t^* H = H$$

where $\varphi_t : M \to M$ is the flow of the Hamiltonian vector field X_H.

Proof. Since the flow of a vector field preserves connected components of M, it suffices to show that

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\varphi_t^* H) = 0$$

By definition of the Lie Derivative, we see that:

$$\begin{aligned}\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\varphi_t^* H) &= \mathcal{L}_{X_H}(H) \\ &= \{H, H\} \\ &= 0\end{aligned}$$ □

This gives more testament to the fact that symplectic mechanics generalizes classical mechanics, since the key feature of time independent Hamiltonian mechanics is that the energy is conserved.

4 Cotangent Bundles

The quintessential example of a symplectic manifold is the cotangent bundle, T^*Q, of a given smooth manifold, Q. Physically, Q is the configuration space of a classical system, i.e. the manifold of allowable physical configurations. Elements of T^*Q correspond to the generalized coordinate positions and corresponding momenta of the system. Given the second order equations of motion typically seen in classical mechanics, the specification of a initial position and momentum is enough to determine the forward evolution of the given situation.

Example 3. Say we wanted to study the motion of a particle confined to a circle. Since the set of positions the particle may take is that of a circle, we take $Q = S^1$. In order to specify the momentum of the particle, we must specify a point in T^*S^1.

In order to apply the general theory of symplectic manifolds to cotangent bundles, we must first define a symplectic form on T^*Q.

Definition 4.1. Let Q be a smooth manifold and $\pi\colon T^*Q \to Q$ be the cotangent bundle of Q. The tautological one form is the unique one form $\Theta \in \Omega^1(T^*Q)$ satisfying the property:

$$\langle \Theta, v_{\alpha_q}\rangle = \langle \alpha_q, T_q\pi(v)\rangle$$

for all $v \in T_{\alpha_q}(T^*Q)$ and $q \in Q$.

Since we defined Θ for all $v \in T(T^*Q)$ it is the unique such one form.

Definition 4.2. The canonical 2-form on T^*Q is defined as $\omega = -\,\mathrm{d}\Theta$.

Proposition 4.3. *Let Q be a smooth manifold and $\pi\colon T^*Q \to Q$ be the cotangent bundle. Then (T^*Q, ω) is a symplectic manifold.*

Proof. In order to see that ω is closed and non-degenerate, we will begin by computing Θ in local coordinates. Let (U, φ) be a chart containing $q \in Q$ with coordinates (q^i). This then induces natural coordinates on T^*U as an open subset of T^*Q. We will label these coordinates (q^i, p_i). Given $v \in T_{\alpha_q}(T^*U)$ we can decompose v into its q and p components, i.e. $v = v^i \frac{\partial}{\partial q^i} + v^j \frac{\partial}{\partial p_j}$ for which $T_q\pi(v) = v^i \frac{\partial}{\partial q^i} \in T_qQ$. Taking $\alpha_q = \alpha_i\,\mathrm{d}q^i$ we have

$$\begin{aligned}\langle \Theta, v\rangle &= \langle \alpha_q, T_q\pi(v)\rangle \\ &= \langle \alpha_i\,\mathrm{d}q^i, v^i \frac{\partial}{\partial q^i}\rangle \\ &= \alpha_i v^i\end{aligned}$$

This means that $\Theta = p_i\,\mathrm{d}q^i$.

Taking the exterior derivative, we find that $\omega = -\mathrm{d}\Theta = \mathrm{d}q^i \wedge \mathrm{d}p_i$. This takes the same form as the canonical two-form on $\mathbb{R}^{2n}$, so ω is nondegenerate. Since the form is the sum of simple wedge products, we can also conclude that it is closed. $\square$

Now that we have shown that T^*Q is a symplectic manifold, we will see how to transform diffeomorphisms between manifolds into symplectic diffeomorphisms between cotangent bundles. This will allow us to more effectively study the symmetries of phase space, T^*Q, through the symmetries of Q.

Definition 4.4. Let Q and R be smooth manifolds and $F\colon Q \to R$ be a diffeomorphism. The *cotangent lift* of F, $T^*F\colon T^*R \to T^*Q$ is the smooth map defined as follows. Given $v \in T_qQ$ and $\alpha \in T_{F(q)}R$,

$$\langle T^*F\alpha, v\rangle = \langle \alpha, T_{F(p)}F \cdot v\rangle$$

In order to show that a diffeomorphism preserves the tautological one form if and only if it is the cotangent lift of a diffeomorphism on the base space, we need to prove the following proposition:

Proposition 4.5. *Let $F_t\colon T^*N \to T^*N$ be the time dependent diffeomorphism $\alpha \mapsto \exp(t)\alpha$ and X_N be the vector field with flow F_t. Then*

$$\langle \Theta_N, X_N \rangle = 0, \quad \mathcal{L}_{X_N}\Theta_N = \Theta_N, \quad \iota_{X_N}\omega_N = -\Theta_N$$

These facts mean that X_N is a Liouville vector field.

Proof. Since F_t only translates along fibers, we can conclude that $T\pi \cdot X_N = 0$. From this we see that

$$\langle \Theta_N(\alpha_n), X_N(\alpha_n) \rangle = \langle \alpha_n, T\pi \cdot X_n(\alpha_n) \rangle = 0$$

In proving our second equation, we start by applying Cartan's magic formula. This tells us:

$$\mathcal{L}_{X_N}\Theta_N = \mathrm{d}\iota_{X_N}\Theta_N + \iota_{X_N}\mathrm{d}\Theta_N.$$

The first equation of our claim tells us that the first term of this sum vanishes. All that remains in proving the second equation is to prove the last.

In coordinates, F_t takes the form $F_t(q_i, p^i) = (q_i, e^t p^i)$. Differentiating with respect to t and taking $t = 0$, we see that $X_N = p^i \frac{\partial}{\partial p^i}$. Then:

$$\begin{aligned} i_{X_N}\omega_N &= \mathrm{d}q^i(p^j \frac{\partial}{\partial p^j})\mathrm{d}p^i - \mathrm{d}p^i(p^j \frac{\partial}{\partial p^j})\mathrm{d}q^i \\ &= 0 - \delta_i^j p_j \mathrm{d}q^i \\ &= -p_i \mathrm{d}q^i = -\Theta_N \end{aligned}$$

□

Proposition 4.6. *Let $\psi\colon T^*R \to T^*Q$ be a diffeomorphism. This diffeomorphism preserves the tautological one form, i.e. $\psi^*\Theta_Q = \Theta_R$, if and only if φ is the cotangent lift of some diffeomorphism $F\colon Q \to R$.*

Proof. Assume $\psi\colon T^*R \to T^*Q$ be a diffeomorphism such that $\psi^*\Theta_Q = \Theta_R$. First we will show that $\psi_* X_Q = X_R$. Since the Liouville vector field is uniquely determined by the properties in Proposition 4.5 it suffices to show the following:

$$\langle \Theta_Q, \psi_* X_R \rangle = 0, \quad \mathcal{L}_{\psi_* X_R}\Theta_Q = \Theta_Q, \quad \iota_{\psi_* X_R}\omega_Q = -\Theta_Q$$

For the first property, we see that:

$$\langle \Theta_Q, \psi_* X_R \rangle = \langle \psi^*\Theta_Q, X_R \rangle = \langle \Theta_R, X_R \rangle = 0$$

The second property is verified thusly:

$$\Theta_R = \mathcal{L}_{X_R}\Theta_R = \mathcal{L}_{X_R}\psi^*\Theta_Q = \psi^*\mathcal{L}_{\psi_* X_R}\Theta_Q,$$

and we conclude that $\mathcal{L}_{\psi_* X_R}\Theta_Q = \Theta_Q$.

We immediately obtain the third property through Cartan's formula:

$$\begin{aligned}\mathcal{L}_{\psi_* X_R}\Theta_Q &= \mathrm{d}\iota_{\psi_* X_R}\Theta_Q + \iota_{\psi_* X_R}\mathrm{d}\Theta_Q \\ &= 0 - \iota_{\psi_* X_R}\omega_Q.\end{aligned}$$

By our second property, we can say $\iota_{\psi_* X_R}\omega_Q = -\Theta_Q$.

Since $\psi_* X_R = X_Q$, we can conclude that $\psi(e^t\alpha) = \psi \circ F_t(\alpha) = e^t\psi(\alpha)$. Taking $t \to -\infty$ one sees that $\psi(0) = 0$. Taking $\iota\colon Q \to T^*Q$ to be the inclusion of Q into the zero section of T^*Q (a diffeomorphism onto the image), we may now conclude that $\varphi = \pi_R \circ \psi \circ \iota\colon Q \to R$ is a diffeomorphism. To complete this argument we must demonstrate that $T^*\varphi = \psi$. We start by applying the definition of the cotangent lift, the tautological one form, and the chain rule to arrive at the conclusion. Let v be an element of T_qQ and $\alpha \in T_{\varphi(q)}R$

$$\begin{aligned}\langle T^*\varphi \cdot \alpha, v\rangle &= \langle \alpha, T_q\varphi \cdot v\rangle \\ &= \langle \alpha, T_q(\pi_r \circ \psi \circ \iota) \cdot v\rangle = \langle \alpha, T_{\psi\circ\iota(q)}\pi_R \cdot T_q(\psi \circ \iota) \cdot v\rangle \\ &= \langle \Theta_R(\alpha), T_{\iota(q)}\psi \cdot (T_q\iota \cdot v)\rangle \\ &= \langle \psi^*\Theta_R(\psi(\alpha)), T_q\iota \cdot v\rangle = \langle \Theta_Q(\psi(\alpha)), T_q\iota \cdot v\rangle \\ &= \langle \psi(\alpha), T_{\pi(\iota(q))}\pi_Q \cdot (T_q\iota \cdot v)\rangle = \langle \psi(\alpha), T_q(\pi_q \circ \iota) \cdot v\rangle \\ &= \langle \psi(\alpha), v\rangle\end{aligned}$$

Now to prove the converse, assume that $\varphi\colon Q \to R$ is a diffeomorphism and $\psi = T^*\varphi\colon T^*R \to T^*Q$. Let α_q be an element of T_q^*Q and $v \in T_{\psi(\alpha_q)}T^*R$. We then prove the identity by applying the definition of the pullback, the cotangent lift and the product rule.

$$\begin{aligned}\langle (T^*\varphi)^*\Theta_Q, v\rangle &= \langle \Theta_Q, T_{\psi(\alpha_q)}(T^*\varphi) \cdot v\rangle \\ &= \langle \psi(\alpha_q), T_{\alpha_q}\pi_Q(T_{\psi(\alpha_q)}(T^*\varphi) \cdot v\rangle \\ &= \langle \psi(\alpha_q), T_{\psi(\alpha_q)}(\pi_Q \circ T^*\varphi) \cdot v\rangle\end{aligned}$$

One can see that $T^*\varphi\big|_{T_r^*R}\colon T_r^*R \to T^*_{\varphi^{-1}(r)}Q$ So $\pi_Q \circ T^*\varphi(\alpha_r) = \varphi^{-1}(q)$. This tells us that the following diagram commutes, i.e. $\pi_Q \circ T^*\varphi = \varphi^{-1} \circ \pi_R$ (and that

$(\varphi^{-1}, T^*\varphi)$ is a bundle map):

$$\begin{tikzcd} T^*Q \arrow[d, "\pi_Q"'] & T^*R \arrow[l, "T^*\varphi"'] \arrow[d, "\pi_R"] \\ Q \arrow[r, "\varphi^{-1}"'] & R \end{tikzcd}$$

From this we conclude

$$\begin{aligned} \langle \psi(\alpha_q), T_{\psi(\alpha_q)}(\pi_Q \circ T^*\varphi) \cdot v\rangle &= \langle \psi(\alpha_q), T_{\psi(\alpha_q)}(\varphi{-}1 \circ \pi_R) \cdot v\rangle \\ &= \langle \alpha_q, T_{\psi(\alpha_q)}(\varphi^{-1} \circ \varphi \circ \pi_R) \cdot v\rangle \\ &= \langle \alpha_q, T_{\psi(\alpha_q)}\pi_R \cdot v\rangle = \langle \Theta_Q, v\rangle \end{aligned}$$

□

Given $T^*\varphi\colon T^*R \to T^*Q$ we can see that such a map is canonical, i.e. it preserves the symplectic form. Since the exterior derivative commutes with the pullback, we have

$$(T^*\varphi)^*\omega_Q = -(T^*\varphi)^*\mathrm{d}\Theta_Q = -\,\mathrm{d}((T^*\varphi)^*\Theta_Q) = -\mathrm{d}\Theta_R = \omega_R.$$

This gives us a convenient way to convert between the dynamics occurring on diffeomorphic configuration spaces. By taking the cotangent lift of the dynamics on T^*R we obtain "the same" dynamics on T^*Q. The method applied in the preceding proof were inspired by the approach in [5].

5 Lie Groups and Group actions

Now that we have described the common method for generating the dynamics of a system using symplectic geometry, we will now look at a systematic way to investigate the symmetries of a system. In order to describe smooth symmetries of manifolds we must first describe what a set of smooth symmetries aught to look like. In describing general symmetries, we use the language of group theory. This allows us to speak abstractly about the most important properties a set of symmetries has, i.e. the law of composition, the existence of the identity and inverses. Since we are focusing on smooth manifolds, our set of symmetries should be smooth, motivating the following definition.

Definition 5.1. A *Lie group* G is a smooth manifold with a smooth group structure. That is $\mu\colon G \times G \to G, (g, h) \mapsto g \cdot h$ and $i\colon G \to G, g \mapsto g^{-1}$ are smooth maps. This effectively defines a smooth set of symmetries.

Definition 5.2. By the definition of μ, we can define an action of G on itself, by taking $L_g \colon G \to G, h \mapsto gh$. Similarly, we define a right action R_g by $h \mapsto hg$.

Now that we have restricted ourselves to smooth symmetries, we gain a new tool for describing these symmetries, the tangent space and vector fields.

Definition 5.3. A vector field $X \in \Gamma(TG)$ is said to be left invariant if $L_{g*}X = X$. These are vector fields that are unchanged by the left action of G on itself and are a powerful way to study the properties of a Lie group.

Proposition 5.4. *Let G be a Lie group and $\Gamma_G(TG)$ the set of left invariant vector fields. $\Gamma_G(TG)$ has a natural vector space structure on it, inherited from $\Gamma(TG)$. Then $\Gamma_G(TG)$ is isomorphic as a vector space to $T_eG = \mathfrak{g}$, the tangent space to the identity.*

Proof. Given $\xi \in \mathfrak{g}$, we define a vector field at every $g \in G$ by $X_\xi(g) = T_eL_g(\xi)(g)$. We see that X_ξ is left invariant, since $T_gL_h(X_\xi(g)) = T_gL_h \cdot T_eL_g(\xi) = T_e(L_h \circ L_g)(\xi) = T_eL_{hg}(\xi) = X_\xi(hg)$. We can see that such a map assigns a unique smooth vector field for every $\xi \in \mathfrak{g}$.

Now, assume that Y is a left invariant vector field on G. If $\eta = Y(e) \in \mathfrak{g}$, then given $h \in G$, $Y(h) = T_eL_h(Y(e)) = T_eL_h(\eta) = X_\eta(h)$, this means that the map $\eta \mapsto X_\eta$ is a surjection from $\mathfrak{g}$ to L so we can conclude that $\mathfrak{g} \cong \Gamma_G(TG)$. □

One property that distinguishes $\mathfrak{g}$ from the tangent space of a generic manifold is that we can imbue it with a Lie algebra structure by way of the left invariant vector fields. To do so, we must first prove the following proposition:

Proposition 5.5. *Let G be a Lie group. $\Gamma_G(TG)$ is a Lie subalgebra of $\Gamma(TG)$.*

Proof. Since $\Gamma_G(TG)$ is already a vector subspace of $\Gamma(TG)$, it suffices to show that $\Gamma_G(TG)$ is closed under the Jacobi-Lie bracket. Take $X, Y \in \Gamma_G(TG)$. Since the pushforward commutes with the Jacobi-Lie bracket, we have

$$L_{g*}[X,Y] = \left[L_{g*}X, L_{g*}Y\right] = [X,Y]$$

hence $[X,Y] \in \Gamma_G(TG)$ and $\Gamma_G(TG)$ is a Lie subalgebra of $\Gamma(TG)$. □

We can give the Lie algebra structure on $\Gamma_G(TG)$ to $\mathfrak{g}$ by evaluation at the identity.

Definition 5.6. Let G be a Lie group and $\mathfrak{g} = T_eG$. We say $(\mathfrak{g}, [\cdot,\cdot])$ is the Lie algebra of G where $[\cdot,\cdot]$ is defined as follows. Given $\xi, \eta \in \mathfrak{g}$,

$$[\xi, \eta] := [X_\xi, X_\eta](e).$$

One easily checks that this bracket gives $\mathfrak{g}$ a Lie algebra structure.

Using the left invariant vector fields, we can define the following map from $\mathfrak{g}$ into G.

Definition 5.7. We define the exponential map $\exp\colon \mathfrak{g} \to G$ as the solution to the equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(g \cdot \exp(t\xi)) = X_\xi(g)$$

meaning that $\exp(t\xi)$ is the flow of the left invariant vector field X_ξ.

The theory of ordinary differential equations guarantees the existence of $\exp(t\xi)$ for t small enough and one can prove existence for all t by left multiplication.

Example 4. The most prevalent examples of Lie groups in physics and mathematics are the matrix Lie groups. Our first example is the the simplest, representing the group of automorphisms on a vector space. Let $K \in \{\mathbb{R}, \mathbb{C}\}$ and $\mathrm{Mat}_n(K)$ be the set of $n \times n$ matrices. The general linear group is defined to be $\mathrm{GL}_n(K) = \{A \in \mathrm{Mat}_n(K) |\ \det(A) \neq 0\}$, the set of invertible linear transformations from K^n to itself. One can see that $\mathrm{GL}_n(K)$ is a smooth manifold, owing the the fact that $\det^{-1}(0)$ is a closed subset of $K^{n\times n}$ implies that $\mathrm{GL}_n(K) = K^{n\times n} \setminus \det^{-1}(0)$ is open and hence a submanifold of $K^{n\times n}$. This carries a group structure from regular matrix multiplication. Non-zero determinant means that each element has an inverse, the identity has determinant 1, and the set is closed under multiplication since given $A, B \in \mathrm{GL}_n(K)$ we have $\det(AB) = \det(A)\det(B) \neq 0$. Since matrix multiplication is a polynomial in the entries of multiplicants it is a smooth map, from this we can conclude that $\mathrm{GL}_n(K)$ is a Lie group. Because $\mathrm{GL}_n(K)$ is an open subset of $K^{n\times n}$ then $\mathfrak{gl}_n(K) = T_e\mathrm{GL}_n(K) \cong K^{n\times n}$. Examining our definition for the exponential map at the identity, we see that it satisfies $\frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0}\exp(t\xi) = \xi$ so the exponential map coincides with the matrix exponential for matrix Lie groups.

Since $\mathrm{GL}_n(K)$ is the largest set of invertible matrices in $\mathrm{Mat}_n(K)$, one can think of it as the group of symmetries of a generic n-dimensional vector space with no additional structure. Since we often imbue our vector spaces with extra structures, e.g. inner products, volume forms, symplectic forms, and the like, we often restrict ourselves to Lie subgroups of $\mathrm{GL}_n(K)$ that preserve these structures.

Orthogonal group. If we want to look at the symmetries of an inner product space, which we will take to be $\mathbb{R}^n$ with the standard dot product, we need to describe all of the linear maps that preserve the dot product. Given $u, v \in \mathbb{R}^n$, we can write $u \cdot v = u^T v$. Then a linear map $A \in \mathrm{Mat}_n(\mathbb{R})$ will preserve the dot product if for all $u, v \in \mathbb{R}^n$ we have $(Au)\cdot(Av) = u \cdot v$. This implies that $(Au)^T(Av) = u^T A^T A v = u^T v$. Since this must hold for every $u, v \in \mathbb{R}^n$, as a bilinear form $A^T A = I_n$. Taking the determinant of this equation and applying the fact that $\det A^T = \det A$, we see that all such matrices must satisfy $(\det A)^2 = 1$ and are therefore invertible. We then define the Orthogonal group as $\mathrm{O}(n) = \{A \in \mathrm{Mat}_n(\mathbb{R}) | A^T A = I_n\}$. This inherits the multiplication from $\mathrm{GL}_n(\mathbb{R})$. We can see that this set forms a subgroup of $\mathrm{GL}_n(\mathbb{R})$ as follows. Given the fact that $(AB)^T = B^T A^T$, we then see that if $A, B \in \mathrm{O}(n)$ we have $(AB)^T(AB) = B^T A^T AB = I_n$. Therefore $AB \in \mathrm{O}(n)$. Since these maps preserve the inner product, they preserve the length of all vectors. These form the group of rotations and reflections of $\mathbb{R}^n$. We typically restrict further to the group of proper rotations $\mathrm{SO}(n) = \{A \in \mathrm{O}(n) | \det A = 1\}$, the connected component containing I_n.

Since the restriction from $\mathrm{O}(n)$ to $\mathrm{SO}(n)$ corresponds to choosing the connected component of the identity, we can conclude that $\mathfrak{so}(n) = T_e\mathrm{SO}(n) = T_e\mathrm{O}(n) = \mathfrak{o}(n)$. To calculate the form of $\mathfrak{so}(n)$, we start by assuming $A = \exp(t\xi)$ for $A \in \mathrm{SO}(n)$ and some $\xi \in \mathfrak{so}(n)$. If we differentiate the expression: $I_n = A^T A = \exp(t\xi)^T \exp(t\xi)$ with respect to t we arrive at $0 = (\xi \exp(t\xi))^T \exp(t\xi) + \exp(t\xi)^T (\xi \exp(t\xi))$ by the Leibniz rule. Letting $t = 0$ we have $\xi^T + \xi = 0$. This means that $\mathfrak{so}(n)$ is the set of skew symmetric matrices of dimension n.

Special Linear Group. When a volume form on a vector space is an object of importance, the special linear group is the relevant Lie group to keep in mind. One recalls that a volume form on an oriented vector space, take it to be $\mathbb{R}^n$ for definiteness, is a non-zero n-form on that space. Taking the standard basis e_i for $\mathbb{R}^n$, we have the standard volume form constructed by wedging the dual basis vectors $\mathrm{d}x^i$ in order. This gives the volume form $\mu = \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \cdots \wedge \mathrm{d}x^n$, this means that the oriented volume of the unit n-cube is 1. We might reasonably define a volume preserving symmetry as a linear map $A \in \mathrm{Mat}_n(\mathbb{R})$ such that $\mu(Av_1, \cdots, Av_n) = \mu(v_1, \cdots, v_n)$ for all ordered sets of vectors (v_i). Using the fact that $\mu(v_1, \cdots, v_n) = \det\left(\begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}\right)$, one sees that our previous equation implies that $\det(AV) = \det(V)$ for all $V \in \mathrm{Mat}_n(\mathbb{R})$. This means that A must satisfy $\det A = 1$ in order to preserve μ. We then define the special orthogonal group as $\mathrm{SL}_n(\mathbb{R}) = \{A \in \mathrm{Mat}_n(\mathbb{R}) | \det A = 1\}$. We can easily show that $\mathrm{SL}_n(\mathbb{R})$ forms a Lie subgroup of $\mathrm{GL}_n(\mathbb{R})$ as follows. Let (A_n) be a convergent sequence of matrices with $\det A_n = 1$ for all $n \in \mathbb{N}$. Continuity of the determinant implies that $A = \lim A_n$ must also satisfy $\det A = 1$ and hence $A \in \mathrm{SL}_n(\mathbb{R})$. The closed-subgroup theorem then implies that $\mathrm{SL}_n(\mathbb{R})$ forms a Lie subgroup of $\mathrm{GL}_n(\mathbb{R})$.

Symplectic Group. Since much of the exposition seen previously relates to the properties of a symplectic form, i.e. a non-degenerate, closed, two-form, it is natural to ask what kinds of symmetries such a form has. As before, we will look at the symmetries of a vector space with the extra structure of a symplectic form. We define a symplectic vector space as follows.

Definition 5.8. Let V be a vector space and $\omega\colon V \times V \to \mathbb{R}$ an alternating bilinear form. (V, ω) is said to be a *symplectic vector space* if ω is nondegenerate.

The non-degenerative condition requires that V be even dimensional, since it is equivalent to $\omega^{\dim V/2}$ being a volume form. One can prove that symplectic vector spaces of the same dimension are isomorphic so we can restrict ourselves to the case $V = \mathbb{R}^{2n}, \omega = \epsilon^i \wedge \epsilon_i$ where ϵ^i is the dual basis element for $e^i, i \leq n$ and ϵ_i is the dual basis element for e_i, without loss of generality. Since any bilinear form can be represented as $v^t W u$ for some $W \in \mathrm{Mat}_{2n}(\mathbb{R})$, the relevant symmetry group will be that for which $A^t W A = W$ where W represents ω as a bilinear form. To find W, we start by letting $v = v^i e_i + v_i e^i$ and $w = w^i e_i + w_i e^i$, then

$$\begin{aligned}
\omega(v, w) &= \epsilon^j (v^i e_i + v_i e^i)\epsilon_j (w^i e_i + w_i e^i) - \epsilon^j (w^i e_i + w_i e^i)\epsilon_j (v^i e_i + v_i e^i) && (5.0.1)\\
&= v^i \delta_i^j w_i \delta_j^i - w^i \delta_i^j v_i \delta_j^i && (5.0.2)\\
&= v^j w_j - w^j v_j && (5.0.3)
\end{aligned}$$

One can see that this corresponds to the block matrix

$$W = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$$

We define $\mathrm{Sp}(2n, \mathbb{R}) = \{A \in \mathrm{Mat}_{2n}(\mathbb{R}) | A^T W A = W\}$.

In order to talk about the symmetries of mechanical systems, we must establish a concrete way of talking about how symmetries manifest themselves on smooth manifolds. This is done through the language of group actions.

Definition 5.9. Let G be a Lie group and M a smooth manifold. A map $\Psi\colon G \times M \to M$, $m \mapsto \Psi_g(m)$ is said to be a *left Lie group action* if it is smooth, $\Psi_e(m) = m$ for all $m \in M$ and $\Psi_h \circ \Psi_g = \Psi_{hg}$ for all $h, g \in G$. Ψ is said to be a *right Lie group action* if it is smooth, $\Psi_e(m) = m$ and $\Psi_h \circ \Psi_g = \Psi_{gh}$. The criteria $\Psi_h \circ \Psi_g = \Psi_{hg}$ means that all group actions must be diffeomorphisms, since $(\Psi_g)^{-1}$ is explicitly given as $\Psi_{g^{-1}}$.

We can define a left action on a Lie group G by itself using conjugation. Define $\mathrm{Ad}\colon G \times G \to G$ by $\mathrm{Ad}_g(h) := ghg^{-1}$. This is called the adjoint action. Taking $h = \exp(t\xi)$ for $\xi \in \mathfrak{g}$ we can define an action of G on $\mathfrak{g}$ as follows: $\mathrm{Ad}_g(\xi) := \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} g \exp(t\xi) g^{-1}$.

One of the common threads of differential geometry is studying smooth objects by linear approximation, i.e. studying the tangent maps. Given a group action $\Psi\colon G \times M \to M$, we can induce a map from $\mathfrak{g}$ to $\Gamma(TM)$ by mapping $\xi \in \mathfrak{g}$ to the vector field whose flow corresponds to $\Psi_{\exp(t\xi)}$, we will call this vector field ξ_M. This can be defined formally as:

$$\xi_M = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \Psi_{\exp(t\xi)}.$$

We will now prove an important proposition regarding the vector fields generated by a group action, also know as infinitesimal symmetries.

Proposition 5.10. *Let G be a Lie group with a left Lie group action on a smooth manifold M. The induced map $\mathfrak{g} \to \Gamma(TM)$ by $\xi \mapsto \xi_M$ is a Lie algebra anti-homomorphism, i.e.*

$$[\xi, \eta]_M = -[\xi_M, \eta_M]$$

for all $\xi, \eta \in \mathfrak{g}$.

To prove this, we must first prove a lemma regarding how the adjoint action of G on $\mathfrak{g}$ interacts with the Lie group action on M.

Lemma 5.11. *Given $\xi \in \mathfrak{g}$ and $g \in G$, the following identity holds:*

$$(\mathrm{Ad}_g \xi)_M = \Psi^*_{g^{-1}} \xi_M.$$

Proof. We start our proof from the definition of $(\mathrm{Ad}_g \xi)_M$.

$$\begin{aligned}
(\mathrm{Ad}_g \xi)_M &= \left(\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \Psi_{\exp\left(\frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} g \exp(t\xi) g^{-1}\right)} \right) \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} (\Psi_{g \exp(t\xi) g^{-1}}) \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} (\Psi_g \circ \Psi_{\exp(t\xi)} \circ \Psi_{g^{-1}}) \\
&= T_{\Psi_{g^{-1}}} \Psi_g (\xi_M(\Psi_{g^{-1}})) \\
&= \Psi^*_{g^{-1}} \xi_M
\end{aligned}$$

□

Proof. One then sees the proof of our original proposition by taking $g = \exp(t\eta)$ and differentiating with respect to t.

$$[\xi, \eta]_M := \left(\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \mathrm{Ad}_{\exp(t\eta)} \xi \right)_M = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \Psi^*_{\exp(-t\eta)} \xi_M = -[\xi_M, \eta_M]$$

□

If we think of a Lie group action as a Lie group homomorphism $G \to \mathrm{Diff}(M) = \{\varphi\colon M \to M | \varphi \text{ is a diffeomorphism}\}$ then this induces a Lie algebra (anti)homomorphism $\mathfrak{g} \to \Gamma(TM)$ where $\Gamma(TM)$ is viewed as the Lie algebra of $\mathrm{Diff}(M)$, in the case where M is compact.

At this point, we can begin applying group actions to our previously mentioned symplectic manifolds. This will lead us towards the topic of moment maps, a generalized version of the Hamiltonian, and eventually to the ideas of Noether's theorem and Marsden-Weinstein reduction.

Definition 5.12. Let (M, ω) be a symplectic manifold and G be a Lie group that acts on M on the left by Ψ. The action Ψ is said to be symplectic if for each $g \in G$, the diffeomorphism $\Psi_g\colon M \to M$ is symplectic. That is

$$\Psi_g^* \omega = \omega.$$

Taking $g = \exp(t\xi)$ for $\xi \in \mathfrak{g}$ and differentiating with respect to t yields

$$\mathcal{L}_{\xi_M} \omega = 0.$$

By Cartan's magic formula,

$$\mathcal{L}_{\xi_M} \omega = \mathrm{d}\iota_{\xi_M} \omega + \iota_{\xi_M} \mathrm{d}\omega = \mathrm{d}\iota_{\xi_M} \omega$$

This means that locally $\xi_M = X_f$ for $f \in C^\infty(U)$ with $U \subseteq M$ open. One can view such an H as the generalized energy corresponding to the infinitesimal symmetry ξ. By the definition of X_f, f will be preserved along the flow of ξ_M where defined. We will soon specialize to the case where H is globally defined, this motivates the so called momentum map.

Definition 5.13. Let $\mathfrak{g}$ be a Lie algebra with a left action $\xi \mapsto \xi_M$ on a symplectic manifold M. A *Moment mapping* corresponding to the left action of $\mathfrak{g}$ is a map $J : M \to \mathfrak{g}^*$ such that $\xi_M = X_{J(\xi)}$, this of course means that $\iota_{\xi_M} \omega = \mathrm{d}J(\xi)$.

Just as regular linear momentum is the dual of the velocities, which generate translations when exponentiated, moment maps are the dual of Lie algebra elements, which generate general symmetries based on the specific group action. Of course, we can simply be given a moment map $J\colon M \to \mathfrak{g}^*$ and a group action $\Psi\colon G \times M \to M$, and we might like to know if these are compatible in some sense. This motivates the following definition.

Definition 5.14. Let (M, ω) be a symplectic manifold, and $\Psi\colon G \times M \to M$ a symplectic G-action. A moment map $J\colon M \to \mathfrak{g}^*$ is said to be G-equivariant if for each $g \in G$,

$$\Psi_g^* J = \mathrm{Ad}_{g^{-1}}^* J$$

that is $J(\Psi_g(x))(\xi) = \langle J(x), \mathrm{Ad}_g\xi\rangle$. This means that the following diagram commutes:

$$\begin{tikzcd} M \arrow[r, "J"] \arrow[d, "\Psi_g"'] & \mathfrak{g}^* \arrow[d, "\mathrm{Ad}^*_{g^{-1}}"] \\ M \arrow[r, "J"'] & \mathfrak{g}^* \end{tikzcd}$$

In the cotangent bundle case, one immediately has such a moment map. One can take $J(\xi) = \iota_{\xi_{T^*Q}}\Theta$ where Θ is the canonical one form mentioned previously.

Now that we have formulated the ideas of continuous symmetries and moment maps, we can now state Noether's theorem for symplectic mechanics.

Theorem 5.15. *(Noether) Let M be a symplectic manifold,G a Lie group with a left action $\Psi : G \times M \to M$ and a corresponding moment map $J : M \to \mathfrak{g}^*$. Assume $H \in C^\infty(M)$. If $\mathcal{L}_{\xi_M}(H) = 0$ then $\mathcal{L}_{X_H}(J(\xi)) = 0$.*

This means that if ξ_M is an infinitesimal symmetry of H, then $J(\xi)$ is conserved along the flow generated by H.

Proof. By the definition of J, we know that $X_{J(\xi)} = \xi_M$, then

$$0 = \mathcal{L}_{\xi_M}(H) = \mathcal{L}_{X_{J(\xi)}}(H) = \{H, J(\xi)\} = -\{J(\xi), H\} = -\mathcal{L}_{X_H}(J(\xi)) \quad \square$$

At this point, a physicist may be contented in calling J a conserved quantity, and "spending" this symmetry to reduce the degrees of freedom of the problem. This will lead to difficulty if care is not taken. In order to apply a reduction procedure and spend our symmetry, we must first show, under the proper assumptions, that identifying points under symmetry leaves us with a smooth manifold. To do so, we will state the quotient manifold theorem, which is instrumental in reduction.

Theorem 5.16. *(Quotient Manifold Theorem) Let M be a manifold and G be a compact lie group with a left action $\Psi : G \times M \to M$ that is proper and free; i.e. $G \times M \to M \times M$, $(g, x) \mapsto (\Psi_g x, x)$ is a proper map and $\Psi_g x = \Psi_h x$ implies that $g = h$. Then*

$$M/G = \{[x] \mid x \in M,\ x \sim y \ \text{ if } \ \Psi_g x = y \text{ for some } g \in G\}$$

is a smooth manifold and $\pi : M \to M/G$, $x \mapsto [x]$ is a smooth submsersion.

The proof of this result can be found in Abraham and Marsden [6]. This gives a concrete way to tell when the process of identifying points under symmetry gives a smooth manifold. Before we can describe how to reduce symmetries, we must state the following definition.

Definition 5.17. Let G be a Lie group with a group on a set X. Given $x \in X$, we define the isotropy subgroup G_x as follows:

$$G_x := \{g \in G | gx = x\}$$

This is instrumental in proving the following theorem, which tells us how to reduce a symplectic manifold under the action of a symmetry.

Theorem 5.18. *(Marsden-Weinstein Reduction Theorem) Let (M, ω) be a symplectic manifold, equipped with a G-action Ψ and an equivariant moment map J such that $\eta \in \mathfrak{g}^*$ is a regular value of J, i.e. T_pJ has full rank for all $p \in J^{-1}(\eta)$. Take*

$$\pi \colon J^{-1}(\eta) \to J^{-1}(\eta)/G_\eta$$

to be the natural projection $p \mapsto [p]$. If G_η acts properly and freely on $J^{-1}(\eta)$, there exists a symplectic form $\omega_{red} \in \Omega^2(M_{red})$ such that

$$\iota^*\omega = \pi^*\omega_{red}.$$

In simple terms, $J^{-1}(\eta)$ is the submanifold of M with moment map value η. This is the set of all phase points p such that $J(p) = \eta$. We require that η is a regular value of J to guarantee that $J^{-1}(\eta)$ is a regular submanifold of M. We then identify all points in the orbits under the action of G, since G is a set of symmetries of M and hence $J^{-1}(\eta)$, this means we are removing the redundant degrees of freedom associated to these symmetries. A proof of this theorem can be found in [4].

Given a G-invariant Hamiltonian $H \in C^\infty(M)$, i.e. $\Psi_g^*H = H$, we get a Hamiltonian on the reduced space $H_{red} \in C^\infty(M_{red})$ with $H\big|_{J^{-1}(\eta)} = H_{red} \circ \pi$. We then generate the dynamics on the reduced space as before, using the vector field $X_{H_{red}}$ such that $\iota_{X_{H_{red}}}\omega_{red} = \mathrm{d}H_{red}$.

6 The Classical Kepler Problem

We now will examine reduction of the classical two-body problem. Our configuration space will be $Q = \mathbb{R}^6 \backslash \{\mathbf{q}_1 = \mathbf{q}_2\}$. Q is the set of positions of the two bodies in question, where they are not permitted to collide. Then the phase space of the system is $T^*Q = \mathbb{R}^6 \times Q$ since Q is an open subset of $\mathbb{R}^6$. The Hamiltonian of this system is

$$H = \frac{\|\mathbf{p}_1\|^2}{2m_1} + \frac{\|\mathbf{p}_2\|^2}{2m_2} - \frac{k}{\|\mathbf{q}_1 - \mathbf{q}_2\|}.$$

This is the sum of the kinetic energies of the bodies plus the potential energy between them.

6.1 First Reduction

We begin with a symmetry that will correspond to the conservation of total linear momentum. One notes that if we map $(\mathbf{q}_1, \mathbf{q}_2, \mathbf{p}_1, \mathbf{p}_2)$ to $(\mathbf{q}_1 + \mathbf{x}, \mathbf{q}_2 + \mathbf{x}, \mathbf{p}_1, \mathbf{p}_2)$, the Hamiltonian remains unchanged. We will take this map to be a left $\mathbb{R}^3$ action on T^*Q. Differentiating the action tells us that the infinitesimal symmetry of this action is $X_\eta = \eta^i \frac{\partial}{\partial q_1^i} + \eta^j \frac{\partial}{\partial q_1^j}$ for η belonging to the Lie algebra of $\mathbb{R}^3$ canonically identified with $\mathbb{R}^3$ itself.

The moment map arising from this action must satisfy $\mathrm{d}(J(\eta)) = \iota_{X_\eta}\omega$. We see that

$$\begin{aligned} i_{X_\eta}\omega &= (\mathrm{d}q_1^i \wedge \mathrm{d}p_i^1 + \mathrm{d}q_2^j \wedge \mathrm{d}p_j^2)\left(\eta^k \frac{\partial}{\partial q_1^k} + \eta^l \frac{\partial}{\partial q_1^l}, \cdot\right) \\ &= \eta^k \,\mathrm{d}q_1^i \left(\frac{\partial}{\partial q_1^k}\right) \mathrm{d}p_i^1 + \eta^l \,\mathrm{d}q_2^j \left(\frac{\partial}{\partial q_2^l}\right) \mathrm{d}p_j^2 \\ &= \eta^k \delta_k^i \,\mathrm{d}p_i^1 + \eta^l \delta_l^j \,\mathrm{d}p_j^2 = \eta^i \,\mathrm{d}p_i^1 + \eta^j \,\mathrm{d}p_j^2 \end{aligned}$$

Then $\frac{\partial J(\eta)}{\partial p_i^j} = \eta^i$, $\frac{\partial J(\eta)}{\partial q_i^j} = 0$ so we can conclude that $J(\eta) = \eta^i p_i^1 + \eta^j p_j^2 = \eta \cdot (\mathbf{p}_1 + \mathbf{p}_2)$ and $\mathbf{J} = \mathbf{p}_1 + \mathbf{p}_2$, where we implicitly identify $\mathbb{R}^{3*}$ with $\mathbb{R}^3$. We can immediately see that this moment map is equivariant because $\mathbb{R}^3$ is an abelian group.

To start the reduction process, we look at the preimage of $0 \in \mathbb{R}^3$ under the moment map. We see that $\mathbf{J}^{-1}(0) = \{(\mathbf{q}_1, \mathbf{q}_2, \mathbf{p}_1, \mathbf{p}_2) \in T^*Q \mid \mathbf{p}_1 + \mathbf{p}_2 = 0\}$. This is precisely the set where the total linear momentum of the system is zero. Now when we take the quotient by the group action, we make the identification $(\mathbf{q}_1, \mathbf{q}_2, \mathbf{p}_1, \mathbf{p}_2) \sim (\mathbf{q}_1 + x, \mathbf{q}_2 + x, \mathbf{p}_1, \mathbf{p}_2)$. To know that the result of the quotient is still a smooth manifold, we must show that our action is free and proper. We start by assuming that there is some $(\mathbf{q}, \mathbf{p}) \in T^*Q$ for which $\mathbf{x} \cdot (\mathbf{q}, \mathbf{p}) = \mathbf{y} \cdot (\mathbf{q}, \mathbf{p})$ for some $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$. This implies that $\mathbf{q} + \mathbf{x} = \mathbf{q} + \mathbf{y}$ and hence $\mathbf{x} = \mathbf{y}$, therefore our action is free. To show that this group action is proper it suffices to show that the action $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ by $(\mathbf{q}, \mathbf{x}) \mapsto \mathbf{q} + \mathbf{x}$ is a proper map.

Let K be a compact subset of $\mathbb{R}^3$ then for $\mathbf{x} + K$ to have non-trivial intersection with K, $\|x\| < \operatorname{diam}(K)$ since the bounding balls of K and $x+K$ must intersect for K and $x+K$ to intersect. this means that $\{x \in \mathbb{R}^3 | (x+K) \cap K \neq \emptyset\}$ is bounded and thus has compact closure. We can then conclude that our action is proper and $\mathbf{J}^{-1}(0)/\mathbb{R}^3$ is a smooth manifold. Since $\mathbf{J}^{-1}(0)$ has $\mathbf{p}_2 = -\mathbf{p}_1$ and $(\mathbf{q}_1 + x) - (\mathbf{q}_2 + x) = \mathbf{q}_1 - \mathbf{q}_2$ we can take our coordinates on $M_{red} = \mathbf{J}^{-1}(0)/\mathbb{R}^3 \cong T^*(\mathbb{R}^3 \backslash \{0\}) \cong (\mathbb{R}^3 \backslash \{0\}) \times \mathbb{R}^3$ to be

$$(\mathbf{q}, \mathbf{p}) = (\mathbf{q}_1 - \mathbf{q}_2, \mathbf{p}_1 - \mathbf{p}_2) = (\mathbf{q}_1 - \mathbf{q}_2, \mathbf{p}_1 + \mathbf{p}_1),$$

with $\mathbf{q} \neq 0$. Our Hamiltonian is invariant under the $\mathbb{R}^3$ action since $\|\mathbf{q}_1 + x - (\mathbf{q}_2 + x)\| =$

$\|\mathbf{q}_1 - \mathbf{q}_2\|$ and it descends to a Hamiltonian on the reduced space. Our new Hamiltonian is

$$H_{red} = \frac{\|\mathbf{p}\|^2}{2}\left(\frac{m_1+m_2}{m_1 m_2}\right) + \frac{k}{\|\mathbf{q}\|} = \frac{\|\mathbf{p}\|^2}{2\mu} + \frac{k}{\|\mathbf{q}\|}$$

6.2 Second Reduction

Now that we have reduced the phase space of the system to half the dimension, we can now examine the rotational invariance of the system. If we take the standard SO(3) action on $\mathbb{R}^3\backslash\{0\}$ by $\Psi_A(\mathbf{q}) = A\mathbf{q}$ we can then take the cotangent lift of this action. We start by noticing that $T\Psi_A(v) = Av$, then $\langle T^*\Psi_A \mathbf{p}, v\rangle = \mathbf{p}^T A^{-1} v = \mathbf{p}^T A^T v = (A\mathbf{p})^T v$ and hence $T^*\Psi_A(\mathbf{q},\mathbf{p}) = (A\mathbf{q}, A\mathbf{p})$. Since SO(3) does not affect the norm of $\mathbf{q}$ or $\mathbf{p}$ we see that H_{red} is invariant under this action.

In order to find the moment map corresponding to the $SO(3)$ action we have to find the vector field corresponding to $\xi \in \mathfrak{so}(3)$. Taking $A = \exp(t\xi)$ we have

$$\begin{aligned} X_\xi &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} T^*\Psi_{\exp(t\xi)}(\mathbf{q},\mathbf{p}) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\exp(t\xi)\mathbf{q}, \exp(t\xi)\mathbf{p}) \\ &= (\xi\mathbf{q}, \xi\mathbf{p}) \\ &= q^j \xi^i_j \frac{\partial}{\partial q^i} + p^l \xi^k_l \frac{\partial}{\partial p^k}. \end{aligned}$$

Now,

$$\begin{aligned} \iota_{X_\xi}\omega &= \mathrm{d}q^m \wedge \mathrm{d}p_m \,(X_\xi) \\ &= q^j \xi^i_j \delta^m_i \,\mathrm{d}p_m - p^l \xi^k_l \delta^m_k \,\mathrm{d}q_m \\ &= q^j \xi^i_j \,\mathrm{d}p_i - p^l \xi^k_l \,\mathrm{d}q_k \end{aligned}$$

This implies that $\frac{\partial \mu(\xi)}{\partial p_i} = q^j \xi^i_j$ and we can conclude that $\mu(\xi) = \xi^i_j q^j p_i = \mathbf{p}\cdot(\xi\mathbf{q})$. Identifying $\mathfrak{so}(3)$ with $\mathbb{R}^3$ by

$$\begin{bmatrix} 0 & -\xi_3 & \xi_2 \\ \xi_3 & 0 & -\xi_1 \\ -\xi_2 & \xi_1 & 0 \end{bmatrix} \mapsto \tilde{\xi} = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix}$$

and noting that

$$\begin{aligned}
\xi\mathbf{q} &= \begin{bmatrix} 0 & -\xi_3 & \xi_2 \\ \xi_3 & 0 & -\xi_1 \\ -\xi_2 & \xi_1 & 0 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} \\
&= \begin{bmatrix} \xi_2 q_3 - \xi_3 q_2 \\ \xi_3 q_1 - \xi_1 q_3 \\ \xi_1 q_2 - \xi_2 q_1 \end{bmatrix} = \tilde{\xi} \times \mathbf{q}
\end{aligned}$$

Then $\mathbf{p} \cdot (\xi\mathbf{q}) = \mathbf{p} \cdot (\tilde{\xi} \times \mathbf{q}) = \tilde{\xi} \cdot (\mathbf{q} \times \mathbf{p})$. We then conclude that $\mathbf{J} = \mathbf{q} \times \mathbf{p}$ under our identification, this is the angular momentum of the system.

Now we will show that the moment map is equivariant. We have

$$\begin{aligned}
\langle \mu \circ T^*\Psi_A(\mathbf{q}, \mathbf{p}), \xi \rangle &= \xi \cdot (A\mathbf{q} \times A\mathbf{p}) \\
&= \xi \cdot A(\mathbf{q} \times \mathbf{p}) = (A^T \xi) \cdot (\mathbf{q} \times \mathbf{p}) \\
&= \langle \mathrm{Ad}^*_{A^{-1}}(\mu(\mathbf{q}, \mathbf{p})), \xi \rangle
\end{aligned}$$

Since we have shown that we have an equivariant moment map arising from a symplectic SO(3) action, we can begin with the reduction procedure.

Instead of taking $\mu^{-1}(0)$ to be the manifold we will quotient, we now take $\mu^{-1}(L)$,where $L \in \mathfrak{so}(3)^* \cong \mathfrak{so}(3)$ is the angular momentum of the system, this means that every point in our reduced phase space will satisfy $\mathbf{q} \times \mathbf{p} = L \neq 0$. We must first examine the form of $\mathrm{SO}(3)_L$. We can imediately observe that $\mathrm{SO}(3)_L \cong \mathrm{SO}(2)$, since every rotation that fixes L corresponds to a rotation of the plane perpendicular two L. Now we must show a few properties of the group action in order to fulfill the assumptions of the Marsden-Weinstein Reduction theorem.

We automatically know that our SO(3) action is proper since SO(3) is compact. Therefore, to show that $\mu^{-1}(L)/\mathrm{SO}(3)_L$ is a smooth manifold, we only need to demonstrate that our action is free. Because $\mathbf{q} \times \mathbf{p} \neq 0$, $\mathbf{q}$ and $\mathbf{p}$ are linearly independent and neither is parallel to L. We then see that there is no transformation that fixes both $\mathbf{p}$ and $\mathbf{q}$. This tells us the chosen action is free and $M_{red'} = \mu^{-1}(L)/\mathrm{SO}(3)_L \cong T^*\mathbb{R}^+ \cong \mathbb{R}^+ \times \mathbb{R}$ is a smooth manifold.

We will now choose coordinates on $M_{red'}$, our first coordinate being $r = \|\mathbf{q}\|$. Since $\mathbf{q} \times \mathbf{p}$ is constant, $\mathbf{p}$ is uniquely determined by its component parallel to $\mathbf{q}$.

Letting $\mathbf{p} = \mathbf{p}^{\perp} + \mathbf{p}^{\parallel}$, we then have $L = \mathbf{q} \times (\mathbf{p}^{\perp} + \mathbf{p}^{\parallel}) = \mathbf{q} \times \mathbf{p}^{\perp}$ and

$$\begin{aligned}\frac{\|\mathbf{p}\|^2}{2m} &= \frac{(\mathbf{p}^{\perp} + \mathbf{p}^{\parallel}) \cdot (\mathbf{p}^{\perp} + \mathbf{p}^{\parallel})}{2\mu} \\ &= \frac{\left\|\mathbf{p}^{\perp}\right\|^2}{2\mu} + 2\frac{\mathbf{p}^{\perp} \cdot \mathbf{p}^{\parallel}}{2\mu} + \frac{\left\|\mathbf{p}^{\parallel}\right\|^2}{2\mu} \\ &= \frac{\left\|\mathbf{p}^{\perp}\right\|^2}{2\mu} + \frac{\left\|\mathbf{p}^{\parallel}\right\|^2}{2\mu}\end{aligned}$$

Letting $\mathbf{q} = r\hat{v}$ and $\mathbf{p}^{\perp} = p^{\perp}\hat{w}$ where then we have $L = rp^{\perp}(\hat{v} \times \hat{w}) = L$ This means that $r^2(p^{\perp})^2 = \|L\|^2$ and hence $\left\|\mathbf{p}^{\perp}\right\|^2 = \frac{\|L\|^2}{r^2}$. We now induce the coordinate p_r being the momentum associated to r which corresponds to $\pm\left\|\mathbf{p}^{\parallel}\right\|$, this can be written as $p_r = \frac{\mathbf{q} \cdot \mathbf{p}}{\|\mathbf{q}\|}$.

Our previous reduced Hamiltonian now gives:

$$H_{red'} = \frac{\|L\|^2}{2\mu r^2} + \frac{p_r^2}{2\mu} - \frac{k}{r}$$

We see that

$$\mathrm{d}H_{red'} = \left(\frac{k}{r^2} - \frac{\|L\|^2}{\mu r^3}\right) \mathrm{d}r + \frac{p_r}{\mu}\,\mathrm{d}p_\mu$$

Let $X_{H'} = X_r \frac{\partial}{\partial r} + X_{p_r} \frac{\partial}{\partial p_r}$, then

$$\iota_{X_{H'}}\omega = X_r\,\mathrm{d}p - X_{p_r}\,\mathrm{d}r\,.$$

Taking $\mathrm{d}H'_{red} = \iota_{X_{H'}}\omega$ we have $X_r = \frac{p_r}{\mu}$ and $X_{p_r} = \frac{\|L\|^2}{\mu r^3} - \frac{k}{r^2}$. We then obtain the equations of motion:

$$\dot{r} = \frac{p_r}{\mu}, \quad \dot{p}_r = \frac{\|L\|^2}{\mu r^3} - \frac{k}{r^2}$$

7 Deformation Quantization

7.1 Motivation

In doing typical quantum mechanics, we begin with some classical phase space, typically $M \cong \mathbb{R}^n \times \mathbb{R}^{n*}$ with coordinates (q^i, p_i), and a Hamiltonian, a smooth function $H \colon \mathbb{R}^{2n} \to \mathbb{R}$ which is a low degree polynomial in p^i with coefficients in $C^\infty(\mathbb{R}^n)$. From here, we take our complex Hilbert space $(\mathcal{H}, \langle\cdot,\cdot\rangle) = L^2(\mathbb{R}^n)$, which describes the set of complex probability density functions on $\mathbb{R}^n$ with the usual inner product of functions, and we would like to find a representation of a suitable subspace of $C^\infty(M)$ containing H on our Hilbert space. To do this, we make a guess. We map p^i to the operator $-i\hbar\frac{\partial}{\partial q^i}$ and map q^i to the operator $\psi \mapsto q^i \cdot \psi$. We call the operator form of H, $\hat{H}$.

At this point, our guess appears to have little to do with the classical case. We make reference to a classical system, but apparently no reference to the geometry of phase space. This is not quite true, there is a hidden reference to the classical symplectic geometry outlined in the first half of this thesis. When we look at the commutation relations between the $\hat{q}^i$'s and $\hat{p}_j$'s we see a glimpse of the classical regime shining through. Let $\psi \in D(\hat{q}^i) \cap D(\hat{p}^j)$, then:

$$\begin{aligned}\left[\hat{q}^i, \hat{p}_j\right]\psi &= -i\hbar q^i \frac{\partial \psi}{\partial q^j} + i\hbar \frac{\partial}{\partial q^j}\left(q^i \psi\right) \\ &= -i\hbar q^i \frac{\partial \psi}{\partial q^j} + i\hbar q^i \frac{\partial \psi}{\partial q^j} + i\hbar \delta^i_j \psi \\ &= i\hbar \delta^i_j \psi\end{aligned}$$

Since functions commute with functions and derivatives commute by the Schwartz theorem, we have the following set of commutation relations:

$$\left[\hat{q}^i, \hat{p}_j\right] = i\hbar\delta^i_j, \quad \left[\hat{q}^i, \hat{q}^j\right] = 0, \quad [\hat{p}_i, \hat{p}_j] = 0.$$

Now, using the canonical Poisson bracket on $\mathbb{R}^n \times \mathbb{R}^{n*}$ we come to the following similar result:

$$\begin{aligned}\left\{q^i, p_j\right\} &= \sum_k \frac{\partial q^i}{\partial q_k}\frac{\partial p_j}{\partial p_k} - \frac{\partial q^i}{\partial p_k}\frac{\partial p_j}{\partial q^k} \\ &= \delta^i_k \delta^k_j = \delta^i_j\end{aligned}$$

and:

$$\left\{q^i, q^j\right\} = \sum_k \frac{\partial q^i}{\partial q^k}\frac{\partial q^j}{\partial p_k} - \frac{\partial q^i}{\partial p_k}\frac{\partial q^j}{\partial q_k} = \delta^i_k \cdot 0 - 0 \cdot \delta^j_k = 0$$

One similarly confirms that $\{p_i, p_j\} = 0$. Comparing these equations to the previous ones, we recognize that the quantum commutation relations are exactly $i\hbar$ times the Poisson commutation relations. If we wanted to generalize this scheme, we might like to take the Poisson algebra of observables and find a new associative product on it that satisfies $i\hbar$ times the Poisson commutation relations. It turns out that enforcing the commutation relations for all observables is an impossible task, even on a set of observables polynomial in q^i and p_j of bounded degree. This fact coming from the Gronewold-van Hove no go theorems. Instead of strictly enforcing these relations we will want our algebra to follow these relations “asymptotically”, a notion which will be the basis for deformation quantization.

7.2 Basic Definitions

Before we get to defining what a deformation is, we will describe the basic algebraic structure used to describe the asymptotics of deformation quantization.

Definition 7.1. Let R be a unital ring, $R[[\lambda]]$ is the ring of formal power series in λ with coefficients in R, i.e. formal linear combinations $r = \sum_{i=0}^{\infty} r_i \lambda^i$ with $r_i \in R$. The ring structure is given by

$$r \cdot s = (\sum r_i \lambda^i)(\sum s_j \lambda^j) = \sum r_i s_j \lambda^{i+j}$$

with addition defined as $r + s = \sum (r_i + s_i)\lambda^i$.

Definition 7.2. Let R be a unital ring and M a left R-module. Then $M[[\lambda]]$ is the left $R[[\lambda]]$ module of formal power series in λ with coefficients in M. We define the module structure analogously to the ring structure on $R[[\lambda]]$.

We commit an abuse of notation and identify R and M with R_0 and M_0, the subring/module of $R[[\lambda]]$ and $M[[\lambda]]$ of zeroth order. We then define a formal deformation quantization as follows:

Definition 7.3. Let A be a Poisson algebra with bracket $\{\cdot,\cdot\}$, i.e. a commutative algebra over $\mathbb{C}$, $(A, \cdot, +)$, where $\{\cdot,\cdot\}$ satisfies the axioms from Proposition 3.8. The bilinear map $\star\colon A[[\lambda]] \times A[[\lambda]] \to A[[\lambda]]$ is said to be a star product if it satisfies the following: assume $f, g, h \in A$

$$f \star g = f \cdot g + o(\lambda), \tag{7.2.1}$$

$$f \star g - g \star f = i\lambda\{f, g\} + o(\lambda^2), \tag{7.2.2}$$

$$(f \star g) \star h = f \star (g \star h). \tag{7.2.3}$$

These can be extended billinearly to all of $A[[\lambda]]$. Of the conditions, Equations 7.2.1 and 7.2.2 are the modified Dirac quantization conditions and Equation 7.2.3 ensures that $(A[[\lambda]], \star)$ is an associative algebra. We now see how to enforce the original Dirac quantization rules asymptotically. We require the $\star$-commutator to agree with $i\lambda\{\cdot,\cdot\}$ up to order λ^2. As a matter of convention one can take $\lambda = \hbar$ so that the commutation relation matches with Dirac's original condition. $\hbar$ is nothing more than a formal parameter within the scope of deformation quantization.

Since the typical way we find said star products is through differential operators on some domain, we define the following:

Definition 7.4. Let $A \subset C^\infty(U)$ a sub-algebra, where U is an open subset of some smooth manifold M. We define $\mathrm{Diff}(A)$ to be the set of differential operators on A, i.e. for each $a \in \mathrm{Diff}(A)$ and $x \in U$

$$(af)\Big|_V = \sum_{|\alpha|\leq l} a_\alpha \partial^\alpha f$$

for some V an open neighborhood of x, $a_\alpha \in C^\infty(U)$.

7.3 Examples of Star Products

It is easy to map polynomials purely in p or q into differential operators, since $\left[\hat{q}^i, \hat{q}^j\right] = \left[\hat{p}^i, \hat{p}^j\right] = 0$ we can map $\prod (q^j)^{\alpha_j}$ to $(\hat{q}^1)^{\alpha_1} \circ \ldots \circ (\hat{q}^n)^{\alpha_n}$ and $\prod (p_j)^{\beta_j}$ to $(\hat{p}_1)^{\beta_1} \circ \ldots \circ (\hat{p}_n)^{\beta_n}$ with no issues.

We will construct a prototypical example on $\mathbb{C}[q,p]$, the space of complex polynomials in q and p, and then see how this might be extended to $\mathbb{C}[q^i, p_i]$. We will begin by constructing a star product $\star$ that extends the typical quantization $q \mapsto \hat{q}$, $\psi \mapsto q\psi$; $p \mapsto -i\hbar \frac{\mathrm{d}}{\mathrm{d}q}$.

Naively, one might map polynomials into the differential operators as follows:

Definition 7.5. We define the standard ordering representation as the map $\rho_s \colon \mathbb{C}[q,p] \to \mathrm{Diff}(\mathbb{C}[q])$ with:

$$q^n p^m \mapsto \rho_s(q^n p^m) := q^n \left(-i\hbar \frac{\mathrm{d}}{\mathrm{d}q}\right)^m.$$

This is an injective linear map.

Definition 7.6. To deform the associative algebra on $\mathbb{C}[q,p]$ we define

$$f \star_s g := \rho_s^{-1}(\rho_s(f) \circ \rho_s(g)).$$

Proposition 7.7. *The preceding map $\star_s$ is a star product on $\mathbb{C}[q,p]$*

Proof. We begin by computing $\star_s$ using monomials and extending bilinearly. For

$m, n, j, k \in \mathbb{N}$ we have

$$\begin{aligned}
\rho_s(q^m p^n) \circ \rho_s(q^j p^k) &= q^m (-i\hbar)^n \frac{\mathrm{d}^n}{\mathrm{d}q^n}\left((q^j)(-i\hbar)^k \frac{\mathrm{d}^k}{\mathrm{d}q^k}\right) \\
&= q^m (-i\hbar)^{n+k} \frac{\mathrm{d}^n}{\mathrm{d}q^n}\left(q^j \frac{\mathrm{d}^k}{\mathrm{d}q^k}\right) \\
&= q^m (-i\hbar)^{n+k} \sum_{r=0}^{n} \binom{n}{r} \frac{j!}{(j-r)!} q^{j-r} \frac{\mathrm{d}^{k+n-r}}{\mathrm{d}q^{k+n-r}} \\
&= \sum (-i\hbar)^r \frac{n!}{(n-r)!r!} \frac{j!}{(j-r)!} q^{j+m-r} (-i\hbar)^{n+k-r} \frac{\mathrm{d}^{k+n-r}}{\mathrm{d}q^{k+n-r}} \\
&= \sum \frac{(-i\hbar)^r}{r!} \frac{j!}{(j-r)!} q^{j+m-r} \frac{n!}{(n-r)!} (-i\hbar)^{n+k-r} \frac{\mathrm{d}^{k+n-r}}{\mathrm{d}q^{k+n-r}}
\end{aligned}$$

If we convert back to q and p and say $f(q,p) = q^m p^n, g(q,p) = q^j p^k$ this sum can be written as

$$\sum \frac{(-i\hbar)^r}{r!} \frac{\partial^r f}{\partial p^r} \frac{\partial^r g}{\partial q^r}$$

which is a terminating sum for $f, q \in \mathbb{C}[q,p]$, so we might define

$$f \star_s g := fg - i\hbar \frac{\partial f}{\partial p}\frac{\partial g}{\partial q} + \sum_{r=2}^{\infty} \frac{(-i\hbar)^r}{r!} \frac{\partial^r f}{\partial p^r} \frac{\partial^r g}{\partial q^r}. \tag{7.3.1}$$

We see that this satisfies the asymptotic Dirac condition:

$$[f,g]_{\star_s} = f \star_s g - g \star_s f = -i\hbar \frac{\partial f}{\partial p}\frac{\partial g}{\partial q} + i\hbar \frac{\partial g}{\partial p}\frac{\partial f}{\partial q} + o(\hbar^2) = i\hbar\{f,g\} + o(\hbar^2)$$

We can show associativity by taking $f, g, h \in \mathbb{C}[q,p]$, we have

$$
\begin{aligned}
(f \star_s g) \star_s h &= \left(\sum_{r=0}^{\infty} \frac{(-i\hbar)^r}{r!} \frac{\partial^r f}{\partial p^r} \frac{\partial^r g}{\partial q^r} \right) \star_s h \\
&= \sum_{k=0}^{\infty} \frac{(-i\hbar)^k}{k!} \frac{\partial^k}{\partial p^k} \left(\sum_{r=0}^{\infty} \frac{(-i\hbar)^r}{r!} \frac{\partial^r f}{\partial p^r} \frac{\partial^r g}{\partial q^r} \right) \frac{\partial^k h}{\partial q^k} \\
&= \sum_{k,r=0}^{\infty} \frac{(-i\hbar)^k}{k!} \frac{(-i\hbar)^r}{r!} \sum_{n \leq k} \binom{k}{n} \frac{\partial^{r+n} f}{\partial p^{r+n}} \frac{\partial^{k-n}}{\partial p^{k-n}} \left(\frac{\partial^r g}{\partial q^r} \right) \frac{\partial^k h}{\partial q^k} \\
&= \sum_{k,r=0}^{\infty} \frac{(-i\hbar)^k}{k!} \frac{(-i\hbar)^r}{r!} \sum_{n \leq k} \binom{k}{n} \frac{\partial^{r+n} f}{\partial p^{r+n}} \frac{\partial^{r}}{\partial q^{r}} \left(\frac{\partial^{k-n} g}{\partial p^{k-n}} \right) \frac{\partial^k h}{\partial q^k} \\
&= \sum_{k,r=0}^{\infty} \frac{(-i\hbar)^k}{k!} \frac{(-i\hbar)^r}{r!} \sum_{n \leq k} \frac{k!}{(k-n)!n!} \frac{\partial^{r+n} f}{\partial p^{r+n}} \frac{\partial^{r}}{\partial q^{r}} \left(\frac{\partial^{k-n} g}{\partial p^{k-n}} \right) \frac{\partial^k h}{\partial q^k} \\
&= \sum_{k,r=0,n\leq k}^{\infty} \frac{(-i\hbar)^{k+r}}{r!} \frac{\partial^{r+n} f}{\partial p^{r+n}} \frac{1}{(k-n)!n!} \frac{\partial^{r}}{\partial q^{r}} \left(\frac{\partial^{k-n} g}{\partial p^{k-n}} \right) \frac{\partial^n}{\partial q^n} \left(\frac{\partial^{k-n} h}{\partial q^{k-n}} \right)
\end{aligned}
$$

Let $j = r + n$, then $n = j - r$ and our sum becomes

$$
\begin{aligned}
&\sum \frac{(-i\hbar)^{k+r}}{r!} \frac{\partial^j f}{\partial p^j} \frac{1}{(k-j+r)!(j-r)!} \frac{\partial^r}{\partial q^r} \left(\frac{\partial^{k-j+r} g}{\partial p^{k-j+r}} \right) \frac{\partial^{j-r}}{\partial q^{j-r}} \left(\frac{\partial^{k-j+r} h}{\partial q^{k-j+r}} \right) \\
&= \sum \frac{(-i\hbar)^{j}}{j!} \frac{(-i\hbar)^{k-j+r}}{(k-j+r)!} \frac{\partial^j f}{\partial p^j} \frac{j!}{r!(j-r)!} \frac{\partial^r}{\partial q^r} \left(\frac{\partial^{k-j+r} g}{\partial p^{k-j+r}} \right) \frac{\partial^{j-r}}{\partial q^{j-r}} \left(\frac{\partial^{k-j+r} h}{\partial q^{k-j+r}} \right)
\end{aligned}
$$

Taking the new index $l = (k - j + r)$ we have

$$
\begin{aligned}
&\sum \frac{(-i\hbar)^{j}}{j!} \frac{(-i\hbar)^{l}}{l!} \frac{\partial^j f}{\partial p^j} \sum_{r \leq j} \binom{j}{r} \frac{\partial^r}{\partial q^r} \left(\frac{\partial^{l} g}{\partial p^{l}} \right) \frac{\partial^{j-r}}{\partial q^{j-r}} \left(\frac{\partial^{l} h}{\partial q^{l}} \right) \\
&= \sum \frac{(-i\hbar)^{j}}{j!} \frac{\partial^j f}{\partial p^j} \frac{\partial^j}{\partial q^j} \left(\sum \frac{(-i\hbar)^{l}}{l!} \frac{\partial^l g}{\partial p^l} \frac{\partial^l h}{\partial q^l} \right) = f \star_s (g \star_s h) \qquad \square
\end{aligned}
$$

In a similar way, we can define $\star_s$ on $C^\infty(\mathbb{R}^{2n})$ as follows:

$$
f \star_s g := \sum \frac{(-i\hbar)^r}{r!} \sum_{k=1}^{n} \frac{\partial^r f}{\partial p_k^r} \frac{\partial^r g}{\partial q_k^r}
$$

and readily sees that the Dirac-Poisson relation and associativity follow from the preceding proof.

If we take $\rho_s(\mathbb{C}[q,p])$ to be operators on $C^\infty_{cpt}(\mathbb{R})[[\hbar]] \subset L^2(\mathbb{R})[[\hbar]]$ equipped with its standard inner product, we soon see an issue with our naive guess. One recalls that $\rho_s(q) = \hat{q}, \rho_s(p) = \hat{p} = -i\hbar\frac{\mathrm{d}}{\mathrm{d}q}$ are self adjoint operators on $C^\infty_{cpt}(\mathbb{R})$ which is an important property of observables, since one expects the expectation value of an observable to correspond to a real number. We see that $\rho_s(qp)^\dagger = (\hat{q}\hat{p})^\dagger = \hat{p}\hat{q}$ meaning that we have not assigned a self-adjoint operator to a real valued polynomial. In remedying this, we might map monomials to their symmetrization, which we will describe in the next definition:

Definition 7.8. We now define the so-called Weyl-Moyal representation, $\rho_w \colon \mathbb{C}[q,p] \to \mathrm{Diff}(\mathbb{C}[q])$, as

$$q^m p^n \mapsto \frac{1}{(m+n)!} \sum_{\sigma \in S_{m+n}} \sigma(\hat{q}, \underset{m \text{ times}}{\ldots}, \hat{q}, \hat{p}, \underset{n \text{ times}}{\ldots}, \hat{p})$$

Where $\sigma(a_1, \ldots, a_{m+n}) = a_{\sigma(1)} \ldots a_{\sigma(m+n)}$.

For example $\rho_w(qp) = \frac{1}{2}(\hat{q}\hat{p} + \hat{p}\hat{q})$. One verifies that $\rho_w(q^m p^n)$ is self-adjoint and ρ_w is injective.

In order to define the Weyl-Moyal product in a more convenient way, we will define the next operation.

Definition 7.9. Let A be a commutative associative algebra over $\mathbb{C}$. We define $\mu \colon A \otimes A \to A$ by bilinear extension of the following property:

$$\mu(u \otimes v) := uv$$

for all $u, v \in A$.

Definition 7.10. Let (M, ω) be a symplectic manifold. The Poisson tensor $P \in \bigwedge^2 TM$ is defined by

$$P(f,g) := \omega(X_f, X_g).$$

In canonical coordinates $P(f,g) = \sum \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i}$ so

$$P = \sum \frac{\partial}{\partial q^i} \wedge \frac{\partial}{\partial p_i}$$

We then define $\Lambda \colon C^\infty(M) \otimes C^\infty(M) \to C^\infty(M) \otimes C^\infty(M)$ by

$$\Lambda(f \otimes g) := \sum \frac{\partial f}{\partial q^i} \otimes \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \otimes \frac{\partial f}{\partial q^i}$$

In non-canonical coordinates, we simply have

$$\Lambda(f \otimes g) = \sum P^{ij} \frac{\partial f}{\partial x^i} \otimes \frac{\partial g}{\partial x^j}.$$

Definition 7.11. When $M = \mathbb{R}^{2n}$ and $\omega = \sum \mathrm{d}q^i \wedge \mathrm{d}p_i$ we define the Weyl-Moyal Product as follows:

$$f \star_w g := \sum \frac{(i\hbar)^n}{2^n n!} \mu(\Lambda^n (f \otimes g)).$$

Equivalently, we could define it as

$$\left(\exp\left(\frac{i\hbar P_{xy}}{2}\right) f(x, p_x) g(y, p_y)\right)\Big|_{x=y=q, p_x=p_y=p}$$

Where

$$P_{xy} f(x, p_x) g(y, p_y) = \sum \frac{\partial f}{\partial x^i}(x, p_x) \frac{\partial g}{\partial p_y^i}(y, p_y) - \frac{\partial f}{\partial p_x^i}(x, p_y) \frac{\partial g}{\partial y^i}(y, p_y)$$

Proposition 7.12. *The map $\star_w \colon C^\infty(\mathbb{R}^{2n}) \times C^\infty(\mathbb{R}^{2n}) \to C^\infty(\mathbb{R}^{2n})$ defined above is a star product.*

Proof. By simple computation we immediately see:

$$f \star_w g - g \star_w f = fg - gf + \frac{i\hbar}{2}\{f, g\} - \frac{i\hbar}{2}\{g, f\} + o(\hbar^2) = i\hbar\{f, g\} + o(\hbar^2)$$

Now, let $f, g, h \in C^\infty(\mathbb{R}^{2n})$. We then have the following

$$\begin{aligned}
(f \star_w g) \star_w h &= \left(\exp\left(\frac{i\hbar P_{xy}}{2}\right)(f \star_w g)(x, p_x) h(y, p_y)\right)\Big|_{x=y, p_x=p_y} \\
&= \left(\exp\left(\frac{i\hbar P_{xy}}{2}\right)\right)\left(\exp\left(\frac{i\hbar P_{xz}}{2}\right) f(x, p_y) g(z, p_z)\Big|_{x=z, p_x=p_z} h(y, p_y)\right)\Big|_{x=y, p_x=p_y} \\
&= \left(\exp\left(\frac{i\hbar(P_{xy} + P_{xz} + P_{yz})}{2}\right) f(x, p_x) g(z, p_z) h(y, p_y)\right)\Big|_{x=y=z, p_x=p_y=p_z} \\
&= f \star_w (g \star_w h)
\end{aligned}$$

□

and we conclude that this defines a star product on $C^\infty(\mathbb{R}^{2n})$

8 The Hydrogen Atom

The classical hydrogen atom's Hamiltonian takes the exact same form as the one from the Kepler problem, where this time $k = \frac{e^2}{4\pi\epsilon_0}$, rather than the typical $k = Gm_1m_2$. Though this detail is not mathematically important. The fact that the hydrogen atom and the two body problem have functionally equivalent Hamiltonians comes from the inverse square force law, which tells us that point objects (or spherical bodies) which have potential energy fields that obey $\Delta\Phi = \alpha\rho$ have force laws of the form $F = \frac{k}{r^2}\hat{r}$.

8.1 First Reduction

Throughout this section, we will take the typical star product $\star_{1/2}$ on $T^*(\mathbb{R}^6\backslash\Delta)$ and find a natural star product on the reduced phase space $T^*\mathbb{R}^+$. To find these reductions, we will mirror the classical case, starting with the reduction of phase space from $T^*(\mathbb{R}^6\backslash\Delta)$ to $T^*(\mathbb{R}^3\backslash\{0\})$.

The fashion in which we have created our projection $\pi\colon J^{-1}(0) \to J^{-1}(0)/\mathbb{R}^3$ allows us to extend this to $\pi_t\colon T^*(\mathbb{R}^6\backslash\Delta) \to T^*(\mathbb{R}^3\backslash\{0\})$ by taking $\pi_t(\mathbf{q}_1, \mathbf{q}_2, \mathbf{p}_1, \mathbf{p}_2) := (\mathbf{q}_1 - \mathbf{q}_2, \mathbf{p}_1 - \mathbf{p}_2)$. We see that $\pi_t\colon T^*(\mathbb{R}^6\backslash\Delta) \to T^*(\mathbb{R}^3\backslash\{0\})$ has a right inverse. Our right inverse is given by $j\colon T^*(\mathbb{R}^3\backslash\{0\}) \to T^*(\mathbb{R}^6\backslash\Delta)$, $(q, p) \mapsto (\mathbf{q}, 0, \mathbf{p}/2, -\mathbf{p}/2)$. We immediately see that $(\pi_t \circ j)(\mathbf{q}, \mathbf{p}) = \pi_t(\mathbf{q}, 0, \mathbf{p}/2, -\mathbf{p}/2) = (\mathbf{q}, \mathbf{p}/2 - (-\mathbf{p}/2)) = (\mathbf{q}, \mathbf{p})$. Then, to define our star product on $C^\infty(T^*(\mathbb{R}^3\backslash\{0\}))$ we take

$$f \tilde{\star}_{1/2} g := j^*(\pi_t^* f \star_{1/2} \pi_t^* g)$$

for all $f, g \in C^\infty(T^*(\mathbb{R}^3\backslash\{0\}))$.

Proposition 8.1. *The map*

$$\tilde{\star}_{1/2}\colon C^\infty(T^*(\mathbb{R}^3\backslash\{0\}))[[\nu]] \times C^\infty(T^*(\mathbb{R}^3\backslash\{0\}))[[\nu]] \to C^\infty(T^*(\mathbb{R}^3\backslash\{0\}))[[\nu]]$$

as defined above is a star product.

Proof. We start with our zeroth order condition. Given $f, g \in C^\infty(T^*(\mathbb{R}^3\backslash\{0\}))$ we see

$$\begin{aligned} f \tilde{\star}_{1/2} g &= j^*(\pi_t^*(f) \star_{1/2} \pi_t^*(g)) \\ &= j^*(\pi_t^*(f)\pi_t^*(g) + o(\nu)) \\ &= j^*(\pi_t^*(f))(\pi_t^*(g)) + o(\nu) \\ &= (\pi_t \circ j)^*(f)(\pi_t \circ j)^*(g) + o(\nu) \\ &= fg + o(\nu) \end{aligned}$$

So $\tilde{\star}_{1/2}$ fulfills the zero order quantization condition.

We now move on to the first order condition. Taking f and g as before, we have:

$$\begin{aligned} f \,\tilde{\star}_{1/2}\, g - g \,\tilde{\star}_{1/2}\, f &= j^*(\pi_t^* f \ \star_{1/2} \ \pi_t^* g) - i^*(\pi_t^* g \ \star_{1/2} \ \pi_t^* f) \\ &= j^*(\pi_t^* f \ \star_{1/2} \ \pi_t^* g - \pi_t^* g \ \star_{1/2} \ \pi_t^* f) \\ &= j^*(i\nu\{\pi_t^* f, \pi_t^* g\} + o(\nu^2)) \\ &= i\nu j^*\{\pi_t^* f, \pi_t^* g\} + o(\nu^2) \end{aligned}$$

From the preceding calculation, we know that in order to prove the first order condition, we need to prove the following:

$$j^*\{\pi_t^* f, \pi_t^* g\} = \{f, g\}.$$

We can do this by a simple coordinate calculation.

$$\begin{aligned}
\{\pi_t^* f, \pi_t^* g\}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{p}_1, \mathbf{p}_2) &= \boldsymbol{\nabla}_{\mathbf{q}_1} f(\mathbf{q}_1 - \mathbf{q}_2, \mathbf{p}_1 - \mathbf{p}_2) \cdot \boldsymbol{\nabla}_{\mathbf{p}_1} g(\mathbf{q}_1 - \mathbf{q}_2, \mathbf{p}_1 - \mathbf{p}_2) \\
&- \boldsymbol{\nabla}_{\mathbf{p}_1} f(\mathbf{q}_1 - \mathbf{q}_2, \mathbf{p}_1 - \mathbf{p}_2) \cdot \boldsymbol{\nabla}_{\mathbf{p}_1} g(\mathbf{q}_1 - \mathbf{q}_2, \mathbf{p}_1 - \mathbf{p}_2) \\
&+ \boldsymbol{\nabla}_{\mathbf{q}_2} f(\mathbf{q}_1 - \mathbf{q}_2, \mathbf{p}_1 - \mathbf{p}_2) \cdot \boldsymbol{\nabla}_{\mathbf{p}_2} g(\mathbf{q}_1 - \mathbf{q}_2, \mathbf{p}_1 - \mathbf{p}_2) \\
&- \boldsymbol{\nabla}_{\mathbf{p}_2} f(\mathbf{q}_1 - \mathbf{q}_2, \mathbf{p}_1 - \mathbf{p}_2) \cdot \boldsymbol{\nabla}_{\mathbf{p}_2} g(\mathbf{q}_1 - \mathbf{q}_2, \mathbf{p}_1 - \mathbf{p}_2) \\
&= (\boldsymbol{\nabla}_{\mathbf{q}} f \boldsymbol{\nabla}_{\mathbf{p}} g - \boldsymbol{\nabla}_{\mathbf{p}} f \boldsymbol{\nabla}_{\mathbf{q}} g)(\mathbf{q}_1 - \mathbf{q}_2, \mathbf{p}_1 - \mathbf{p}_2) \\
&+ (-1)(-1)(\boldsymbol{\nabla}_{\mathbf{q}} f \cdot \boldsymbol{\nabla}_{\mathbf{p}} g - \boldsymbol{\nabla}_{\mathbf{p}} f \cdot \boldsymbol{\nabla}_{\mathbf{q}} g)(\mathbf{q}_1 - \mathbf{q}_2, \mathbf{p}_1 - \mathbf{p}_2) \\
&= \pi_t^*\{f, g\}
\end{aligned}$$

Then we have $j^*\{\pi_t^* f, \pi_t^* g\} = j^* \pi_t^* \{f, g\} = (\pi_t \circ j)^* \{f, g\} = \{f, g\}$.

We will now prove the associativity condition.

$$\begin{aligned}
(f \,\tilde{\star}_{1/2}\, g) \,\tilde{\star}_{1/2}\, h &= j^*(\pi_t^* j^*(\pi_t^* f \ \star_{1/2} \ \pi_t^* g)) \ \star_{1/2} \ \pi_t^* h) \\
&= j^*((\pi_t^* f \ \star_{1/2} \ \pi_t^* g) \ \star_{1/2} \ \pi_t^* h) \\
&= j^*(\pi_t^* f \ \star_{1/2} \ (\pi_t^* g \ \star_{1/2} \ \pi_t^* h)) \\
&= j^*(\pi_t^* f \ \star_{1/2} \ \pi_t^* j^*(\pi_t^* g \ \star_{1/2} \ \pi_t^* h)) \\
&= f \,\tilde{\star}_{1/2}\, (g \,\tilde{\star}_{1/2}\, h)
\end{aligned}$$

Since $\tilde{\star}_{1/2}$ satisfies all three of the Dirac quanitization conditions, we conclude that it defines a star algebra structure on $C^\infty(T^*(\mathbb{R}^3 \backslash \{0\}))$

8.2 Second Reduction

Using this same argument, we can see that $\Lambda(\pi^*(f \otimes g)) = \pi^* \Lambda(f \otimes g)$ meaning that $\tilde{\star}_{1/2}$ coincides with $\star_{1/2}$ defined on $C^\infty(\mathbb{R}^6 \backslash \{0\})$ □

Now that we have created a star algebra on the first reduced space, we would like to do so in a similar fashion for the second reduced space. Without loss of generality, we will take $\tilde{L} = l\hat{z}$. We can then define an inclusion $\iota \colon T^*\mathbb{R}^+ \to \mu^{-1}(L)$

by $(q,p) \mapsto (q,0,0,p,\frac{l}{q},0)$. We see that

$$\begin{aligned}(\pi_2 \circ \iota)(q,p) &= \pi(q,0,0,p,\frac{l}{q},0)\\ &= \left(\sqrt{q^2}, \frac{1}{q}(q\hat{x}\cdot(p\hat{x}+\frac{l}{q}\hat{y}))\right)\\ &= (q, \frac{qp}{q}) = (q,p)\end{aligned}$$

We then define our new star product on $C^\infty(T^*\mathbb{R}^+)$ by

$$f \ \star'_{1/2} \ g := \iota^*(\pi^* f \,\tilde{\star}_{1/2}\, \pi^* g)$$

Proposition 8.2. *The map* $\star'_{1/2}\colon C^\infty(T^*\mathbb{R}^+)[[\lambda]]\times C^\infty(T^*\mathbb{R}^+)[[\lambda]] \to C^\infty(T^*\mathbb{R}^+)[[\lambda]]$ *as defined above defines a star product on* $C^\infty(T^*\mathbb{R}^+)$.

Proof. Our proof is almost verbatim to that of Proposition 5.0.1. We begin with the zeroth order condition. Given $f,g \in C^\infty(T^*\mathbb{R}^+)[[\lambda]]$ we have

$$\begin{aligned}f \ \star'_{1/2} \ g :&= \iota^*(\pi^* f \,\tilde{\star}_{1/2}\, \pi^* g)\\ &= \iota^*(\pi^* f \cdot \pi^* g + o(\nu))\\ &= \iota^*(\pi^* f \cdot \pi^* g) + o(\nu)\\ &= (\pi \circ \iota)^* f \cdot \iota^*\pi^* g + o(\nu)\\ &= f \cdot g + o(\nu)\end{aligned}$$

We now consider the second order condition. Taking f and g as before, we see that

$$\begin{aligned}f \ \star'_{1/2} \ g - g \ \star'_{1/2} \ f &= \iota^*(\pi^* f \,\tilde{\star}_{1/2}\, \pi^* g) - \iota^*(\pi^* g \,\tilde{\star}_{1/2}\, \pi^* f)\\ &= \iota^*(\pi^* f \,\tilde{\star}_{1/2}\, \pi^* g - \pi^* g \,\tilde{\star}_{1/2}\, \pi^* f)\\ &= \iota^*(i\nu\{\pi^* f, \pi^* g\} + o(\nu^2))\\ &= i\nu\iota^*\{\pi^* f, \pi^* g\} + o(\nu^2)\end{aligned}$$

Once again it suffices to show that π is a Poisson map. One sees that

$$\pi^* f(\mathbf{q},\mathbf{p}) = f\left(\|\mathbf{q}\|, \frac{\mathbf{q}\cdot\mathbf{p}}{\|\mathbf{q}\|}\right)$$

We can rewrite the Poisson bracket as

$$\{\pi^* f, \pi^* g\} = \boldsymbol{\nabla}_{\mathbf{q}}\pi^* f \cdot \boldsymbol{\nabla}_{\mathbf{p}}\pi^* g - \boldsymbol{\nabla}_{\mathbf{q}}\pi^* g \cdot \boldsymbol{\nabla}_{\mathbf{p}}\pi^* f$$

The chain rule then gives $\nabla_{\mathbf{q}}\pi^*f = \nabla_{\mathbf{q}}(f\circ\pi) = \nabla_{\mathbf{q}}(\|\mathbf{q}\|)\left(\frac{\partial f}{\partial q}\circ\pi\right)+\nabla_{\mathbf{q}}(\frac{\mathbf{q}\cdot\mathbf{p}}{\|\mathbf{q}\|})\left(\frac{\partial f}{\partial p}\circ\pi\right)$ and $\nabla_{\mathbf{p}}\pi^*f = \nabla_{\mathbf{p}}(\frac{\mathbf{q}\cdot\mathbf{p}}{\|\mathbf{q}\|})(\frac{\partial f}{\partial p}\circ\pi)$. In the first case we have

$$\nabla_{\mathbf{q}}\|\mathbf{q}\| = \frac{\mathbf{q}}{\|\mathbf{q}\|}$$

and

$$\nabla_{\mathbf{q}}\frac{\mathbf{q}\cdot\mathbf{p}}{\|\mathbf{q}\|} = \frac{\mathbf{p}}{\|\mathbf{q}\|} - \frac{\mathbf{q}\cdot\mathbf{p}}{\|\mathbf{q}\|^2}\frac{\mathbf{q}}{\|\mathbf{q}\|} = \frac{\mathbf{p}}{\|\mathbf{q}\|} - \frac{(\mathbf{q}\cdot\mathbf{p})\mathbf{q}}{\|\mathbf{q}\|^3}.$$

The second case simply gives $\nabla_{\mathbf{p}}(\frac{\mathbf{q}\cdot\mathbf{p}}{\|\mathbf{q}\|}) = \frac{\mathbf{q}}{\|\mathbf{q}\|}$. Plugging these expressions into our Poisson bracket equation then gives

$$\begin{aligned}
\{\pi^*f,\pi^*g\} &= \left(\frac{\mathbf{q}}{\|\mathbf{q}\|}\left(\frac{\partial f}{\partial q}\circ\pi\right) + \left(\frac{\mathbf{p}}{\|\mathbf{q}\|} - \frac{(\mathbf{q}\cdot\mathbf{p})\mathbf{q}}{\|\mathbf{q}\|^3}\right)\left(\frac{\partial f}{\partial p}\circ\pi\right)\right)\cdot\frac{\mathbf{q}}{\|\mathbf{q}\|}\left(\frac{\partial g}{\partial p}\circ\pi\right)\\
&\quad - \left(\frac{\mathbf{q}}{\|\mathbf{q}\|}\left(\frac{\partial g}{\partial q}\circ\pi\right) + \left(\frac{\mathbf{p}}{\|\mathbf{q}\|} - \frac{(\mathbf{q}\cdot\mathbf{p})\mathbf{q}}{\|\mathbf{q}\|^3}\right)\left(\frac{\partial g}{\partial p}\circ\pi\right)\right)\cdot\frac{\mathbf{q}}{\|\mathbf{q}\|}\left(\frac{\partial f}{\partial p}\circ\pi\right)\\
&= \frac{\mathbf{q}\cdot\mathbf{q}}{\|\mathbf{q}\|^2}\left(\frac{\partial f}{\partial q}\frac{\partial g}{\partial p}\right)\circ\pi + \left(\frac{\mathbf{p}\cdot\mathbf{q}}{\|\mathbf{q}\|} - \frac{(\mathbf{p}\cdot\mathbf{q})(\mathbf{q}\cdot\mathbf{q})}{\|\mathbf{q}\|^3}\right)\left(\frac{\partial f}{\partial p}\frac{\partial g}{\partial p}\right)\circ\pi\\
&\quad - \frac{\mathbf{q}\cdot\mathbf{q}}{\|\mathbf{q}\|^2}\left(\frac{\partial g}{\partial q}\frac{\partial f}{\partial p}\right)\circ\pi - \left(\frac{\mathbf{p}\cdot\mathbf{q}}{\|\mathbf{q}\|} - \frac{(\mathbf{p}\cdot\mathbf{q})(\mathbf{q}\cdot\mathbf{q})}{\|\mathbf{q}\|^3}\right)\left(\frac{\partial g}{\partial p}\frac{\partial f}{\partial p}\right)\circ\pi\\
&= \left(\frac{\partial f}{\partial q}\frac{\partial g}{\partial p}\right)\circ\pi + 0 - \left(\frac{\partial g}{\partial q}\frac{\partial f}{\partial p}\right)\circ\pi - 0\\
&= \pi^*\{f,g\}
\end{aligned}$$

From this we conclude that $\iota^*\{\pi^*f,\pi^*g\} = \iota^*\pi^*\{f,g\} = \{f,g\}$ and the first order quantization condition is fulfilled. Because our definition takes the same form as in Proposition 8.1 we omit the associativity argument. □

Due to the simplicity of the preceding arguments, it is easy to come up with the following proposition:

Proposition 8.3. *Let M be a symplectic manifold equipped with a G-equivariant momentum map $J\colon M\to\mathfrak{g}^*$ with $\mu\in\mathfrak{g}^*$ a regular value. Let $\star_{1/2} : C^\infty(M)[[\lambda]]\times C^\infty(M)[[\lambda]] \to C^\infty(M)[[\lambda]]$ be a star product on M. If $J^{-1}(\mu)$ is a trivial G_μ bundle with a section $\iota\colon J^{-1}(\mu)/G_\mu \to J^{-1}(\mu)$ and there is a global projection $\pi\colon M\to J^{-1}(\mu)$ which is Poisson then the following formla defines a star product on $J^{-1}(\mu)/G_\mu$:*

$$f\,\tilde{\star_{1/2}}\,g := \iota^*(\pi^*f \ \star_{1/2}\ \pi^*g)$$

Proof. The proof follows in the same fashion as Propositions 8.3 and 8.2. □

Bibliography

[1] F Bayen et al. "Deformation theory and quantization. I. Deformations of symplectic structures". en. In: *Annals of Physics* 111.1 (Mar. 1978), pp. 61–110. ISSN: 0003-4916. DOI: 10.1016/0003-4916(78)90224-5. URL: https://www.sciencedirect.com/science/article/pii/0003491678902245 (visited on 12/07/2021).

[2] Paul Adrien Maurice Dirac. *The fundamental equations of quantum mechanics*. en. DOI: 10.1098/rspa.1925.0150. URL: http://royalsocietypublishing.org/doi/epdf/10.1098/rspa.1925.0150 (visited on 03/28/2022).

[3] Leon van Hove. *Sur certaines representations unitaires d'un groupe infini de transformations.* 1951.

[4] Jerrold Marsden and Alan Weinstein. "Reduction of symplectic manifolds with symmetry". In: *Reports on Mathematical Physics* 5.1 (1974), pp. 121–130. ISSN: 0034-4877. DOI: https://doi.org/10.1016/0034-4877(74)90021-4. URL: https://www.sciencedirect.com/science/article/pii/0034487774900214.

[5] Jerrold E. Marsden and Tudor S. Ratiu. *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems*. en. Ed. by Jerrold E. Marsden et al. Vol. 17. Texts in Applied Mathematics. New York, NY: Springer New York, 1999. ISBN: 978-1-4419-3143-6 978-0-387-21792-5. DOI: 10.1007/978-0-387-21792-5. URL: http://link.springer.com/10.1007/978-0-387-21792-5 (visited on 09/26/2021).

[6] Ralph Abraham and Jerrold E. Marsden. *Foundations of Mechanics*. Second Edition. ISBN: 978-0-8218-4438-0.