Master thesis

An Introduction to the Orbit Method

J. Maes

Department of Mathematics, Utrecht University
Institute for Theoretical Physics, Utrecht University

Utrecht, April 4, 2011

...from the same principles, I now demonstrate the frame of the System of the World.

Isaac Newton ([17])

As time goes on, it becomes increasingly evident that the rules which the mathematician finds interesting are the same as those which Nature has chosen.

Paul Dirac at age 36, ([13])

Abstract

The Orbit Method is a method to determine all irreducible unitary representations of a Lie group. It is entangled with its physical counterpart geometric quantization, which is an extension of the canonical quantization scheme to general curved manifolds. The main ingredient of the Orbit Method is the notion of coadjoint orbits, which will be explained. Coadjoint orbits of a Lie group have the natural structure of a symplectic manifold, as does the phase space of a classical mechanical system. Naturally, geometric quantization will be treated next, since it attempts to provide a geometric interpretation of quantization within an extension of the mathematical framework of classical mechanics (symplectic geometry). In particular, the axioms imposed on a quantization will be discussed. Finally, as an application, coadjoint orbits and geometric quantization will be brought together by indicating how to determine the irreducible unitary representations of SU(2) by means of the Orbit Method.

Acknowledgments

This master thesis is not the work of just one person, but the result of the dedication of many people. First of all, I would like to thank my supervisors. Dr. Gleb Arutyunov from the Institute for Theoretical Physics, who had the brilliant ability to give strong intuition for abstract notions in physics and mathematics, without doing deficit to the theory itself, and Dr. Andre Henriques from the Department of Mathematics, who amazed me with his deep, fundamental understanding of the concepts we were discussing. He succeeded in significantly improving my skills in abstract mathematics the past year, by going back to the very fundamentals. Both of my supervisors were always willing to help me. Some of our appointments even took more then two hours.

Secondly, I want to thank my parents (John and Jolanda Maes), friends (Erwin Engelman, Martine Gennisse, Rolf van Vugt, Tristan van Heijst) and fellow students for their emotional support. At the beginning of my research I was dissapointed by the rate at which I was able to understand the new abstract theory and the growing consciousness that the final goal, to study the Virasoro group, could probably not be achieved. My parents and fellow students were always there to put things in a more realistic perspective and encouraged me to look for goals in other things then technical achievement too. This led me to write a thesis which is intended for both physics students with little knowledge of abstract mathematics and mathematics students interested in the physical ideas behind quantum mechanics as well. Only knowledge of calculus, linear algebra, elementary set theory and some topology is presupposed.

In realizing the, in my opinion, beautiful pictures in this thesis I was technically supported by Bram Jenneskens, who posseses skills in realizing vector pictures. Many thanks go to him as well. Further thanks go Tristan van Heijst, who corrected some earlier versions of my thesis in a critical and motivating way, Vivian Jacobs, Maarten Kater and Thessa Fokkema who helped me with computer tecnnical problems, Marleen Kooiman, who helped me with Mathematica, Thomas Wasserman, who answered many of my questions about representation theory, Daniël Prins, whose presentation on Frobenius theorem was useful in the subsection on polarizations and Jules Lamers, whose latex code I could use to give form to my own master thesis talk.

Considering the contents, for the vocabulary section I owe many thanks to the lecture notes of Erik van den Ban ([3]) and Eduard Looijenga ([10]). I used the notes of Matthias Blau ([9]) for both the introduction and the section on Geometric quantization. I am grateful for using the notes of Alexandre Kirillov ([1]) in an intensive manner for both sections about Coadjoint orbits. The section on geometric quantization could not have been written without the book of Nicholas Woodhouse ([7]). The lectures notes of Stefan Vandoren ([6]) where useful for both section 5 and section 6. Finally, for the section about the Orbit Method and the representaions of SU(2) the lecture notes of Erik van den Ban were again very important.

I am sure I forget to mention people here which do owe credit to this thesis. This is my final thanks to all of those people.

Contents

1 Conventions and notation

-Today's shocks are tomorrow's conventions.-
Carolyn Heilbrun ([15])

Although sometimes the notation used is already explained in the thesis itself, the total amount of notation introduced is quite big. Therefore, this section serves as a handy overview of (often) used notation troughout this thesis. Furthermore, the conventions used will be given.

Notation:
$\mathbb{K}$: the field of either the real or complex numbers.
$U \equiv \{U_i\}_{i \in A}$: an open cover of a manifold. Here A denotes some index set.
S^2: the two-dimensional sphere.
(M, ω): a symplectic manifold M with symplectic 2-form ω.
(M, J, ω): a Kähler manifold M with symplectic 2-form ω and complex structure J.
$T_m M$: the tangent space of a manifold M at the point $m \in M$.
TM: the tangent bundle of a manifold M.
T^*M: the cotangent bundle of a manifold M.
$\Omega^k(M)$: the set of k-forms on a manifold M.
$\mathrm{Vect}(M)$: the set of vector fields on a manifold M.
$\mathrm{Vect}(M; P)$: the set of vector fiels X such that $X|_m \in P_m$ for every $m \in M$. Here P denotes a polarization.
$\mathrm{End}(X)$: the set of endomorphisms of a mathematical object X.
$\mathrm{Aut}(X)$: the set of automorphisms of a mathematical object X.
$\mathrm{Mat}(n, \mathbb{K})$: the group of invertible $n \times n$-matrices with entries in $\mathbb{K}$.
$GL(V)$: the group of invertible, linear transformations on a vector space V.
$C^\infty(M, \mathbb{K})$: the space of either real or complex-valued functions on a manifold M.
$C^\infty(M, \mathbb{K}; P)$: the functions in $C^\infty(M, \mathbb{K})$ whose flow leaves the polarization invariant.
$L^2(M, \mathbb{C})$: the space of square-integrable, smooth, complex-valued functions on a manifold M.
$P(M)$: the space of polynomials on an manifold M.
$P_n(M)$: the space of homogeneous polynomials of degree n on an manifold M.
$\Gamma_X(M)$: the set of smooth sections over a real manifold M of a line bundle X.
h_*: the push-forward corresponding to the diffeomorphism $h : M \to N$ with M, N manifolds.

h^*: the pull-back corresponding to the smooth map $h : M \to N$ with M, N manifolds.
exp: the exponential map.
$\mathfrak{g}$: the Lie algebra of a Lie group G.
Id: the Identity map.
$\pi = (\pi, V)$: a representation of a Lie group G in a vector space V.
$\pi^* = (\pi^*, V^*)$: the representation dual to π in the dual space V^*.
$V^{\mathbb{C}}$: the complexification of a vector space V.
$TM^{\mathbb{C}}$: the complexified tangent bundle. $TM^{\mathbb{C}} \equiv \coprod_{m\in M} (T_m M)^{\mathbb{C}}$.
V^*: the dual of a vector space V.
Ad: the adjoint representation.
Ad*: the coadjoint representation.
ad: ad $\equiv T_e Ad$.
ad*: $\mathrm{ad}^*(X) \equiv (-\mathrm{ad}(X))^*$ for $X \in \mathfrak{g}$.
X_f: a Hamiltonian vector field corresponding to a real scalar function f.
$\mathcal{L}_v$: the Lie derivative along a vector field v.
∇: the connection.
$[,]$: the Lie bracket.
$\{,\}$: the Poisson bracket.
$\langle,\rangle$: evaluated on a linear functional f in the first slot and a vector v in the second slot it denotes the value of the linear functional f on the vector v. Otherwise, it denotes the inner product.
$C^p(U,\mathbb{R})$: the set of p-cochains corresponding to real, locally constant functions. Here U denotes an open covering of a manifold M.
$Z^p(U,\mathbb{R})$: the set of p-cocycles corresponding to real, locally constant functions. Here U denotes an open covering of a manifold M.
$H^p(U,\mathbb{R})$: the pth Čech cohomology group with coefficients in $\mathbb{R}$. Here U denotes an open covering of a manifold M.
$H^p(M,\mathbb{R})$: the pth de Rham cohomology group of a manifold M with coefficients in $\mathbb{R}$.

Conventions:
Skew symmetrization: $T^{[ab\ldots c]} = \frac{1}{p!}\sum_\sigma \mathrm{sign}(\sigma) T^{\sigma(a)\sigma(b)\ldots\sigma(c)}$,
where T is a p-index tensor and the summations are over all permutations of $a, b, ..., c$.
Interior product: $i(X)\omega = k\omega(X, ...)$ for a vector field X and $\omega \in \Omega^k(M)$, where M denotes a manifold.

2 Introduction

- When one looks back over the development of physics, one sees that it can be pictured as a rather steady development with many small steps and superposed on that a number of big jumps. These big jumps usually consist in overcoming a prejudice... And then a physicist has to replace this prejudice by something more precise, and leading to some entirely new conception of nature.-
Paul Dirac on Quantum Theory ([13])

At the macroscopic level our world seems to be pretty much gouverned by the laws of classical physics, that is, by Newtonian mechanics on the one hand and by Maxwell theory on the other. Newtonian mechanics describes the motion of particles under the influence of forces acting on them. Maxwell theory covers almost the entire spectrum of phenomena occuring in electromagnetism and optics. The in these two theories is governed by deterministic equations of motion and it is possible, given the initial conditions, to predict the results of measurements on the system at any later time.

At first sight, classical physics could thus provide us with a very satisfactory description of the world we live in. However, it fails to give an explanation for a number of phenomena observed at the microscopic level and, moreover, is in plain contradiction with experimental evidence.

As an illustration of what these phenomena are and what they may be trying to tell us about a theory which will have to replace classical physics, we mention two examples. The first is regarding the stability of atoms. From scattering and other experiments it had been deduced that atoms consist of a tiny positively charged nucleus orbited at quite some distance by negatively charged point-like particles, the electrons. Classical physics would predict such structures to collapse within fractions of a second, in contradiction with the stability of the world around us: orbital motion is accelerated motion, and according to Maxwell theory accelerated charges emit radiation; consequently, the electron would radiate away energy and spiral into the nucleus of the atom. It was also observed that simple atoms (like hydrogen) were able to adsorb and emit energy only in certain discrete quantities. Combining these two observations it thus appeared to be necsesary to postulate the existence of stable orbits for electrons at certain discrete radii (energy levels). This suggests that at the microscopic level, as opposed to in classical physics, nature allows for a discrete (or quantized) structure.

The second example is the set of experiments related to the nature of light. On the one hand, there is the famous double-slit experiment which investigated the diffraction and interference patterns of beams of particles like electrons. The experiment is depicted in figure 1 (from [11]). It indicated

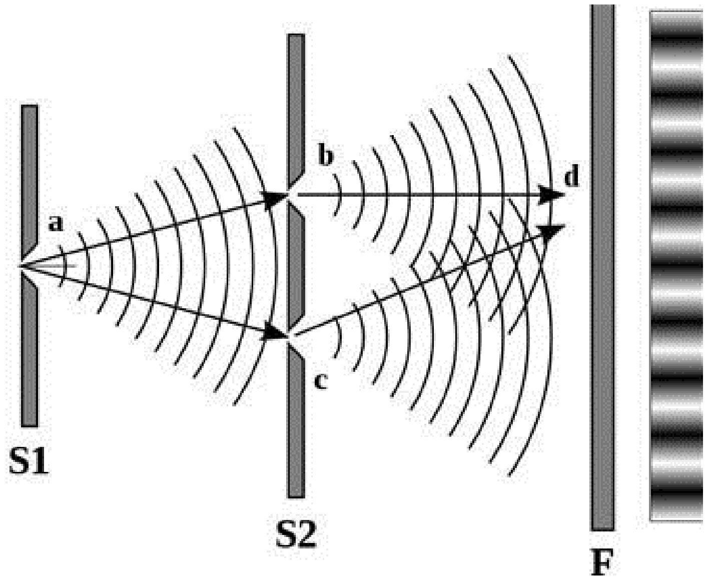

Figure 1: The double-slit experiment. A single photon (being a probability wave) departs from slit **a** in plate **S1**. When the wavefront reaches slits **b** and **c** in plate **S2** it creates two new probability waves, one emerging from slit **b** and one from slit **c**. The single photon will show up on the detector screen **F** (at say **d**) according to the net probability values resulting from the co-incidence of the probability waves coming by way of the two slits **b** and **c**.

that under certain circumstances particles (like electrons) can show interference patterns and thus wave-like nature. On the other hand, Einstein's interpretations of the photo-electric effect made clear that light showed behaviour characteristic of particles and not of waves. In the words of the great Dirac (see [12, p.3]),

'We have here a very striking and general example of the breakdown of classical mechanics - not merely an inaccuracy of its laws of motion, but an inadequacy of its concepts to supply us with a description of atomic events.'

Furthermore, the outcome of these experiments appeared to depend on the measuring process itself, i.e. whether or not one was checking through which slit the electron went. This, for the first time, implied that in order to describe a system the effects of observation of that system had to be taken into account. In particular, it implied that one cannot make any observations on a suitably 'small' system without perturbing the system itself (expressed in the uncertainty principle).

Heisenberg and Schrödinger provided two mathematical models, later shown to be equivalent, which were able to reproduce the above results and make many other succesfully tested predictions. These models, collectively known as **quantum mechanics** , describe the quantum behaviour of particles in flat space under the influence of external forces. Supplemented by some interpretation of, roughly speaking, the role of the measuring process or observer, they constitute a major step forward in the understanding of quantum physics in general.

However, at the conceptual level the situation was not satisfactory. In particular, it was not clear how general the proposed models were, which features were to be regarded as fundamental to any quantum version of a classical theory and which were to be attributed to particular properties of the systems considered so far. To gain some insight into this question, it was in particular Dirac who emphasized the formal similarities between classical and quantum mechanics and the necessity of properly understanding these. Abstracting from the analogy found between classical mechanics and Schrödinger and Heisenberg quantum mechanics, Dirac formulated a general quantum condition, a guideline for passing from a given classical system to the corresponding quantum theory. This process in general is known as **quantization** . Roughly speaking, quantization consists of replacing the classical algebra of observables (smooth functions on phase space) by an algebra of operators acting on some Hilbert space, the quantum condition relating the commutator of two operators to the Poisson bracket of their classical counterparts:

$$Poisson\ \ Brackets \rightarrow Commutators.$$

At first sight, the very concept of quantization appears to be ill-founded since it attempts to construct a 'correct' theory from a theory which is only approximately correct. After all, our world is quantum, and while it seems a legitimate task to try to extract classical mechanics in some limit from quantum mechanics, there seems too little reason to believe that the inverse construction can always be performed:

Classical limit: Quantum mechanics $\Longrightarrow$ Classical mechanics
Quantization: Quantum mechanics $\overset{?}{\Longleftarrow}$ Classical mechanics.

Furthermore, there is no reason to believe that such a construction would be unique as there could well be (and, in fact, are) lots of different quantum theories which have the same classical limit. Unfortunately, however, it is conceptually very difficult to describe a quantum theory from scratch, without the help of a reference classical theory. Moreover, there is enough to the analogy between classical and quantum mechanics to make quantization a worthwile approach. Perhaps, ultimately, the study of quantization will tell us enough about quantum theory itself to do without the very concept of quantization.

Unfortunately, Dirac's quantum condition is not as general as one might have hoped it to be, or, at least, not sufficiently unambiguous (see, [12]). Thus, it is desirable to find a more intrinsic and constructive description of this quantization procedure. This is the aim of **geometric quantization** which attempts to provide a geometric interpretation of quantization within an extension of the mathematical framework of classical mechanics (symplectic geometry).

Geometric Quantization is the physical counterpart of the Orbit Method . The Orbit Method is a method to determine all irreducible unitary representations of a Lie group. This is useful, since representation theory remains the method of choice for simplifying the physical analysis of systems possesing symmetry. The main ingredient of the Orbit Method is the notion of coadjoint orbits. What is so special about coadjoint orbits is that they have the natural structure of a symplectic manifold, as does the phase space of a classical mechanical system. Indeed, Geometric Quantization is another ingredient of the Orbit Method. That is why the Orbit Method brings the ingredients Geometric Quantization and the notion of coadjoint orbits together in a beautiful way. The idea behind it is to unite harmonic analysis with symplectic geometry. This idea can be considered as a part of the more general idea of the unification of mathematics and physics.

3 Vocabulary

-The golden age of mathematics, that was not the age of Euclid, it is ours.-
C.J. Keyser ([16])

To study the orbit method, a good knowledge and understanding of the language of manifolds, Lie groups and representation theory is essential. For this reason this section states the most important definitions and results related to these subjects. In this way the exposition of the orbit method is made as self-contained as possible. The reader can therefore view this section as a dictionary and look back at important definitions whenever necessary. It begins with manifolds, continues with listing some morphisms, then treats Lie groups and Lie algebras and ends with some basic theory about actions and representations. This section would not have been possible without the lecture notes [3] and [10].

3.1 The language of manifolds

Definition 3.1.1. A **k-smooth n-dimensional manifold** is a topological space M that admits a covering by open sets U_α, $\alpha \in A$, endowed with one-to-one continuous maps $\phi_\alpha : U_\alpha \to V_\alpha \subset \mathbb{R}^n$ so that all maps $\phi_{\alpha,\beta} \equiv \phi_\alpha \circ \phi_\beta^{-1}$ are k-smooth (i.e. have continuous partial derivatives of order $\leq k$) whenever defined.

We call the manifold M smooth when all the maps $\phi_{\alpha,\beta}$ are infinitely differentiable.

The following terminology is used:

- the sets U_α are called charts;
- the functions $x^i_\alpha = x^i \circ \phi_\alpha$ are called local coordinates. Here $\{x^i\}_{1\leq i\leq n}$ are the standard coordinates in $\mathbb{R}^n$;
- the functions $\phi_{\alpha,\beta}$ are called the transition functions;
- the collection $\{U_\alpha, \phi_\alpha\}_{\alpha\in A}$ is called the atlas on M.

A given topological space M can be endowed with several different atlases satisfying all the requirements of the definition. Should the corresponding objects be considered as different or as the same manifold? To clear this up, we introduce

Definition 3.1.2. Two atlases $\{U_\alpha\}_{\alpha\in A}$ and $\{U'_\beta\}_{\beta\in B}$ are called **equivalent** if the transition functions from any chart of the first atlas to any chart of the second one are k-smooth (whenever defined). The **structure of a smooth manifold** on M is an equivalence class of atlases.

Remark 3.1.1. When dealing with a smooth manifold, we usually follow a common practice and each time use only one, the most appropriate, atlas, while keeping in mind that we can always replace it by an equivalent one. In practical computations people usually prefer minimal atlases. For example, for the n-dimensional sphere S^n and for the n-dimensional torus T^n there exist atlases with only two charts.

The following definition tells us when a subset of a (C^k-)manifold is a (C^k-)manifold itself.

Definition 3.1.3. Given a C^k-manifold N of dimension n, then we say that a subset $M \subset N$ is a C^k**-submanifold** of dimension m if M can be covered by charts (U, ϕ) of N with the property that $M \cap U = \phi^{-1}(\mathbb{R}^m \times \{0\})$. (If $k = \infty$ we will also call this a smooth submanifold.)

Hence we must have $m \leq n$ and $\mathbb{R}^n$ is written as $\mathbb{R}^m \times \mathbb{R}^{n-m}$. Coming to the geometry of manifolds, the following notion is essential:

Definition 3.1.4. A vector bundle over the manifold M (called the **base manifold**) is a real manifold X (called the **total space**) together with a smooth map $\mathrm{pr} : X \to M$ (called the **projection**) such that

(VB1) For each $m \in M$, $X_m = \mathrm{pr}^{-1}(m)$ is a vector space over $\mathbb{K}$ of constant dimension r (X_m is called the **fibre** over $\{m\}$ and r is called the **fibre dimension**),

(VB2) For each $m_0 \in M$, there is a neighbourhood $U \in m_0$ in M together with a diffeomorphism $\psi : U \times \mathbb{K}^r \to \mathrm{pr}^{-1}(U)$ such that $\psi|_m : \{m\} \times \mathbb{K}^r \to X_m$ is a linear isomorphism for each $m \in U$ (the pair (U, ψ) is called a **local trivialization**).

In the definition above M is a n-dimensional real manifold and $\mathbb{K}$ denotes the field of either the real or complex numbers. A map $s : U \subset B \to X$ is called a **section** of a vector bundle X if $\mathrm{pr} \circ s = \mathrm{Id}$.

The most important example of a vector bundle is the tangent bundle TM. It was the source of the whole theory of vector bundles. In order to understand what the space TM is we first introduce an auxiliary definition. A **Curve** at $m \in M$ is a smooth map γ from an open neighbourhood of 0 in $\mathbb{R}$ to M with $\gamma(0) = m$. On the collection of curves at m we define an equivalence relation: $\gamma_1 \equiv \gamma_2$ if for a chart U at m and the corresponding diffeomorphism α, $\alpha \circ \gamma_1$ and $\alpha \circ \gamma_2$ have the same derivative in 0. The equivalence class of γ relative to this equivalence relation will be called the **derivative of γ at 0** and will be denoted by $\dot{\gamma}(0)$. This relation is independent of the choice of U. This leads us to the following definition:

Definition 3.1.5. A **tangent vector** of M at m is the derivative at 0 of a curve at m. The collection of tangent vectors of M at m is the **tangent space** of M at m and will be denoted by T_mM. The union $TM \equiv \bigcup_{m \in M} T_mM$ is called the **tangent bundle** .

The union TM forms a vector bundle over M with the projection pr that maps all elements of T_mM to m: Indeed, let (U, ϕ), with $U \subset M$, be a local chart with coordinates $(x^1, ..., x^n)$. First of all, T_mM is a vector space. How is addition and scalar multiplication defined on this vector space? If we take two curves $\gamma(t)$ and $\tilde{\gamma}(t)$ trough m with $\gamma(0) = \tilde{\gamma}(0) = m$, then adding them or multiplying them by a scalar need not produce a curve in M. However, we can add these curves, using the chart ϕ, as follows:
(1) Addition: $(\gamma + \tilde{\gamma})(t) = \phi^{-1}((\phi(\gamma(t))) + \phi(\tilde{\gamma}(t)))$
(2) Scalar multiplication: $(\lambda\gamma)(t) = \phi^{-1}(\lambda\phi(\gamma(t)))$ $(\lambda \in \mathbb{R})$
Taking the derivatives of $\gamma + \tilde{\gamma}$ and $\lambda\gamma$ at the point $0 \in \mathbb{R}$ will, by the chain rule, produce the sum and scalar multiples of the corresponding tangent vectors (see also [30]). Furthermore, any tangent vector $a \in T_mM$ can be written as $a = a^k \partial_k$ where $a^k = a^k(x)$ are numerical coefficients and $\partial_k \equiv \partial/\partial x^k$ denotes the partial derivative. So the set $\mathrm{pr}^{-1}(U)$ is identified with $U \times \mathbb{R}^n$. It is by definition a local chart on TM with coordinates $(x^1, ..., x^n; a^1, ..., a^n)$. This proves the claim.

The smooth sections of TM are called smooth **vector fields** on M. We denote the space of all smooth vector fields on M by $\mathrm{Vect}(M)$.

The second important example of a vector bundle is the so-called cotangent bundle.

Definition 3.1.6. Given a manifold M, then the **cotangent space** of M at $m \in M$ is the dual T^*_mM of the tangent space T_mM. The **cotangent bundle** of M is the dual of the tangent bundle and is denoted by T^*M. A **differential 1-form** on M is a section of its cotangent bundle.

We denote the space of all smooth differential 1-forms on M by $\Omega^1(M)$, which is dual to Vect(M). Locally, we can write a differential 1-form as the expression $\omega = \omega_i dx^i$ for some chart U, where $\{dx^i\}_{1\leq i\leq n}$ forms the basis of $\Omega^1(U)$ dual to the basis $\{\partial_i\}_{1\leq i\leq n}$ of Vect(U).

The space $\Omega^k(M)$ of smooth **differential k-forms** on M can be defined can be defined as the k-th exterior power of $\Omega^1(M)$. In a local coordinate system a k-form ω looks like

$$\omega = \sum_{i_1<...<i_k} \omega_{i_1,...,i_k}(x) dx^{i_1} \wedge ... \wedge dx^{i_k}, \tag{1}$$

where all indices i_s, $1 \leq s \leq k$, run from 1 to n and the wedge product is bilinear, associative, and anti-symmetric.

We will now define the **exterior derivative** . The exterior derivative d is the unique $\mathbb{R}$-linear mapping from k-forms to $(k+1)$-forms satisfying the following properties:

(1) df is the differential 1-form of f for smooth functions f.

(2) Flatness: $\mathrm{d} \circ \mathrm{d} = 0$.

(3) Leibniz rule: $\mathrm{d}(\alpha \wedge \beta) = \mathrm{d}\alpha \wedge \beta + (-1)^p(\alpha \wedge \mathrm{d}\beta)$ where α is a differential p-form.

We will end this subsection with a very important object in differential geometry, which is the **Lie derivative $\mathcal{L}$**. Let M be a manifold. Three particular cases of the Lie derivative are:

(a) The ordinary directional derivative $v \cdot f$ of a real smooth scalar function $f : M \to \mathbb{R}$ along the vector field v is just $\mathcal{L}_v f$;

(b) The so-called **Lie bracket** of two vector fields $v, w \in$ Vect(M) is $[v, w] = \mathcal{L}_v w = -\mathcal{L}_w v$;

(c) For a general differential k-form ω the **Cartan formula** holds: $L_v\omega = i_v\mathrm{d}\omega + \mathrm{d}(i_v\omega)$ where i_v is the interior multiplication, that is, it is the map $i_v : \Omega^k(M) \to \Omega^{k-1}(M)$ which sends a k-form ω to the $(k-1)$-form $i_v\omega$ defined by the property that $(i_v\omega)(v_1, ..., v_{k-1}) = \omega(v, v_1, ..., v_{k-1})$ for any vector fields $v_1, ..., v_{k-1}$.

The Lie derivative acting on differential k-forms (case (c)) has some particular properties. In order to quote them we first need to define an auxiliary notion.

Definition 3.1.7. A **flow** on a manifold M is a smooth map $H : \mathbb{R} \times M \to M$ with the property that if $H_t : M \to M$ is defined by $H_t(p) = H(t,p)$, then H_0 is the identity map and $H_s H_t = H_{s+t}$ for all $s, t \in \mathbb{R}$.

You may think of a flow in the definition above as a fluid in motion which does not change in time (the map in the definition above then describes how a given fluid particle moves in time). The **orbit** of a point $p \in M$ is the image of the map $\gamma_p : \mathbb{R} \to M$ $\gamma_p(t) \equiv H(t,p)$. The orbit map γ_p defines a tangent vector $\dot{\gamma}_p(0) \in T_pM$. In the physical situation of a fluid in motion this can be interpreted as the velocity of the particle at p. By letting p vary, we get a vector field on M. We call this vector field the **infinitesimal generator** of the flow and denote it by $\frac{\partial H}{\partial t}\big|_{t=0}$.

We now quote without proof three properties satisfied by the Lie derivative acting on differential k-forms.

Theorem 3.1.1. The Lie derivative $\mathcal{L}$ obeys the following three properties.

(1) commutation with d: $d\mathcal{L}_V = \mathcal{L}_V d$,

(2) Leibniz rule: if $\alpha, \beta \in \Omega^k(M)$, then
$\mathcal{L}_V(\alpha \wedge \beta) = \mathcal{L}_V(\alpha) \wedge \beta + \alpha \wedge \mathcal{L}_V(\beta)$,

(3) Infinitesimal pull-back: if V is the infinitesimal generator of a flow H on M, then for every k-form α on M,
$\frac{\partial}{\partial t}\big|_{t=0} H_t^* \alpha = \mathcal{L}_V \alpha$.

3.2 Structure-preserving maps (morphisms)

The notion of **morphism** (structure-preserving map) plays a central role in many modern branches of mathematics. In this thesis, treating a subject in mathematics and physics, this notion is dominantly present. In particular, in the process of quantization we want to preserve as much structure as possible, when going from classical to quantum physics. The general study of morphisms and of the structures (called **objects**) over which they are defined is treated in **category theory** . Much of the terminology of morphisms, as well as the intuition underlying them, comes from concrete categories, where the objects are simply sets with some additional structure, and morphisms are functions preserving the structure. For the purposes of this thesis it suffices to define some particular, often used, types of morphisms. From the context it should always be clear what the mathematical objects under consideration are. Giving the abstract definition of morphism from category theory is unnecessary here and would only distract the reader from the main goal. This subsection is therefore an enumeration of the concrete types of morphisms used in this thesis and their definitions.

We start with:

Definition 3.2.1. *(Homomorphism)*
Given are two groups $(G, *)$, $(H, \cdot)$. A **Homomorphism** from $(G, *)$ to $(H, \cdot)$ is a function $h : G \to H$ such that for $u, v \in G$ it holds that
$h(u * v) = h(u) \cdot h(v)$.

From this definition follows:

- h maps the unit element e_G of G to the unit element e_H of H. Namely:
  $h(e_G) = h(e_G * e_G) = h(e_G) \cdot h(e_G) \Rightarrow h(e_G) = e_H$.
- h maps inverses to inverses. Namely:
  $h(u) \cdot h(u^{-1}) = h(u * u^{-1}) = h(e_G) = e_H \Rightarrow h(u^{-1}) = h(u)^{-1}$.

We say that h is compatible with the group structure. In general, the definition of homomorphism depends on the type of algebraic structure under consideration. The common theme is that a homomorphism is a function between two algebraic objects that respects the algebraic structure. Let us now define:

Definition 3.2.2. *(Isomorphism)*
An **Isomorphism** is a bijective map f such that both f and its inverse f^{-1} are homomorphisms.

Definition 3.2.3. *(Endomorphism)*
An **endomorphism** is a morphism from a mathematical object to itself. An endomorphism of a vector space V is a linear map $f : V \to V$. An endomorphism of a group G is a homomorphism $f : G \to G$.

Definition 3.2.4. *(Automorphism)*
An **automorphism** is an endomorphism which is also an isomorphism. In case the mathematical objects are groups an automorphism is simply an isomorphism from the group to itself. In case the mathematical objects are vector spaces an automorphism is an invertible linear map from the vector space to itself.

The set of all automorphisms is a subset of $\mathrm{End}(X)$, which denotes the set of all endomorphisms of a mathematical object X, with a group structure, called the **automorphism group** of X and denoted by Aut(X). The definitions and their relations are summarized in the diagram below:

$$\begin{array}{ccc} \textit{automorphism} & \longrightarrow & \textit{isomorphism} \\ \downarrow & & \downarrow \\ \textit{endomorphism} & \longrightarrow & (\textit{homo})\textit{morphism}. \end{array}$$

Here an arrow denotes implication.

We finally define:

Definition 3.2.5. *(Diffeomorphism)*
Given two manifolds M and N, a bijective map f from M to N is called a **diffeomorphism** if both $f : M \to N$ and its inverse $f^{-1} : N \to M$ are differentiable (if these functions are r times continuously differentiable f is called a C^r-diffeomorphism). A diffeomorphism is an isomorphism in the category of smooth manifolds. Two manifolds M and N are (C^r-)diffeomorphic if there is an (r times) continuously differentiable bijective map f from M to N with an (r times) continuously differentiable inverse.

Two useful notions related to morphisms are the **push-forward** and **pull-back** .
Push-forward: given a diffeomorphism $h : M \to N$, then objects on M (such as a function $f : M \to \mathbb{R}$ or a vector field $V : M \to TM$) can be transfered to N to give a corresponding object on N. This is called push-forward and the transfering map is often denoted by h_*. For instance, $h_* f \equiv f \circ h^{-1}$ and $h_* V \equiv Dh \circ V \circ h^{-1} : N \to TN$.
Pull-back: similarly, the pull-back refers to transfering objects on N via h to objects on M and is denoted by h^*. This is in fact the push-forward of h^{-1}. The difference is however that a pull-back often makes sense when h only is a C^∞-map (without insisting that it is a diffeomorphism). For instance, the pull-back of a function g on N is simply $h^* g \equiv g \circ h$.

3.3 Lie groups and Lie algebras

Definition 3.3.1. A **Lie group** is a smooth (C^∞) manifold G equipped with a group structure so that the maps $\mu : G \times G \to G \ (x, y) \mapsto xy$ and $\iota : G \to G \ x \mapsto x^{-1}$ are smooth.

We quote without proof (for the proof, see [3, p.6]):

Lemma 3.3.1. Let G be a Lie group, and let $H \subset G$ be both a subgroup and a smooth submanifold. Then H is a Lie group.

Next we come to the concept of isomorphic Lie groups. Let G and H be Lie groups.

(a) A **Lie group homomomorphism** from G to H is a smooth map $\phi : G \to H$ that is a homomorphism of groups.

(b) A **Lie group isomorphism** from G to H is a bijective Lie group homomorphism $\phi : G \to H$ whose inverse is also a Lie group homomorphism.

(c) A **Lie group automorphism** of G is a Lie group isomorphism of G onto itself.

If $\phi : G \to H$ is a Lie group isomorphism, then ϕ is smooth and bijective and its inverse is smooth as well. Hence ϕ is a diffeomorphism.

Before we define the **exponential map** we first need to define what **left invariant** vector fields are. A vector field $v \in \mathrm{Vect}(G)$ is called left invariant if $(l_x)_* v = v$ for all $x \in G$, where l_x is the translation map $l_x : G \to G \quad y \mapsto xy$, which is a diffeomorphism. Equivalently, v is left invariant, when

$$v(xy) = T_y(l_x)v(y) \quad (x, y \in G). \tag{2}$$

From the above equation with $y = e$ we see that a left invariant vector field is completely determined by its value $v(e) \in T_eG$. Now

Definition 3.3.2. The exponential map $\exp = \exp_G : T_eG \to G$ is defined by $\exp(X) = \alpha_X(1)$ where α_X is the maximal integral curve with initial point e of the left invariant vector field v_X on G determined by $v_X(e) = X$, i.e.

$$\frac{d}{dt}\alpha_X(t) = v_X(\alpha_X(t)). \tag{3}$$

The following lemma, of which the proof can be found in for example [3, p.16], states some properties of the exponential map and at the same times justifies its name.

Lemma 3.3.2. For all $s, t \in \mathbb{R}$, $X \in T_eG$ we have

(i) $\exp(sX) = \alpha_X(s)$

(ii) $\exp((s+t)X) = \exp(sX)\exp(tX)$.

Moreover, the map exp: $T_eG \to G$ is smooth and a local diffeormorphism at 0. Its tangent map at the origin is given by $T_0\exp = I_{T_eG}$.

We now come to a very important application of the exponential map. The proof of the lemma can be found in [3, p.17].

Lemma 3.3.3. Let $\phi : G \to H$ be a homomorphism of Lie groups. Then the following diagram commutes:

$$\begin{array}{ccc} G & \xrightarrow{\phi} & H \\ \uparrow{\scriptstyle exp_G} & & \uparrow{\scriptstyle exp_H} \\ T_eG & \xrightarrow{T_e\phi} & T_eH \end{array}$$

We now define the notion of **Lie algebra**. Note that in general this definition does not need to have any connection with Lie groups. However, the fact that there is such a connection in specific cases will become clear in the next section.

Definition 3.3.3. A real Lie algebra is a real linear space $\mathfrak{a}$ equipped with a bilinear map $[\cdot,\cdot] : \mathfrak{a} \times \mathfrak{a} \to \mathfrak{a}$ such that for all $X, Y, Z \in \mathfrak{a}$ we have

(1) $[X,Y] = -[Y,X]$ (anti-symmetry)

(2) $[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$ (Jacobi-identity)

Note that:

- condition (1) may be replaced by the equivalent condition
  (1') $[Z,Z] = 0$ for all $Z \in \mathfrak{a}$: without loss of generality we can assume $Z = X + Y$. Then, using (1),
  $[Z,Z] = [X+Y, X+Y] = [X,X] + [Y,X] + [Y,X] + [Y,Y] =$
  $[X,Y] + [Y,X] = [X,Y] - [X,Y] = 0$,
  and the claim follows.

- In view of antisymmetry (1), condition (2) may be replaced by the equivalent condition (2') $[[X,Y],Z] = [X,[Y,Z]] - [Y,[X,Z]]$. Another equivalent form of the Jacobi identity is given by the Leibniz type rule $[X,[Y,Z]] = [[X,Y],Z] + [Y,[X,Z]]$.

Definition 3.3.4. Let $\mathfrak{a}, \mathfrak{b}$ be Lie algebras. A **Lie algebra homomorphism** from $\mathfrak{a}$ to $\mathfrak{b}$ is a linear map $\phi : \mathfrak{a} \to \mathfrak{b}$ such that

$$\phi([X,Y]_{\mathfrak{a}}) = [\phi(X), \phi(Y)]_{\mathfrak{b}} \tag{4}$$

for all $X, Y \in \mathfrak{a}$.

3.4 Actions and representations

To understand what a representation is we first need some theory about group actions.

Definition 3.4.1. Let M be a set and G a group. A **(left) action** of G on M is a map $\alpha : G \times M \to M$ such that

(a) $\alpha(g_1, \alpha(g_2, m)) = \alpha(g_1 g_2, m) \quad (m \in M,\ g_1, g_2 \in G)$;

(b) $\alpha(e, m) = m \quad (m \in M)$.

Instead of the cumbersome notation α we usually exploit the notation $g \cdot m$ or gm for $\alpha(g, m)$. Then the above rules (a) and (b) become:
$g_1 \cdot (g_2 \cdot m) = (g_1 g_2) \cdot m$, and $e \cdot m = m$.

If $g \in G$, then we sometimes use the notation α_g for the map $m \mapsto \alpha(g, m) = gm,\ \ M \to M$. From (a) and (b) we see that α_g is a bijection with inverse map equal to $\alpha_{g^{-1}}$. Let $\mathrm{Bij}(M)$ denote the set of bijections from M onto itself. Then $\mathrm{Bij}(M)$, equipped with the composition of maps, is a group. According to (a) and (b), the map $\alpha : g \mapsto \alpha_g$ is a group homomorphism of G into $\mathrm{Bij}(M)$.

Remark 3.4.1. Similarly, a **right action** of a group G on a set M is defined to be a map $\beta : M \times G \to M,\ \ (m, g) \mapsto mg$, such that $me_G = m$ and $(mg_1)g_2 = m(g_1 g_2)$ for all $m \in M$ and $g_1, g_2 \in G$. Notice that these requirements on β are equivalent to the requirement that the map $\beta^{\vee} : G \times M \to M$ defined by $\beta^{\vee}(g, m) = mg^{-1}$ is a left action. Thus, all results for left actions have natural counterparts for right actions.

As a next step we concentrate on continuous actions. This is most naturally done for topological groups.

Definition 3.4.2. A **topological group** is a group G equipped with a topology such that the multiplication map $\mu : G \times G \to G,\ \ (x, y) \mapsto xy$ and the inversion map $\iota : G \times G,\ \ x \mapsto x^{-1}$ are continuous.

Hence it is important to note that a Lie group is in particular a topological group.

Definition 3.4.3. Let G be a topological group. By a **continuous left action** of G on a topological space M we mean an action $\alpha : G \times M \to M$ that is continuous as a map between topological spaces. A (left) G-space is a topological space equipped with a continuous (left) G-action.

Sets of the form mG $(m \in M)$ are called **orbits** of the action α. Note that for two orbits m_1G, m_2G either $m_1G = m_2G$ or $m_1G \cap m_2G = \emptyset$. Thus, the orbits constitute a partition of M. The set of all orbits, called the **orbit space** , is denoted by M/G. Let us denote the canonical projection $M \to M/G,\ m \mapsto mG$ by π.

The orbit space $X = M/G$ is equipped with the quotient topology. This is the finest topology for which the map $\pi : M \to M/G$ is continuous. Thus, a subset O of X is open if and only if its preimage $\pi^{-1}(O)$ is open in M.

The action of G on M is called **transitive** if it has only one orbit, the full manifold M. In this case the G-space M is said to be a homogeneous space for G.

We are now ready to define a representation, a notion that is essential during this thesis. In the rest of this subsection G will always be a Lie group.

Definition 3.4.4. Let V be a locally convex space. A **continuous representation** $\pi = (\pi, V)$ of G in V is a continuous left action $\pi : G \times V \to V$, such that $\pi(x) : v \mapsto \pi(x)v = \pi(x, v)$ is a linear endomorphism of V, for every $x \in G$. The representation is called **finite dimensional** if $\dim V < \infty$.

Remark 3.4.2. If G is a group, and V a linear space, one defines a representation of G in V similarly, but without the requirement of continuity.

We quote without proof (for the proof the reader is referred to [3, p.69]):

Lemma 3.4.1. Let (π, V) be a finite dimensional representation of G. If π is continuous, then π is smooth.

The next definition contains an important property of continuous representations.

Definition 3.4.5. Let π be a representation of G in a linear space V. By an **invariant subspace** we mean a linear subspace $W \subset V$ such that $\pi(x)W \subset W$ for every $x \in G$. A continuous representation π of G in a complete locally convex space V is called **irreducible** , if 0 and V are the only closed invariant subspaces of V.

Besides, the notion of a representation of a Lie group there also exists the notion of a representation of a Lie algebra.

Definition 3.4.6. Let $\mathfrak{g}$ be a Lie algebra. A **representation of $\mathfrak{g}$** in a complex linear space V is a bilinear map $\iota \times V \to V$, $(X, v) \mapsto Xv$, such that

$$[X, Y]v = XYv - YXv$$

for all $X, Y \in \mathfrak{g}$ and $v \in V$. In other words, the map $X \mapsto X\cdot$ is a Lie algebra homomorphism from $\mathfrak{g}$ into $\mathrm{End}(V)$.

Earlier in this section we already defined what it means for two atlases to be equivalent. We end this section by stating what it means for two representations to be equivalent.

Definition 3.4.7. If $(\pi_j, v_j)(j = 1, 2)$ are continuous representations of G in complete locally convex spaces, then a continuous linear map $T : V_1 \to V_2$ is said to be **equivariant**, or **intertwining** if the following diagram commutes for every $x \in G$:

$$\begin{array}{ccc} V_1 & \xrightarrow{T} & V_2 \\ \uparrow{\scriptstyle \pi_1(x)} & & \uparrow{\scriptstyle \pi_2(x)} \\ V_1 & \xrightarrow{T} & V_2 \end{array}$$

The representations π_1 and π_2 are said to be **equivalent** if there exists a linear isomorphism T from V_1 onto V_2 which is equivariant.

If the above representations are finite dimensional, then one does not need to require T to be continuous, since every linear map $V_1 \to V_2$ has this property.

4 Coadjoint orbits

-Isaac Newton encrypted his discoveries in analysis in the form of an anagram that deciphers to the sentence, 'It is worthwhile to solve differential equations'. Accordingly, one can express the main idea behind the orbit method by saying 'It is worthwhile to study coadjoint orbits'.-
A.A. Kirillov ([1])

The notion of coadjoint orbits is the main ingredient of the orbit method. It is also the most important new mathematical object that has been brought into consideration in connection with the orbit method.

4.1 The coadjoint representation

Let G be a Lie group. Unless stated otherwise, we will assume G to be a matrix group, that is, a subgroup and at the same time a smooth submanifold of $GL(n, \mathbb{R})$. Let $\mathfrak{g} = \mathrm{Lie}(G)$ be the associated Lie algebra. Let $x \in G$. Then the translation maps $l_x : G \to G \;\; y \mapsto xy$ and $r_x : G \to G \;\; y \mapsto yx$ are diffeomorphisms from G onto itself, hence they are automorphisms (the notions of diffeomorphism and isomorphism are identical for Lie groups). As a consequence the conjugation map $C_x = l_x \circ r_x^{-1} : G \to G \;\; y \mapsto xyx^{-1}$ is an automorphism too, also called an inner automorphism. This map fixes the neutral element e. So we can define its tangent map at e which is a linear automorphism of $T_eG = \mathrm{Lie}(G)$ (the last identification will be justified later in this section). Thus, because the automorphisms form a group under composition of morphisms, we have $T_eC_x \in GL(T_eG)$.

Definition 4.1.1. If $x \in G$ we define $\mathrm{Ad}(x) \in GL(T_eG)$ by $\mathrm{Ad}(x) \equiv T_eC_x$. The map Ad: $G \to GL(T_eG)$ is called the **adjoint representation** of G in T_eG.

Because here G is a matrix group the Lie algebra $\mathfrak{g}$ is a subspace of $\mathrm{Mat}(n, \mathbb{R})$. Hence the adjoint representation is simply matrix conjugation:

$$\mathrm{Ad}(g)X = g \cdot X \cdot g^{-1}, \quad X \in \mathfrak{g}, \quad g \in G. \tag{5}$$

Indeed Ad is a representation: for $g_1, g_2 \in G, X \in \mathfrak{g}$ we have
$\mathrm{Ad}(g_1)\mathrm{Ad}(g_2)X = \mathrm{Ad}(g_1)(g_2Xg_2^{-1}) = g_1g_2Xg_2^{-1}g_1^{-1} =$
$g_1g_2X(g_1g_2)^{-1} = \mathrm{Ad}(g_1g_2)X$.

Now consider $\mathfrak{g}^* \equiv \{f : \mathfrak{g} \to \mathbb{R} | \text{f linear}\}$, the vector space dual to $\mathfrak{g}$. For any linear representation (π, V) of a group G one can define a dual representation (π^*, V^*) in the dual space V^*:

$$\pi^*(g) : V^* \to V^* \quad \pi^*(g) \equiv \pi(g^{-1})^* \tag{6}$$

where the asterisk in the right-hand side means the dual operator in V^* defined by:

$$\langle \pi(g^{-1})^* f, v\rangle \equiv \langle f, \pi(g^{-1})v\rangle \qquad \text{for any } \ v \in V, f \in V^* \tag{7}$$

and by $\langle f, v\rangle$ we denote the value of the linear functional f on a vector v. In particular:

Definition 4.1.2. Let $X \in \mathfrak{g}, F \in \mathfrak{g}^*$. Then the coadjoint representation $\mathrm{Ad}^* : G\times\mathfrak{g}^* \to \mathfrak{g}^*$ of G in $\mathfrak{g}^*$ is defined by $\mathrm{Ad}^*(g) : \mathfrak{g}^* \to \mathfrak{g}^* \quad \mathrm{Ad}^*(g) \equiv \mathrm{Ad}(g^{-1})^*$. So by definition, $\langle \mathrm{Ad}(g^{-1})^* F, X\rangle = \langle F, \mathrm{Ad}(g^{-1})X\rangle$.

Indeed Ad^* is a representation:
$\langle \mathrm{Ad}^*(g_1)\mathrm{Ad}^*(g_2)F, X\rangle = \langle \mathrm{Ad}(g_1^{-1})^*\mathrm{Ad}(g_2^{-1})^* F, X\rangle$
$= \langle \mathrm{Ad}(g_2^{-1})^* F, \mathrm{Ad}(g_1^{-1})X\rangle = \langle F, \mathrm{Ad}(g_2^{-1})\mathrm{Ad}(g_1^{-1})X\rangle$
$= \langle F, \mathrm{Ad}(g_2^{-1}g_1^{-1})X\rangle = \langle F, \mathrm{Ad}((g_1g_2)^{-1})X\rangle$
$= \langle \mathrm{Ad}((g_1g_2)^{-1})^* F, X\rangle = \langle \mathrm{Ad}^*(g_1g_2)F, X\rangle$.

Now the groups under consideration are matrix groups. What specific form does the coadjoint representation take for matrix groups? The space $\mathrm{Mat}(n, \mathbb{R})$ has a bilinear form

$$\mathrm{S} : \mathrm{Mat}(n, \mathbb{R}) \times \mathrm{Mat}(n, \mathbb{R}) \to \mathbb{R} \quad (A, B) \mapsto \mathrm{tr}(AB), \tag{8}$$

where tr(-) is the operation of taking the trace of a matrix. It is:

- non-degenerate (meaning $S(A, B) = 0 \ \forall B \in \mathrm{Mat}(n, \mathbb{R}) \Rightarrow A = 0$): first we will show $\mathrm{tr}(A^T A) = 0 \Leftrightarrow A = 0$. Namely, $\mathrm{tr}\big(A^T A\big) = \sum_{i,j=1}^n A^T_{ij} A_{ji} = \sum_{i,j=1}^n (A_{ji})^2 = 0 \Leftrightarrow (A_{ji})^2 = 0 \ \forall i, j \Leftrightarrow A = 0$. Hence $S(A, B) = \mathrm{tr}(AB) = \mathrm{tr}(BA) = 0 \ \forall B \Rightarrow \mathrm{tr}(A^T A) = 0 \Rightarrow A = 0$ (matrix with all entries zero).

- invariant under conjugation: let $C \in \mathrm{Mat}(n, \mathbb{R})$. Then $\mathrm{tr}(CABC^{-1}) = \mathrm{tr}(C^{-1}CAB) = \mathrm{tr}(AB)$. Hence S is invariant under conjugation.

The bilinear form S gives an identification between $\mathrm{Mat}(n, \mathbb{R})$ and $\mathrm{Mat}(n, \mathbb{R})^*$ via the map $\beta : A \to S(A, -) \in \mathrm{Mat}(n, \mathbb{R})^*$ defined by $\beta(A)(B) = S(A, B)$.

Now we want to show that the space $\mathfrak{g}^*$, dual to te subspace $\mathfrak{g} \subset$Mat$(n,\mathbb{R})$ can be identified with the quotient space Mat$(n,\mathbb{R})/\mathfrak{g}^\perp$, where the sign $\perp$ means the orthogonal complement with respect to the the form S:

$$\mathfrak{g}^\perp = \{A \in \mathrm{Mat}(n,\mathbb{R}) | S(A, \iota(B)) = 0 \ \ \forall B \in \mathfrak{g}\} \tag{9}$$

where ι is the inclusion map from $\mathfrak{g}$ to Mat$(n,\mathbb{R})$.

To start with the following table comparing properties of a map and its dual is good to have in mind:

f	f^*
isomorphism	isomorphism
zero image	zero image
injective	surjective
surjective	injective

Now consider the following short exact sequence of vector spaces ($\pi \circ \iota = 0$):

$A \overset{\iota}{\hookrightarrow} B \overset{\pi}{\to} C$

with ι an injective and π a surjective map. Then $C \cong B/\iota(A)$. Furthermore $\pi \circ \iota = 0 \Rightarrow (\pi \circ \iota)^* = \iota^* \circ \pi^* = 0$ (see the second property quoted in the table above). Hence

$A^* \overset{\iota^*}{\leftarrow} B^* \overset{\pi^*}{\hookleftarrow} C^*$

is a short exact sequence of vector spaces with ι^* a surjective and π^* an injective map (see the third and fourth property quoted in the table above). And $A^* \cong B^*/\pi^*(C^*)$ (where the first property quoted in the table is used). Now define:

$$A^\perp \equiv \{\phi \in B^* | \phi(\iota(a)) = 0 \ \ \forall a \in A\}. \tag{10}$$

We will show that $\pi^*(C^*) = A^\perp$ from which follows $A^* \cong B^*/A^\perp$. Indeed:

- $\pi^*(C^*) \subset A^\perp$: take $f \in C^*$, i.e. $f : C \to \mathbb{R}$, then $\pi^*(f) : B \to \mathbb{R}$. So for $a \in A$ we have $(\pi^*(f))(\iota(a)) \equiv (f \circ \pi)(\iota(a)) = (f \circ \pi \circ \iota)(a) = 0$, because $\pi \circ \iota = 0$. Hence $f \in A^\perp$.

- $A^\perp \subset \pi^*(C^*)$: take $g \in A^\perp$. Then $g \in B^*$ and $g(\iota(a)) = 0 \ \forall a \in A$. Now for $f \in C^*$ and $b \in B$ define $f([b]) \equiv g(b)$, where $[b] \equiv (b + \iota(a)) \in C(a \in A)$, is the equivalence class of an element $b \in B$. This definition is independent of the choice of $b \in [b]$. Namely,

$g(b+\iota(a)) = g(b) + g(\iota(a)) = g(b)$, because $g \in A^{\perp}$. It is left to check that $g = \pi^*(f)$, because then $g \in \pi^*(C^*)$. Indeed $g(b) = f([b]) = f(\pi(b)) = (f \circ \pi)(b) \equiv \pi^*(f)(b)$.

From $\pi^*(C^*) \subset A^{\perp}$ and $A^{\perp} \subset \pi^*(C^*)$ follows $\pi^*(C^*) = A^{\perp}$.

Now take $A = \mathfrak{g}$, $B = \mathrm{Mat}(n, \mathbb{R})$, ι the inclusion map from $\mathfrak{g}$ to $\mathrm{Mat}(n, \mathbb{R})$ and π the projection map from $\mathrm{Mat}(n, \mathbb{R})$ to $\mathrm{Mat}(n, \mathbb{R})/\iota(\mathfrak{g})$. Indeed $\pi \circ \iota = 0$ for this choice, by definition of the quotient space $\mathrm{Mat}(n, \mathbb{R})/\iota(\mathfrak{g})$. It follows by the previous argument that $\mathfrak{g}^* \cong \mathrm{Mat}(n, \mathbb{R})^*/\mathfrak{g}^{\perp}$. Here is used $S(A, \iota(B)) = \beta(A)(\iota(B))$ and $\beta(A) \in \mathrm{Mat}(n, \mathbb{R})^*$. Finally, because the bilinear form S is non-degenerate, the map β is an isomorphism. Hence $\mathrm{Mat}(n, \mathbb{R}) \cong \mathrm{Mat}(n, \mathbb{R})^*$. We arrive at $\mathfrak{g}^* \cong \mathrm{Mat}(n, \mathbb{R})^*/\mathfrak{g}^{\perp} \cong \mathrm{Mat}(n, \mathbb{R})/\mathfrak{g}^{\perp}$.

In practice this quotient space $\mathrm{Mat}(n, \mathbb{R})/\mathfrak{g}^{\perp}$ is often identified with a subspace $V \subset \mathrm{Mat}(n, \mathbb{R})$ that is transversal to $\mathfrak{g}^{\perp}$ and has complementary dimension. So we can write $\mathrm{Mat}(n, \mathbb{R}) = V \oplus \mathfrak{g}^{\perp}$. Let p_V be the projection of $\mathrm{Mat}(n, \mathbb{R})$ onto V parallel to $\mathfrak{g}^{\perp}$. Hence for matrix groups the coadjoint representation takes the form:

$$\mathrm{Ad}^*(g) : \mathfrak{g}^* \cong V \to \mathfrak{g}^* \cong V \quad F \mapsto p_V(gFg^{-1}), \quad g \in G = \mathrm{Mat}(n, \mathbb{R}). \tag{11}$$

Remark 4.1.1. If we could choose $V \cong \mathfrak{g}^*$ invariant under $\mathrm{Ad}(g)$, $g \in G$, then $gFg^{-1} \in V \cong \mathfrak{g}^*, F \in V \cong \mathfrak{g}^*$, hence we can drop the projection in the above definition of the coadjoint representation. This can be done if $\mathfrak{g}$ is semisimple.

We can now define the notion of a **coadjoint orbit** :

Definition 4.1.3. Given $F \in \mathfrak{g}^*$. The **coadjoint orbit** $\mathfrak{O}(F)$ is the image of the map $\kappa : G \to \mathfrak{g}^* \quad g \mapsto \mathrm{Ad}^*(g)F$.

This section shall be concluded with the introduction of the so-called infinitesimal version of the coadjoint representation. From the definition of C_x it follows that $C_e = I_G$. Hence $\mathrm{Ad}(e) = I_{T_eG}$. Hence, using $T_I GL(T_eG) = \mathrm{End}(T_eG)$, we can define the map tangent to Ad:

Definition 4.1.4. The linear map ad: $T_eG \to \mathrm{End}(T_eG)$ is defined by ad $\equiv T_e\mathrm{Ad}$

Furthermore,

Definition 4.1.5. We define the Lie bracket $[,] : T_eG \times T_eG \to T_eG$ by $[X, Y] \equiv (\mathrm{ad}X)Y$.

We will again derive what specific form the Lie bracket takes for matrix groups. First of all, note that by the chain rule

$$\mathrm{ad}(X) = \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}(\exp tX), \quad t \in \mathbb{R} \tag{12}$$

where is used the exponential map defined in section 3. We already knew that the adjoint representation takes the form:

$$\mathrm{Ad}(g)X = g \cdot X \cdot g^{-1}, \quad X \in \mathfrak{g}, \quad g \in G$$

Substituting $g = \exp tY$, $Y \in \mathfrak{g}$, and differentiating the resulting expression with respect to t at $t = 0$ we obtain, using (8),

$$[Y, X] = (\mathrm{ad}Y)X = \left.\frac{d}{dt}\right|_{t=0} \exp^{tY} X \exp^{-tY} = YX - XY, \tag{13}$$

the commutator bracket of X and Y.

Hence the Lie bracket:

- is a bilinear map,
- is anti-symmetric: $[X, Y] = XY - YX = -(YX - XY) = -[Y, X]$,
- satisfies the Jacobi identity:
  $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] =$
  $[X, YZ - ZY] + [Y, ZX - XZ] + [Z, XY - YX] =$
  $[X, YZ] - [X, ZY] + [Y, ZX] - [Y, XZ] + [Z, XY] - [Z, YX] =$
  $XYZ - YZX - XZY + ZYX + YZX - ZXY - YXZ + XZY + ZXY -$
  $XYZ - ZYX + YXZ = 0$, $Z \in \mathfrak{g}$.

It follows that T_eG together with the Lie bracket is a Lie algebra. This justifies the identification of T_eG with $\mathfrak{g}$, the Lie algebra of G, already used earlier in this subsection. The Lie bracket is a Lie algebra representation (also called the infitesimal representation associated with the representation of the Lie group) of $\mathfrak{g}$ in $\mathrm{End}(\mathfrak{g})$, which follows from the definition of a Lie algebra representation given in section 3. Namely, for $Z \in \mathfrak{g}$, $[X, Y]\mathrm{ad}(Z) = (XY - YX)\mathrm{ad}(Z) = XY\mathrm{ad}(Z) - YX\mathrm{ad}(Z)$.

We are now ready to define the infinitesimal version of the coadjoint representation:

Definition 4.1.6. Let $X, Y \in \mathfrak{g}, F \in \mathfrak{g}^*$. Then $\mathrm{ad}^* : \mathfrak{g} \times \mathfrak{g}^* \to \mathfrak{g}^*$ is defined by $\mathrm{ad}^*(X) : \mathfrak{g}^* \to \mathfrak{g}^* \quad \mathrm{ad}^*(X) \equiv (-\mathrm{ad}(X))^*$. So by definition, $\langle(-\mathrm{ad}(X))^* F, Y\rangle = \langle F, -\mathrm{ad}(X)Y\rangle = \langle F, -[X,Y]\rangle = \langle F, [Y,X]\rangle$.

For matrix groups it takes the form

$$\mathrm{ad}^*(X)F = p_V([X,F]) \text{ for } X \in \mathfrak{g}, \quad F \in V \cong \mathfrak{g}^*. \tag{14}$$

Again, if $\mathfrak{g}$ is semisimple, the projection can be dropped in the equation above.

4.2 The coadjoint orbits of SU(2)

In this subsection we will apply the theory of the coadjoint orbits to the Lie group SU(2), that is, we will argue what the coadjoint orbits of SU(2) are. The approach in this subsection will be less formal than in previous subsections. Because SU(2) is a relatively simple application of the theory it will give a good intuition for the notion of coadjoint orbits.

4.2.1 Deriving the coadjoint orbits by finding conjugation-invariant functions on $\mathfrak{su}(2)^*$

First of all, SU(2) is the matrix group defined by

$$\mathrm{SU}(2) \equiv \{x \in \mathrm{Mat}(2,\mathbb{C}) | x^\dagger x = \mathrm{I}, \det(x) = 1\}. \tag{15}$$

From the conditions on x contained in the definition we can derive the form of a matrix in SU(2). Let us first take a general matrix in $\mathrm{Mat}(2,\mathbb{C})$. It has the form

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

with $\alpha, \beta, \gamma, \delta \in \mathbb{C}$. By the definition of SU(2) we have that $x^\dagger = x^{-1}$. At the level of matrices this implies, using $\det(x) = \alpha\delta - \beta\gamma = 1$, that

$$\begin{pmatrix} \alpha^* & \gamma^* \\ \beta^* & \delta^* \end{pmatrix} = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}$$

using Cramer's rule to invert the matrix. Hence it follows that
$\alpha^* = \delta, \gamma^* = -\beta, \beta^* = -\gamma$ and $\delta^* = \alpha$. So at the level of matrices we can write

$$\mathrm{SU}(2) = \left\{ \begin{pmatrix} \alpha & -\beta^* \\ \beta & \alpha^* \end{pmatrix} \middle| (\alpha, \beta) \in \mathbb{C}^2, |\alpha|^2 + |\beta|^2 = 1 \right\}. \tag{16}$$

The next step is to calculate the Lie algebra of SU(2), which we will denote by $\mathfrak{su}(2)$. We can write $x = \exp^{\epsilon X}$ for $X \in \mathfrak{g}, \epsilon \in \mathbb{R}$ small and where exp is the exponential map defined in subsection 3.3. Taylor expanding this expression around ϵ and keeping terms up to first order in ϵ we have $x = \mathrm{I} + \epsilon X$, so the condition $x^\dagger x = \mathrm{I}$ translates into

$$(\mathrm{I}+\epsilon X)^\dagger(\mathrm{I}+\epsilon X) = \mathrm{I} \Leftrightarrow \mathrm{I}+\epsilon(X+X^\dagger)+O(\epsilon^2) = \mathrm{I} \Leftrightarrow X+X^\dagger = 0 \Leftrightarrow X = -X^\dagger.$$

Furthermore the condition $\det(x) = \mathrm{I}$ translates into

$$\det(x) = \exp^{\log(\det(x))} = \exp^{\mathrm{tr}\log(x)} = \exp^{\mathrm{tr}\log\left(\exp^{\epsilon X}\right)} = \exp^{\mathrm{tr}(\epsilon X)} = \exp^{\epsilon\mathrm{tr}(X)} \cong \mathrm{I} + \epsilon \mathrm{tr} X = \mathrm{I} \Leftrightarrow \mathrm{tr} X = 0.$$

Hence,

$$\mathfrak{su}(2) = \{X \in \mathrm{Mat}(2, \mathbb{C}) | X = -X^\dagger, \mathrm{tr} X = 0\}. \tag{17}$$

Again, from the conditions on X we can derive the form of a matrix in $\mathfrak{su}(2)$. Take the same general matrix in $\mathrm{Mat}(2, \mathbb{C})$ as before. At the level of matrices the condition $X = -X^\dagger$ becomes

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} -\alpha^* & -\gamma^* \\ -\beta^* & -\delta^* \end{pmatrix}.$$

Hence it follows that
(i) $\alpha^* = -\alpha$, (ii) $\delta^* = -\delta$, (iii) $-\beta = \gamma^*$ and (iv) $-\beta^* = \gamma$.
If we write $\alpha = a + ib, \beta = c + id, \gamma = e + if, \delta = g + ih$ and note that the conditions (iii) and (iv) are equivalent we deduce that
$a = 0, g = 0, d = f, c = -e$.
Finally $\mathrm{tr}(X) = 0$ implies that $b = -h$. We conclude that at the level of matrices

$$\mathfrak{su}(2) = \left\{ \begin{pmatrix} ib & c + id \\ -c + id & -ib \end{pmatrix} \middle| b, c, d \in \mathbb{R} \right\}. \tag{18}$$

It follows that $\mathfrak{su}(2)$ is determined in terms of three free real parameters b, c, d and hence is isomorphic to $\mathbb{R}^3$.

The Lie algebra $\mathfrak{su}(2)$ is semisimple. To show this, it suffices to show that $\mathfrak{su}(2)$ has an invariant inner product, since $SU(2)$ is a compact Lie group (see [31]). An inner product $\langle, \rangle$ is invariant under the action of $\mathfrak{su}(2)$ if, for $X, Y, Z \in \mathfrak{su}(2)$ ([32]),

$$\langle [X, Y], Z\rangle = -\langle Y, [X, Z]\rangle, \tag{19}$$

where $[,]$ denotes the Lie bracket. Since $\mathfrak{su}(2)$ is a matrix group, the Lie bracket $[,]$ is just a commutator bracket. Consequently, an invariant inner product is given by

$$\langle X, Y \rangle = \mathrm{tr}(XY), \tag{20}$$

for $X, Y \in \mathfrak{su}(2)$. Indeed,

$\langle [X, Y], Z\rangle = \mathrm{tr}([X, Y]Z) = \mathrm{tr}(Z[X, Y]) = \mathrm{tr}([Z, X]Y) =$
$\mathrm{tr}(Y[Z, X]) = -\mathrm{tr}(Y[X, Z]) = -\langle Y, [X, Z]\rangle,$

for $X, Y, Z \in \mathfrak{su}(2)$.

By remark 4.1.1. we can now drop the projection p_V in equation (7). Furthermore, we know by the previous subsection that on matrix groups $\mathrm{Mat}(n, \mathbb{R})$ there exists a bilinear form S and the corresponding earlier defined isomorphism β. This isomorphism β tells us that for SU(2) we have $\mathfrak{g}^* \cong \mathfrak{g}$. Take $F \in \mathfrak{g}^* \cong \mathfrak{g} \cong \mathbb{R}^3$. Now by definition 4.1.3. in order to find the coadjoint orbits of SU(2) we look for the points $F' \in \mathfrak{g}^* \cong \mathfrak{g} \cong \mathbb{R}^3$ for which

$$F' = gFg^{-1} \text{ for a certain } g \in \mathrm{SU}(2). \tag{21}$$

Let Q be a funtion Q: $\mathfrak{su}(2) \to \mathbb{C}$ which is conjugation invariant. Then (21) implies that $Q(F') = Q(gFg^{-1}) = Q(F)$ with $F' \in \Omega$, a coadjoint orbit of SU(2). Hence $Q(\Omega) = \text{constant} \in \mathbb{C}$. That is, different coadjoint orbits are characterized by different constants in the equation above. We have to check all conjugation invariant functions on $\mathfrak{su}(2)$ to find all coadjoint orbits of SU(2).

Take for Q the function Q: $\mathfrak{su}(2) \to \mathbb{C}, \quad F \mapsto \mathrm{tr}(F)$. Then Q is indeed conjugation invariant: $\mathrm{tr}(gFg^{-1}) = \mathrm{tr}(F)$, by cyclicity of the trace. But for SU(2) we already know that $\mathrm{tr}(F) = 0$ for all $F \in \mathfrak{su}(2)$, so in particular for points in Ω, so this gives no information about the specific points in Ω.

Now take for Q the function Q : $\mathfrak{su}(2) \to \mathbb{C} \quad F \mapsto \mathrm{tr}((F)^2)$. Indeed this Q is conjugation invariant:

$\mathrm{tr}(gFg^{-1}gFg^{-1}) = \mathrm{tr}(gF^2g^{-1}) = \mathrm{tr}(g^{-1}gF^2) = \mathrm{tr}((F)^2).$

Let us calculate $\mathrm{tr}((F')^2)$ for $F' \in \Omega$,using the matrix representation we have for elements in $\mathfrak{su}(2)$:

$$\mathrm{tr}((F')^2) = \mathrm{tr}\begin{pmatrix} ib & c+id \\ -c+id & -ib \end{pmatrix} \cdot \begin{pmatrix} ib & c+id \\ -c+id & -ib \end{pmatrix} =$$

$$\mathrm{tr}\begin{pmatrix} -b^2-c^2-d^2 & \ldots \\ \ldots & -b^2-c^2-d^2 \end{pmatrix} = \text{-2(b}^2+c^2+d^2) = \text{constant}.$$

Note that we can write without loss of generality, constant$/(-2) = r^2$ for a certain other constant $r \in \mathbb{C}$. So the condition on $b, c, d \in \mathbb{R}$ becomes:

$$b^2 + c^2 + d^2 = r^2. \tag{22}$$

Because the constants b, c, d are real, the expression $b^2 + c^2 + d^2$ must be real. Hence r^2 must be real too. Therefore, expression (21) tells us that at least some coadjoint orbits of SU(2) are spheres $S^2 \subset \mathbb{R}^3$ and we have a point as coadjoint orbit in case $b = c = d = 0$. We will now make clear why all coadjoint orbits of SU(2) are two-dimensional spheres or a point (one-dimensional).

Take for Q the function Q : $\mathfrak{su}(2) \to \mathbb{C}$ $F \mapsto \mathrm{tr}(FF^\dagger)$. Because $F \in \mathfrak{su}(2)$ it holds that $F = -(F)^\dagger$. Hence $\mathrm{tr}(FF^\dagger) = -\mathrm{tr}((F)^2)$. It follows that Q is conjugation invariant and the condition on $b, c, d \in \mathbb{R}$ corresponding with points on the coadjoint orbits becomes the same and we again get a point and two-dimensional spheres as coadjoint orbits.

Take for Q the function Q: $\mathfrak{su}(2) \to \mathbb{C}$ $F \mapsto \mathrm{tr}(F^N)$ for $N \in \mathbb{Z}$. We already proved $\mathrm{tr}(F^2) = -2(b^2 + c^2 + d^2)$ for $b, c, d \in \mathbb{R}$. With induction it follows that $\mathrm{tr}(F^N) = (-b^2 - c^2 - d^2)^{(N-1)/2}\mathrm{tr}(F)$ for N odd and $\mathrm{tr}((F)^N) = (-2)^{N/2}(b^2 + c^2 + d^2)^{N/2}$ for N even. So for N odd we get no information about the specific points in Ω, because $\mathrm{tr}(F) = 0$ for all $F \in \mathfrak{su}(2)$. For N even we can write

$$\mathrm{tr}(F^N) = (-2)^{N/2}(b^2 + c^2 + d^2)^{N/2} = (-2(b^2 + c^2 + d^2))^{N/2} = (\mathrm{tr}(F^2))^{N/2}.$$

Hence Q is conjugation invariant and the condition on $b, c, d \in \mathbb{R}$ again becomes $b^2 + c^2 + d^2 = r^2$ for $r \in \mathbb{R}$ and we get a point and two-dimensional spheres as coadjoint orbits.

The reader can convince himself/herself that all conjugation invariant functions on $\mathfrak{su}(2)$ are similar to the ones described above and hence lead to the same coadjoint orbits: two-dimensional spheres or points. The formal, rigourous proof of the fact that the two-dimensional spheres and a point exhaust the coadjoint orbits of SU(2) will not be given in this thesis.

4.2.2 Deriving the coadjoint orbits by comparing with SO(3)

Another way to see that the coadjoint orbits of SU(2) are two-dimensional spheres or a point is to establish an isomorphism between the Lie algebras su(2) and so(3) and then prove from this that the coadjoint orbits of both Lie groups are equivalent. Because SO(3) is the group of rotations in $\mathbb{R}^3$ it is then clear from visual point of view that its coadjoint orbits are two-dimensional spheres or a point.

Let a Lie group homomorphism $\phi : G \to H$ be given. Then we know from the theory given in subsection 3.3 that the map induced on the Lie algebras $\tilde{\phi} : \mathfrak{g} \to \mathfrak{h}$ is a (Lie algebra) homomorphism. This induces the following commutative diagram:

$$\begin{tikzcd}
G \times \mathfrak{g} \arrow[r, "\mathrm{Ad}"] \arrow[d, "\phi\times\tilde{\phi}"] & \mathfrak{g} \arrow[d, "\tilde{\phi}"] \\
H \times \mathfrak{h} \arrow[r, "\mathrm{Ad}"] & \mathfrak{h}
\end{tikzcd}$$

with Ad the adjoint representation. This diagram is equivalent to the following commutative diagram

$$\begin{tikzcd}
G \times \mathfrak{g} \arrow[rr, "\mathrm{Ad}"] \arrow[d, "\mathrm{Id}_G\times\tilde{\phi}"'] & & \mathfrak{g} \arrow[drr, "\tilde{\phi}"] & & \\
G \times \mathfrak{h} \arrow[rr, "\phi\times\mathrm{Id}_{\mathfrak{g}}"'] & & H \times \mathfrak{h} \arrow[rr, "\mathrm{Ad}"'] & & \mathfrak{h}
\end{tikzcd}$$

with Id_G and $\mathrm{Id}_{\mathfrak{h}}$ the identity mappings on G and $\mathfrak{g}$ respectively. The fact that these diagrams are equivalent is obvious from the fact that
$\phi \times \tilde{\phi} = (\phi \times \mathrm{Id}) \circ (\mathrm{Id} \times \tilde{\phi})$.

In general, if for G a group and V, W vector spaces the maps $\alpha_1 : G\times V \to V$, $\alpha_2 : G \times W \to W$ are continuous G-representations and a homomorphism $\phi : V \to W$ is given this induces the commutative diagram:

$$\begin{tikzcd}
G \times V \arrow[r, "\alpha_1"] \arrow[d, "\mathrm{Id}_G\times\phi"] & V \arrow[d, "\phi"] \\
G \times W \arrow[r, "\alpha_2"] & W
\end{tikzcd}$$

We then say that ϕ is a **homomorphism of G-representations** . In this situation it is generally true that G-orbits in V are sent to G-orbits in W.

So in the particular case considered above we have $V = \mathfrak{g}$, $W = \mathfrak{h}$, $\phi = \tilde{\phi}$, $\alpha_1 = \mathrm{Ad}$ and $\alpha_2 = \mathrm{Ad} \circ (\phi \times \mathrm{Id})$. Indeed α_2 is a continuous representation in this case, because the composition of a continuous representation with a homomorphism is again a continuous representation. It follows that $\tilde{\phi}$ is a homomorphism of G-representations and G-orbits in $\mathfrak{g}$ are sent to G-orbits in $\mathfrak{h}$.

However, we are interested in the coadjoint orbits, not in the adjoint ones. If $\tilde{\phi} : \mathfrak{g} \to \mathfrak{h}$ is a homomorphism, then $\tilde{\phi}^* : \mathfrak{h}^* \to \mathfrak{g}^*$ is a homomorphism. The following commutative diagram is induced:

$$\begin{tikzcd}
G \times \mathfrak{g}^* \arrow[rr, "\mathrm{Ad}^*"] & & \mathfrak{g}^* & \\
G \times \mathfrak{h}^* \arrow[u, "\mathrm{Id}_G \times \tilde{\phi}^*"] \arrow[r, "\phi^* \times \mathrm{Id}_{\mathfrak{h}^*}"'] & H \times \mathfrak{h}^* \arrow[rr, "\mathrm{Ad}^*"'] & & \mathfrak{h}^* \arrow[ul, "\tilde{\phi}^*"']
\end{tikzcd}$$

We conclude that $\tilde{\phi}^*$ is a homomorphism of G-representations and G-orbits in $\mathfrak{h}^*$ are sent to G-orbits in $\mathfrak{g}^*$.

Now take $G = \mathrm{SU}(2)$ and $H = \mathrm{SO}(3)$. We will now show that $\mathfrak{su}(2) \cong \mathfrak{so}(3)$. Because $\mathfrak{su}(2)$ is the algebra of $X \in \mathrm{Mat}(2, \mathbb{C})$ with $X^\dagger = -X$ and $\mathrm{tr}(X) = 0$ one can see from this that as a real linear space $\mathfrak{su}(2)$ is generated by the elements

$$r_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad r_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad r_3 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Note that $r_j = i\sigma_j$, where $\sigma_1, \sigma_2, \sigma_3$ are the famous Pauli spin matrices. One readily verifies that $r_1^2 = r_2^2 = r_3^2 = -I$ and $r_1 r_2 = -r_2 r_1 = r_3$ and $r_2 r_3 = -r_3 r_2 = r_1$.

It follows from the above product rules that the commutator brackets are given by

- $[r_1, r_1] \equiv \mathrm{ad}(r_1) r_1 = 0, \quad [r_1, r_2] \equiv \mathrm{ad}(r_1) r_2 = r_1 r_2 - r_2 r_1 = 2r_3,$
  $[r_1, r_3] \equiv \mathrm{ad}(r_1) r_3 = r_1 r_3 - r_3 r_1 = r_1^2 r_2 + r_3^2 r_2 = -I r_2 - I r_2 = -2r_2.$
- $[r_2, r_1] = -[r_1, r_2] = -2r_3, \quad [r_2, r_2] = 0, \quad [r_2, r_3] = r_2 r_3 - r_3 r_2 = 2r_1.$
- $[r_3, r_1] = -[r_1, r_3] = 2r_2, \quad [r_3, r_2] = -[r_2, r_3] = -2r_1, \quad [r_3, r_3] = 0.$

From this it follows that the endomorphisms $\mathrm{ad}(r_j) \in \mathrm{End}(\mathfrak{su}(2))$ have the following matrices with respect to the basis r_1, r_2, r_3:

$$\mathrm{mat}(\mathrm{ad}\, r_1) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & 2 & 0 \end{pmatrix}, \quad \mathrm{mat}(\mathrm{ad}\, r_2) = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ -2 & 0 & 0 \end{pmatrix}, \quad \mathrm{mat}(\mathrm{ad}\, r_3) = \begin{pmatrix} 0 & -2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The above elements belong to

$$\mathfrak{so}(3) = \{X \in \mathrm{Mat}(3, \mathbb{R}) | X = -X^T\},$$

the Lie algebra of the group SO(3).

If $a \in \mathbb{R}^3$, then the exterior product map $X \mapsto a \times X$, $\mathbb{R}^3 \to \mathbb{R}^3$ has matrix

$$R_a = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}$$

with respect to the standard basis e_1, e_2, e_3 of $\mathbb{R}^3$. Clearly $R_a \in \mathfrak{so}(3)$.

Write $R_j = R_{e_j}$ for $j = 1, 2, 3$. Then by the above formulas for mat ad (r_j) we have

$$\mathrm{mat}(\mathrm{ad}\, r_j) = 2R_j \quad (j = 1, 2, 3).$$

We now define the map ϕ: $\mathrm{SU}(2) \to \mathrm{GL}(3, \mathbb{R})$ by $\phi(x) = \mathrm{mat}\,\mathrm{Ad}\,(x)$, the matrix being taken with respect to the basis r_1, r_2, r_3. Then ϕ is a homomorphism of Lie groups. The theory in subsection 3.3 gives that $\phi(\exp X) = \mathrm{mat}\, e^{\mathrm{ad}\, X} = e^{\mathrm{mat\ ad}\, X}$ from which we see that ϕ maps $\mathrm{SU}(2)_e$ into SO(3). Since SU(2) is connected, we have $\mathrm{SU}(2) = \mathrm{SU}(2)_e$, so that ϕ is a Lie group homomorphism from SU(2) to SO(3). The tangent map of ϕ is given by $\tilde{\phi} : \mathfrak{su}(2) \to \mathfrak{so}(3)$ $X \mapsto \mathrm{mat}\,\mathrm{ad}\, X$. It maps the basis $\{r_j\}$ of $\mathfrak{su}(2)$ onto the basis $\{2R_j\}$ of $\mathfrak{so}(3)$, hence is a linear isomorphism. So $\mathfrak{su}(2) \cong \mathfrak{so}(3)$.

It follows that ϕ is a local diffeomorphism at I, hence its image $\mathrm{im}(\phi)$ contains an open neighbourhood of I in SO(3). By homogeneity, $\mathrm{im}(\phi)$ is an open connected subgroup of SO(3), and we see that $\mathrm{im}(\phi) = \mathrm{SO}(3)_e$. The latter subgroup equals SO(3), by connectedness of the latter group. From this we conclude that $\phi : \mathrm{SU}(2) \to \mathrm{SO}(3)$ is a surjective group homomorphism. Hence $\mathrm{SO}(3) \cong \mathrm{SU}(2)/\ker(\phi)$. The kernel of ϕ may be computed as follows. If $x \in \ker(\phi)$, then $\mathrm{Ad}\,(x) = I$. Hence $xr_j = r_jx$ for $j = 1, 2, 3$.

From this one sees that $x \in \{-I, I\}$. Hence $\ker(\phi) = \{-I, I\}$. So $\mathrm{SU}(2)/\{\pm I\} \cong \mathrm{SO}(3)$.

So what do we have so far? We have that $\tilde{\phi} : \mathfrak{su}(2) \to \mathfrak{so}(3)\ \ X \mapsto \mathrm{ad}(X)$ is a linear isomorphism, hence $\tilde{\phi}^*$ is a linear isomorphism, hence in the commutative diagram above we can replace $\mathfrak{h}^* = \mathfrak{so}(3)^*$ by $\mathfrak{g}^* = \mathfrak{su}(2)^*$. So we have that $\tilde{\phi}^*$ is a homomorphism of SU(2)-representations and SU(2)-orbits in $\mathfrak{so}(3)^* \cong \mathfrak{su}(2)^*$ are send to SU(2)-orbits in $\mathfrak{su}(2)^* \cong \mathfrak{so}(3)^*$.

Moreover, we can now show that the SU(2)-orbits in $\mathfrak{so}(3)^* \cong \mathfrak{su}(2)^*$ are equivalent to the SO(3)-orbits in $\mathfrak{so}(3)^*$. Two points $x, y \in \mathfrak{so}(3)^*$ lie both on a SU(2)-orbit if $g \cdot x = y$, where $g \in \mathrm{SU}(2)$ and '$\cdot$' means 'acting'. We denote this by the equivalence relation $x \sim_{\mathrm{SU}(2)} y$. Because ϕ is a homomorphism $g \cdot x = y \Rightarrow \phi(g) \cdot x = y$. The fact that $\phi(g) \in SO(3)$ thus gives $x \sim_{\mathrm{SU}(2)} y \Rightarrow x \sim_{\mathrm{SO}(3)} y$. Now suppose that the points $x, y \in \mathfrak{so}(3)^*$ lie both on a SO(3)-orbit, that is, $h \cdot x = y$ for $h \in \mathrm{SO}(3)$. As proved earlier, ϕ is surjective, meaning $\exists\, g$ such that $\phi(g) = h$. This implies that $g \cdot x = y$. Hence $x \sim_{\mathrm{SO}(3)} y \Rightarrow x \sim_{\mathrm{SU}(2)} y$ and we are done.

In other words, the coadjoint orbits of SO(3) and SU(2) are equivalent.

5 Symplectic structure on coadjoint orbits

-One feature of coadjoint orbits is eye-catching when you consider a few examples: they always have an even dimension. This is not accidental, but has a deep geometric reason.-
A.A. Kirillov ([1, p.4])

5.1 The Kirillov form

Definition 5.1.1. A **symplectic manifold** (M, ω) is a smooth manifold M on which there exists a closed nondegenerate differential 2-form ω (called the symplectic form). In local coordinates:

$$\omega = \omega_{\mu\nu}(x)dx^\mu \wedge dx^\nu, \quad d\omega = 0,$$

where $\omega_{\mu\nu}$ is an anti-symmetric, invertible matrix.

In the appendix it is proven that the closedness $d\omega = 0$ of the symplectic form ω at the level of the matrix $\omega_{\mu\nu}$ translates to the condition

$$\partial_\kappa \omega_{\mu\nu} + \partial_\mu \omega_{\nu\kappa} + \partial_\nu \omega_{\kappa\mu} = 0,$$

which is known as the **Jacobi identity for symplectic forms** .

From the definition it follows that symplectic manifolds should have even dimension. Indeed,

$$\det(\omega_{\mu\nu}) = \det(-(\omega_{\mu\nu})^T) = (-1)^n \det((\omega_{\mu\nu})^T) = (-1)^n \det(\omega_{\mu\nu}),$$

with $\omega_{\mu\nu}$ an $n \times n$-matrix. In the first identity the anti-symmetry of $\omega_{\mu\nu}$ is used. In the second and third identity properties of the determinant are used. Consequently, if n is odd, $\det(\omega_{\mu\nu}) = 0$ and $\omega_{\mu\nu}$ does not satisfy the invertibility property anymore. Hence a symplectic manifold should be even-dimensional.

The important and beautiful fact that the coadjoint orbits of Lie groups (in our case matrix groups) are symplectic manifolds is encoded in the following theorem:

Theorem 5.1.1. On every coadjoint orbit Ω of a matrix group G, there exists a nondegenerate closed G-invariant differential 2-form B_Ω (called the Kirillov form) defined by $B_\Omega(F)(X, Y) \equiv \langle F, [X, Y] \rangle$ for $X, Y \in \mathfrak{g}, F \in \mathfrak{g}^*$.

Remark 5.1.1. Notice that although this theorem applies only to matrix groups, the result can also be proven for general Lie groups. The proof of this more general theorem can be found in for example [2, p.227-230]. Because in applications in (mathematical) physics the Lie groups are almost always matrix groups the above theorem suffices for the purposes of this thesis.

Proof. In this proof we will proceed as follows. We will first prove the existence of this differential form. Then we will prove that the differential form is nondegenerate, G-invariant and closed, respectively. This together establishes the claim of the theorem.

- Existence: Note that Ω is by definition a transitive G-space (being one orbit). So it follows from the Orbit-Stabilizer theorem (for the proof, see for example [5]) that $\Omega \cong G/\mathrm{Stab}_G(F)$ for $F \in \Omega$. Here $\mathrm{Stab}_G(F) \equiv \{g \in G | \ \mathrm{Ad}^*(g)F = F\}$.

  Take $X \in \mathfrak{g} \equiv \mathrm{Lie}(G)$. Then we know from the theory in subsection 3.1 that X is an equivalence class of curves $\gamma : [-\epsilon, \epsilon] \to G, \ \ 0 \mapsto e$, which we will denote as $[\gamma]$. Let an action of G on a manifold M be given. We can now construct a vector field on M. To achieve this we somehow need to map the interval $[-\epsilon, \epsilon]$ to the manifold M. A natural way to do this is as follows: $[-\epsilon, \epsilon] \xrightarrow{\tilde{\gamma}} G \times M \xrightarrow{\alpha} M \ \ 0 \mapsto (e, m) \mapsto m$ where $\tilde{\gamma} \equiv \iota_m \circ \gamma$ with $\iota_m : G \hookrightarrow G \times M, \ \ g \mapsto (g, m)$, the inclusion map and α the action map. So for a given $X \in \mathfrak{g}$ we can define $\xi_X \in \mathrm{Vect}(M)$ by $(\xi_X)_m \equiv [t \mapsto \gamma(t) \cdot m]$ where '$\cdot$' means 'acting on'.

  Now take for M the orbit space Ω, for the curve the specific realisation $t \mapsto \exp^{tX}$ with $t \in \mathbb{R}$ and for the action the coadjoint representation. Then for $F \in \Omega$

  $$\xi_X(F) = \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}^*\left(e^{tX}\right) F = \left.\frac{d}{dt}\right|_{t=0} p_V\left(e^{tX} F e^{-tX}\right) = p_V([X, F]). \tag{23}$$

  We will construct a differential two-form on Ω. At the point $F \in \Omega$ a differential two form gives an alternating bilinear map $B : T_F\Omega \times T_F\Omega \to \mathbb{R}$. Take $a, b \in T_F\Omega$. We can write $a = \xi_X(F)$, $b = \xi_Y(F)$ for $X, Y \in \mathrm{Lie}(\Omega) = \mathfrak{g}/\mathrm{Lie}(Stab_G(F))$. This follows from the fact that the surjective map $G \to \Omega$ given by the coadjoint representation induces a surjective map on the tangent spaces

$T_eG \equiv \mathfrak{g} \to T_e\Omega \equiv \mathfrak{g}/\mathrm{Lie}(\mathrm{Stab}_G(F))$ and the fact that these choices for a and b do not depend on the choices for $X, Y \in \mathrm{Lie}(\Omega)$. Namely, for $Z \in \mathrm{Lie}(Stab_G(F))$ we have

$\xi_{X+Z}(F) = p_V([X+Z,F]) = p_V([X,F]) + p_V([Z,F]) = p_V([X,F]) = \xi_X(F)$.

That $[Z,F] = 0$, and hence $p_V([Z,F]) = 0$, follows from restricting $\mathrm{ad}^*(Z)$ to $\Omega \cong G/Stab_G(F)$.

We are now in the positon to define $B(a,b) \equiv B_\Omega(F)(X,Y) \equiv \langle F, [X,Y]\rangle$. Indeed this definition does not depend on our choice for $X, Y \in \mathrm{Lie}(\Omega)$: take $Z \in \mathrm{Lie}(\mathrm{Stab}_G(F))$. Then

$B_\Omega(F)(Z,Y) = \langle F, [Z,Y]\rangle = \langle F, \mathrm{ad}(Z)Y\rangle$
$= \langle \mathrm{ad}^*(Z)F, Y\rangle = \langle p_V[Z,F], Y\rangle = 0,$

so it follows that,

$B_\Omega(F)(X+Z,Y) = \langle F,[X+Z,Y]\rangle = \langle F,[X,Y]\rangle + \langle F,[Z,Y]\rangle =$
$\langle F,[X,Y]\rangle = B_\Omega(F)(X,Y)$.

Indeed $B_\Omega(F)(X,Y)$ defines a differential two-form:

$B_\Omega(F)(X,Y) = \langle F,[X,Y]\rangle = -\langle F,[Y,X]\rangle = -B_\Omega(F)(Y,X)$.

Hence we have established the existence of the differential two-form $B_\Omega(F)(X,Y)$ on the coadjoint orbits of the matrix group G.

- Nondegeneracy: To prove that $B(a,b) \equiv B_\Omega(F)(X,Y)$ is nondegenerate is now easy. We want to prove that for $a \neq 0$ there exists a b such that $B(a,b) \neq 0$. This is equivalent to proving that for $X \notin \mathrm{Lie}(\mathrm{Stab}_G(F))$ there exists a Y such that $B_\Omega(F)(X,Y) \neq 0$. Take $Y \in \mathfrak{g}$. Then

  $B_\Omega(F)(X,Y) = \langle F,[X,Y]\rangle = \langle F, \mathrm{ad}(X)Y\rangle = \langle (\mathrm{ad}(X))^*F, Y\rangle =$
  $\langle -(\mathrm{ad}^*(X))F, Y\rangle = \langle -(p_V[X,F]), Y\rangle$.

  We already had that $[X,F] = 0 \Leftrightarrow X \in \mathrm{Lie}(\mathrm{Stab}_G(F))$. Hence $[X,F] \neq 0 \Leftrightarrow X \notin \mathrm{Lie}(\mathrm{Stab}_G(F))$. It follows that $\langle -(p_V[X,F]), Y\rangle \neq 0$ and nondegeneracy is established.

- G-invariance: $B_\Omega(F)(X,Y)$ being G-invariant means that $l_g^* B_\Omega(F)(X,Y) = B_\Omega(F)(X,Y)$ where l_g^* is the pull-back of l_g as defined in section 3. In general, a mathematical object α has a natural group action defined on it and we call the mathematical object G-invariant if $l_g^*\alpha = \alpha$. If α is a function f we get $l_g^*(f) : M \to \mathbb{R} \quad m \mapsto f(g \cdot M)$ and $l_g^* f = f \Leftrightarrow f(g \cdot m) = f(m)$. Here $\cdot$ means 'acting on'. This gives, using $B_\Omega(F)(X,Y) \in C^\infty(\Omega)$, that,

  $l_g^* B_\Omega(F)(X,Y) = B_\Omega(F)(X,Y) = \langle F, [X,Y] \rangle,$

  implying that $\langle F, [X,Y] \rangle = \langle \mathrm{Ad}^*(g)F, [\mathrm{Ad}(g)X, \mathrm{Ad}(g)Y] \rangle$.

  So this last equation is the one that needs to be checked. Indeed,

  $\langle \mathrm{Ad}^*(g)F, [\mathrm{Ad}(g)X, \mathrm{Ad}(g)Y] \rangle = \langle \mathrm{Ad}^*(g)F, [gXg^{-1}, gYg^{-1}] \rangle.$

  Furthermore,

  $[gXg^{-1}, gYg^{-1}] \equiv gXg^{-1}gYg^{-1} - gYg^{-1}gXg^{-1}$
  $= gXYg^{-1} - gYXg^{-1} = g[X,Y]g^{-1}.$

  Hence,

  $\langle \mathrm{Ad}^*(g)F, [gXg^{-1}, gYg^{-1}] \rangle = \langle \mathrm{Ad}^*(g)F, g[X,Y]g^{-1} \rangle =$
  $\langle (\mathrm{Ad}(g^{-1}))^* F, g[X,Y]g^{-1} \rangle = \langle F, \mathrm{Ad}(g^{-1})(g[X,Y]g^{-1}) \rangle =$
  $\langle F, g^{-1}g[X,Y]g^{-1}g \rangle = \langle F, [X,Y] \rangle.$

  This gives G-invariance of $B_\Omega(F)(X,Y)$.

- Closedness: We have to prove that $\mathrm{d}B_\Omega = 0$, where the definition of d is used as given in section 3. For notational convenience we drop the subscript Ω in B_Ω for a moment. The notation c.p. denotes cyclic permutation in X, Y, Z of the expression on the left hand side of it.

  Using the shape of B in local coordinates and taking $X, Y, Z \in \mathrm{Lie}(\Omega)$ gives

  $\mathrm{d}B(X,Y,Z) =$
  $\mathrm{d}\left(\sum_{i<j} B_{ij} \mathrm{d}x^i \wedge \mathrm{d}x^j\right)(X,Y,Z) =$
  $\left(\sum_k \sum_{i<j} \partial_k B_{ij} \mathrm{d}x^k \wedge \mathrm{d}x^i \wedge \mathrm{d}x^j\right)(X,Y,Z) =$
  $\sum_k \sum_{i<j} \partial_k B_{ij} (X^k Y^i Z^j - X^k Y^j Z^i + \text{c.p.}).$

We claim that $\mathrm{d}B(X,Y,Z) = X \cdot B(Y,Z) - B([X,Y],Z) + \text{c.p.}$. Here is used that for a function f and and a vector field v the expression $v \cdot f$ denotes the directional derivative of f in the direction v, that is $v \cdot f \equiv v^i \frac{\partial f}{\partial x^i}$. Let us check this claim by first calculating $X \cdot B(Y,Z) + \text{c.p.}$. We have

$$\begin{aligned} X \cdot B(Y,Z) + \text{c.p.} &= \\ \sum_k X^k \frac{\partial}{\partial X^k} \sum_{i<j} B_{ij}(Y^i Z^j - Y^j Z^i) + \text{c.p.} &= \\ \sum_{i<j} \sum_k \{X^k(\partial_k B_{ij}) Y^i Z^j + X^k B_{ij} \partial_k Y^i Z^j + X^k B_{ij} Y^i \partial_k Z^j\} & \\ - \sum_{i<j} \sum_k \{X^k(\partial_k B_{ij}) Y^j Z^i + X^k B_{ij} \partial_k Y^j Z^i + X^k B_{ij} Y^j \partial_k Z^i\} + \text{c.p.} & \end{aligned} \tag{24}$$

It follows that the terms containing $\partial_k B_{ij}$ match with the expression for $\mathrm{d}B(X,Y,Z)$ found at the beginning. So it is only left to check that the terms not containing $\partial_k B_{ij}$ cancel against $-B([X,Y],Z) + \text{c.p.}$. We have

$-B([X,Y],Z) + \text{c.p.} =$
$-\sum_{i<j} B_{ij}([X,Y]^i Z^j - [X,Y]^j Z^i) + \text{c.p.}$

Using $[X,Y] = X^i \partial_i Y^j - Y^i \partial_i X^j$ this equals

$$\begin{aligned} \sum_k \sum_{i<j} B_{ij}(&-X^k \partial_k Y^i Z^j + Y^k \partial_k X^i Z^j \\ &+ X^k \partial_k Y^j Z^i - Y^k \partial_k X^j Z^i) + \text{c.p.} \end{aligned} \tag{25}$$

Hence the terms not containing $\partial_k B_{ij}$ in (24) indeed cancel against the terms in (25). This verifies the claim.

Notice that because $B(Y,Z)$ is a C^∞-function, its directional derivative along the vector field X is equal to the Lie derivative $L_X B(Y,Z)$ (see subsection 3.1). The same holds for the cyclic permutations of $B(Y,Z)$. Using the Leibniz rule for the Lie derivative,

$X \cdot B(Y,Z) - B([X,Y],Z) =$
$(L_X B)(Y,Z) + B(L_X Y, Z) + B(Y, L_X Z) - B([X,Y],Z) =$
$B([X,Y],Z) + B(Y,[X,Z]) - B([X,Y],Z) = B(Y,[X,Z])$.

Here is used that $L_X B = 0$: earlier in this proof we showed that B is G-invariant. Write $g = \exp^{tA}$ for $t \in \mathbb{R}$, $A \in \mathfrak{g}$, $g \in G$. Then B being G-invariant means $l_g^*(B) = B$ with l_g the left translation map. So $l^*_{\exp(tA)}(B) = B \Leftrightarrow \frac{d}{dt}\big|_{t=0} l^*_{\exp(tA)}(B) \equiv L_{X_A}(B) = 0$ and we are done. One can apply similar mathematics on the cyclic permutations of the expression $X \cdot B(Y,Z) - B([X,Y],Z)$. We end up with the three terms $B(Y,[X,Z]), B(X,[Z,Y]), B(Z,[Y,X])$. Using the anti-symmetry of B and the Lie bracket, rewriting of these three terms gives:

(1) $B(Y,[X,Z]) = -B([X,Z],Y)$

(2) $B(X,[Z,Y]) = -B(X,[Y,Z]) = B([Y,Z],X)$

(3) $B(Z,[Y,X]) = -B(Z,[X,Y]) = B([X,Y],Z)$.

It follows that for $X,Y,Z \in \mathrm{Lie}(\Omega)$ we have (inserting the subscript Ω in B_Ω again):

$$
\begin{aligned}
&\mathrm{d}B_\Omega(X,Y,Z) = B_\Omega([X,Y],Z) - B_\Omega([X,Z],Y) + B_\Omega([Y,Z],X) = \\
&\langle F, [[X,Y],Z]\rangle - \langle F, [[X,Z],Y]\rangle + \langle F, [[Y,Z],X]\rangle = \\
&\langle F, [[X,Y],Z] - [[X,Z],Y] + [[Y,Z],X]\rangle = \\
&\langle F, -[Z,[X,Y]] + [Y,[X,Z]] - [X,[Y,Z]]\rangle = \\
&\langle F, -[Z,[X,Y]] - [Y,[Z,X]] - [X,[Y,Z]]\rangle,
\end{aligned}
$$

where the anti-symmetry of the Lie bracket is used in the last two identities. The last expression is indeed equal to zero by the Jacobi identity. Hence we have established that B_Ω is closed.

□

It follows that the coadjoint orbits of matrix groups are always even-dimensional.

5.2 The Kirillov form on the coadjoint orbits of SU(2)

This second subsection of section 5 will again be an application of the theory described in the foregoing subsection (as was the case in section 4) and will hopefully give more insight into this theory. Here the explicit 'local coordinates shape' of the Kirillov form on the physically interesting coajoint orbits of SU(2) (being the two-dimensional spheres) will be derived.

Before we come to this calculation we make a small excursion to the theory of Poisson manifolds and Poisson brackets.

Definition 5.2.1. A **Poisson manifold** is a differentiable mainifold M such that the the algebra $C^\infty(M)$ of smooth functions over M is equipped with a bilinear map $\{,\} : C^\infty(M) \times C^\infty(M) \to C^\infty(M)$ such that $\{,\}$

(1) is skew-symmetric, i.e. $\{f, g\} = -\{g, f\}$,

(2) obeys the Jacobi identity: $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$,

(3) is a derivation of $C^\infty(M)$ in its first argument:
$\{fg, h\} = f\{g, h\} + g\{f, h\}$ for all $f, g, h \in C^\infty(M)$.

The map $\{,\}$ is called the **Poisson bracket**.

In local coordinates the Poisson bracket can be written as

$$\{f, g\} \equiv -\frac{1}{2}\frac{\partial f}{\partial x^\mu}\omega^{\mu\nu}(x)\frac{\partial g}{\partial x^\nu}, \tag{26}$$

where $x \in M$ and $\omega^{\mu\nu}(x)$ is a matrix such that $\{f, g\}$ satisfies properties (1) and (2) in the above definition (so in particular it must be an anti-symmetric matrix by virtue of property 1). Expression (26) also satisfies property (3):

$\{fg, h\} = -\frac{1}{2}\frac{\partial(fg)}{\partial x^\mu}\omega^{\mu\nu}(x)\frac{\partial h}{\partial x^\nu} =$
$-f\frac{1}{2}\frac{\partial g}{\partial x^\mu}\omega^{\mu\nu}(x)\frac{\partial h}{\partial x^\nu} - g\frac{1}{2}\frac{\partial f}{\partial x^\mu}\omega^{\mu\nu}(x)\frac{\partial h}{\partial x^\nu} =$
$f\{g, h\} + g\{f, h\}$,
for all $f, g, h \in C^\infty(M)$.

Given a symplectic manifold (M, ω) and $f, g \in C^\infty(M)$. We can define the Poisson brackets in local coordinates by

$$\{f, g\} \equiv -\frac{1}{2}\frac{\partial f}{\partial x^\mu}\omega^{\mu\nu}(x)\frac{\partial g}{\partial x^\nu}, \tag{27}$$

where in this case $\omega^{\mu\nu}$ is the inverse of the matrix $\omega_{\mu\nu}$ corresponding to the symplectic form $\omega = \omega_{\mu\nu}(x)dx^\mu \wedge dx^\nu$. The proof that this definition of $\{,\}$ does indeed define a Poisson bracket is postponed untill subsection 6.1.

As a consequence, all symplectic manifolds are Poisson manifolds. The opposite need not be true. This is because on a Poisson manifold, the matrix $\omega^{\mu\nu}$, defining the Poisson bracket, need not be invertible.

Our procedure to calculate the explicit 'local coordinate' shape of the Kirillov form on the two-dimensional spheres will be to calculate the matrix $\omega^{\mu\nu}$ in (26) locally and show that it is invertible. Hence in the SU(2)-case we can calculate the inverse matrix $\omega_{\mu\nu}$ and hence the corresponding symplectic form ω. By showing that $\omega_{\mu\nu}$ is closed and SU(2)-invariant we thus find the Kirillov form on the two-dimensional spheres, the physically relevant coadjoint orbits of SU(2). We will do the calculation in two often used coordinate systems: the spherical coordinate system and the stereographic coordinate system.

5.2.1 The Kirillov form on the coadjoint orbits of SU(2) in the spherical coordinate system

Let $\{x_1, x_2, x_3\}$ be the standard coordinates in $\mathbb{R}^3$. The spheres of radius R in $\mathbb{R}^3$ are given by the equation $x_1^2 + x_2^2 + x_3^2 = R^2$.

We can cover the sphere by two charts. One chart being the whole sphere minus the northpole ($\theta = 0$), which we denote by U_+, the other chart being the whole sphere minus the southpole ($\theta = \pi$), which we denote by U_-. Because we calculate the Kirillov form on the two-dimensional sphere in local coordinates, we have to perform the calculation on both charts U_+ and U_-. Let us first consider U_-.

The well known parametrization of U_- is in terms of the parameters ϕ, θ and is given by

$$\begin{cases} x_1 = R\sin(\theta)\cos(\phi) \\ x_2 = R\sin(\theta)\sin(\phi) \\ x_3 = R\cos(\theta) \\ 0 \leq \phi < 2\pi, \quad 0 \leq \theta < \pi. \end{cases}$$

Indeed,

$$\begin{aligned} &x_1^2 + x_2^2 + x_3^2 = \\ &(R\sin(\theta)\cos(\phi))^2 + (R\sin(\theta)\sin(\phi))^2 + (R\cos(\theta))^2 = \\ &R^2(\sin^2(\theta)\cos^2(\phi) + \sin^2(\theta)\sin^2(\phi) + \cos^2(\theta)) = \\ &R^2(\sin^2(\theta) + \cos^2(\theta)) = R^2. \end{aligned}$$

As a first step towards finding the Kirillov form on U_- in the spherical coordinate system the Poisson brackets $\{\theta, \phi\}$, $\{\phi, \theta\}$, $\{\theta, \theta\}$ and $\{\phi, \phi\}$ need to be calculated. By formula (26) and in particular the antisymmetry of $\omega^{\theta\theta}$, $\omega^{\theta\phi}$ and $\omega^{\phi\phi}$ it immediately follows that

$$\begin{aligned} \{\theta, \theta\} &= 0 \\ \{\phi, \phi\} &= 0 \\ \{\theta, \phi\} &= -\{\phi, \theta\}. \end{aligned}$$

So it is left to calculate $\{\phi, \theta\}$. The most convenient way to do this is to first calculate the bracket $\left\{\frac{x_1}{x_2}, x_3\right\}$ in terms of the parameters θ, ϕ and subsequently, to express this bracket in terms of $\{\phi, \theta\}$.

In subsubsection 4.2.2 we already proved the commutation relations for the generators of $\mathfrak{su}(2) \cong \mathbb{R}^3$. For the choice of generators $\{x_1, x_2, x_3\}$ of $\mathbb{R}^3 \cong \mathfrak{su}(2)$ here we thus have:

$$[x_1, x_2] = 2x_3, \quad [x_1, x_3] = -2x_2, \quad [x_2, x_3] = 2x_1.$$

We claim that a bilinear map $\{,\} : C^\infty(\mathbb{R}^3) \times C^\infty(\mathbb{R}^3) \to C^\infty(\mathbb{R}^3)$ which is a derivation in its first argument and which satisfies the bracket relations

$$\{x_i, x_j\} = c\epsilon_{ijk} x_k,$$

for $c \in \mathbb{R}$ and ϵ_{ijk} the Levi-Civita tensor, defines a Poisson bracket:

- Jacobi-identity: For $f, g \in C^\infty(\mathbb{R}^3)$ we have

$$\{f, g\} = \frac{\partial f}{\partial x_i}\frac{\partial g}{\partial x_j}\{x_i, x_j\} \text{ with } i, j \in \{1, 2, 3\}, \tag{28}$$

  which easily follows from equation (26) and the chain rule. To check the Jacobi identity for general functions in $C^\infty(\mathbb{R}^3)$ it is by (28) sufficient to check it for the coordinates $\{x_1, x_2, x_3\}$ on $\mathbb{R}^3$. Indeed,

  $\{x_1, \{x_2, x_3\}\} + \{x_2, \{x_3, x_1\}\} + \{x_3, \{x_1, x_2\}\} =$
  $\{x_1, cx_1\} + \{x_2, cx_2\} + \{x_3, cx_3\} = 0.$

- Anti-symmetry: follows directly from the definition of the Levi-Civita tensor and (28).

This turns $\mathbb{R}^3$ into a Poisson manifold. Moreover, for the Poisson bracket induced by the Kirillov form the constant c in the bracket relations above equals $c = 2$, since in that case the structure constants of the Poisson bracket relations and the Lie bracket relations between the generators of $\mathfrak{su}(2)$ are the same (without proof). Consequently, for these induced Poisson brackets it holds that,

$$\{x_1, x_2\} = 2x_3, \quad \{x_1, x_3\} = -2x_2, \quad \{x_2, x_3\} = 2x_1.$$

Furthermore, in order to calculate $\left\{\frac{x_1}{x_2}, x_3\right\}$ we use that
$\{f(x_1, x_2), x_3\} = \frac{\partial f}{\partial x_1}\{x_1, x_3\} + \frac{\partial f}{\partial x_2}\{x_2, x_3\}$.

Consequently,

$\left\{\frac{x_1}{x_2}, x_3\right\} = \frac{1}{x_2}\{x_1, x_3\} - \frac{x_1}{x_2^2}\{x_2, x_3\} = \frac{1}{x_2}(-2x_2) - \frac{x_1}{x_2^2}(2x^1) =$
$2\left(-1 - \frac{x_1^2}{x_2^2}\right) = -2(1 + \tan^2(\phi)) = \frac{-2}{\cos^2(\phi)}.$

Again, from (26) and the chain rule follows $\{f(\phi), g(\theta)\} = \frac{\partial f}{\partial \phi}\frac{\partial g}{\partial \theta}\{\phi, \theta\}$. So, on the other hand,

$\left\{\frac{x_1}{x_2}, x_3\right\} = \{\tan(\phi), R\cos(\theta)\} = \left(\frac{-1}{\cos^2(\phi)}\right) R\sin(\theta)\{\phi, \theta\}.$

It thus follows that

$$\{\phi, \theta\} = \frac{2}{R\sin(\theta)}, \tag{29}$$

which has a singularity at the point $\theta = 0$.

Hence, the matrix $\omega^{\mu\nu}$ in (26) is given by

$$\omega^{\mu\nu} = \begin{pmatrix} 0 & \frac{2}{R\sin(\theta)} \\ -\frac{2}{R\sin(\theta)} & 0 \end{pmatrix}.$$

It clearly has a nonvanishing determinant, implying that it is invertible. The inverse is given by

$$\omega_{\mu\nu} = \begin{pmatrix} 0 & -\frac{1}{2}R\sin(\theta) \\ \frac{1}{2}R\sin(\theta) & 0 \end{pmatrix}.$$

Consequently, the symplectic form on U_- equals

$$\begin{aligned} \omega &= 0\, d\phi \wedge d\phi + 0\, d\theta \wedge d\theta - \frac{1}{2}R\sin(\theta)\, d\phi \wedge d\theta + \frac{1}{2}R\sin(\theta)\, d\theta \wedge d\phi \\ &= R\sin(\theta)\, d\theta \wedge d\phi. \end{aligned} \tag{30}$$

Indeed, ω is closed: $d\omega = R\cos(\theta)\, d\theta \wedge d\theta \wedge d\phi = 0$. Hence, in order to show that ω defines the Kirillov form on U_+ in the spherical coordinate system we have to prove it is SU(2)-invariant.

Since we have characterized our points on the sphere by the coordinates x_1, x_2, x_3 satisfying $x_1^2 + x_2^2 + x_3^2 = R^2$ we cannot act on these vectors by elements of SU(2), being 2×2 matrices. However, since we proved earlier that $SU(2)/\{\pm I\} \cong SO(3)$, proving S0(3)-invariance is equivalent to proving SU(2)-invariance of ω. We can of course let SO(3), the rotation group in 3 dimensions, act on the vectors corresponding to points on the sphere. Since our choice of axes is arbitrary it suffices to prove rotation invariance by choosing a specific rotation, being the rotation around the x_3-axis by a constant angle $\tilde{\phi}$. The matrix representing this rotation is given by:

$$\begin{pmatrix} \cos(\tilde{\phi}) & \sin(\tilde{\phi}) & 0 \\ -\sin(\tilde{\phi}) & \cos(\tilde{\phi}) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Letting this matrix act on the vector $(x_1, x_2, x_3)^T$ gives

$$\begin{pmatrix} \cos(\tilde{\phi}) & \sin(\tilde{\phi}) & 0 \\ -\sin(\tilde{\phi}) & \cos(\tilde{\phi}) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} R\sin(\theta)\sin(\phi) \\ R\sin(\theta)\cos(\phi) \\ R\cos(\theta) \end{pmatrix} =$$

$$\begin{pmatrix} R\sin(\theta)\sin(\phi)\cos(\tilde{\phi}) + R\sin(\theta)\cos(\phi)\sin(\tilde{\phi}) \\ -R\sin(\theta)\sin(\phi)\sin(\tilde{\phi}) + R\sin(\theta)\cos(\phi)\cos(\tilde{\phi}) \\ R\cos(\theta) \end{pmatrix} =$$

$$\begin{pmatrix} R\sin(\theta)\sin(\phi + \tilde{\phi}) \\ R\sin(\theta)\cos(\phi + \tilde{\phi}) \\ R\cos(\theta) \end{pmatrix}.$$

So the transformed coordinates are $\phi' = \phi + \tilde{\phi}$ and $\theta' = \theta$. We conclude

$$\omega' = R\sin(\theta')\, d\theta' \wedge d(\phi + \tilde{\phi}) = R\sin(\theta)\, d\theta \wedge d\phi = \omega,$$

showing SU(2)-invariance of ω.

Finally, notice that the calculation on U_+ is fully analogous, except that in this case $0 < \theta \leq \pi$ and $\{\phi, \theta\}$ has a singularity at $\theta = \pi$.

5.2.2 The Kirillov form on the coadjoint orbits of SU(2) in the stereographic coordinate system

Another often used coordinate system to calculate the 'local coordinate' shape of the Kirillov form on the the two-dimensional spheres is the stereographic coordinate system. To this end, consider the stereographic projection that projects the two-dimensional sphere to the complex plane. This projection is depicted in figure 2.

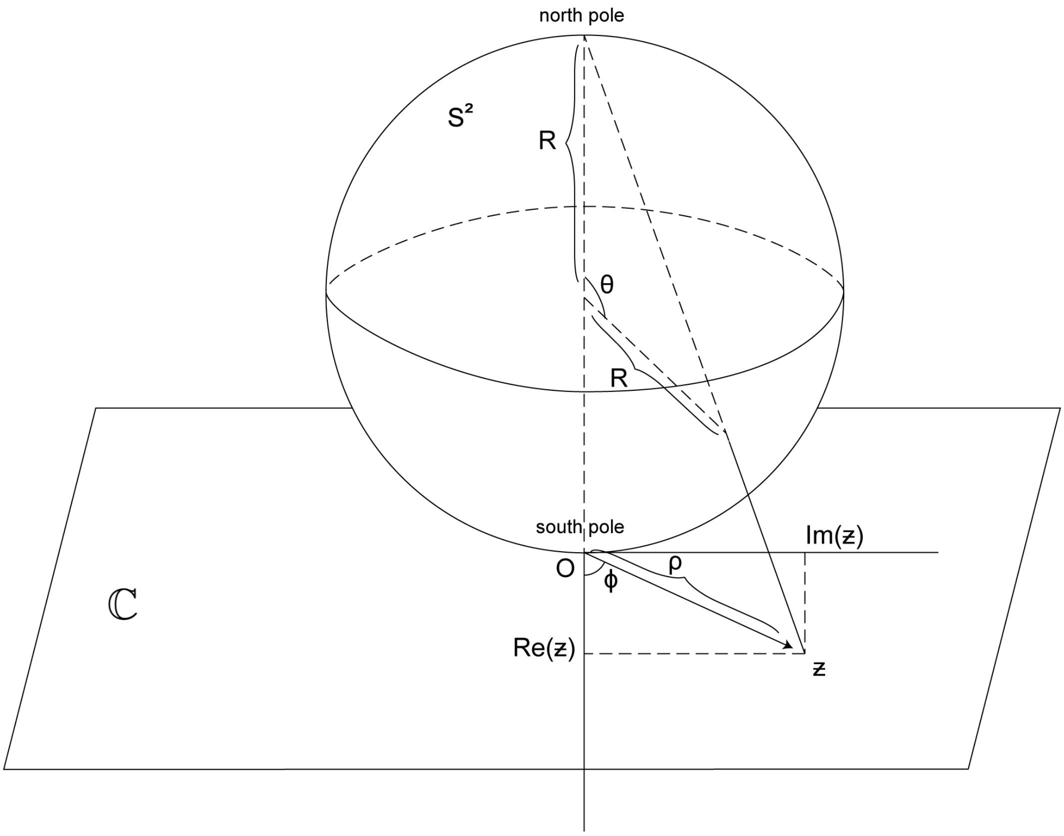

Figure 2: The stereographic projection. For any point P on the sphere (minus the southpole), which we denoted by U_-, there is a unique line trough the northpole and P, and this line intersects the complex plane (which runs trough the southpole of the sphere) in exactly one point z.

The plane $x_3 = 0$ runs trough the southpole of the sphere, which coincides with the origin of $\mathbb{R}^3$. Let M be the sphere minus the southpole (U_-). For any point P on M, there is a unique line trough the northpole and P, and this line intersects the complex plane $x_3 = 0$ in exactly one point z. The stereographic projection of P is defined to be this point z. We denote the distance from the southpole of the sphere to the point z by ρ. The stereographic coordinates z can thus be expressed in terms of the variables ϕ, ρ. We already know that the points on the sphere can be expressed in terms of the variables θ,ϕ. So if we find an expression for ρ in terms of θ (and the constant radius of the sphere R), this enables us to write z as a function of θ and ϕ (and R). As we shall see, this will be helpful in calculating the Poisson brackets between the independent coordinates, $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$, that span the complex plane.

To express ρ in terms of θ, consider the triangle in figure 3. We denote

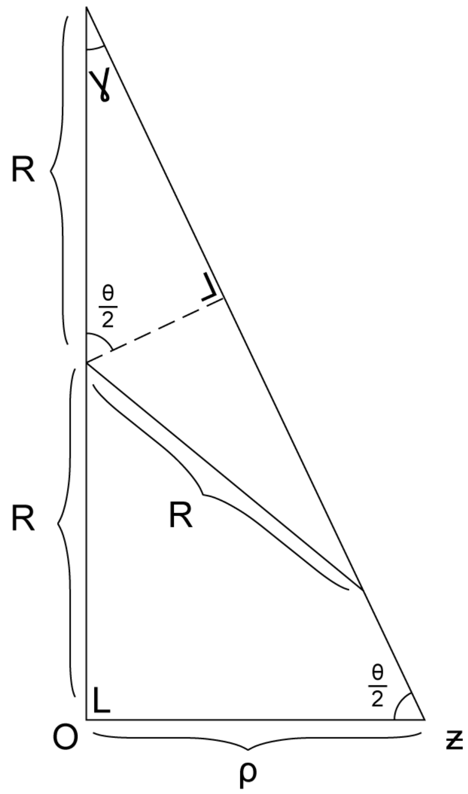

Figure 3: The triangle A lifted from figure 1 and the small rectangular triangle B at the 'top' of A.

the triangle as a whole by A and the rectangular triangle at the 'top' of A by B. Because both triangles A and B have angle γ and a right angle in common, their third angle should be equal too (the sum of the angles being 180° for every triangle). For triangle B we already know that this third angle equals $\theta/2$. Hence, the third angle of triangle A equals $\theta/2$ too (see figure). It follows that

$$\tan(\theta/2) = \frac{2R}{\rho} \Leftrightarrow \rho = 2R\cot(\theta/2). \tag{31}$$

Consequently, $\mathrm{Re}(z) = 2R\cot(\theta/2)\cos(\phi)$ and $\mathrm{Im}(z) = 2R\cot(\theta/2)\sin(\phi)$. Using Euler's formula, $\cos(\phi) + i\sin(\phi) = e^{i\phi}$, we conclude that

$$z = 2R\cot(\theta/2)e^{i\phi}, \tag{32}$$

being the superposition of $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$.

The next step is to calculate the Poisson brackets $\{\mathrm{Re}(z), \mathrm{Re}(z)\}$, $\{\mathrm{Re}(z), \mathrm{Im}(z)\}$, $\{\mathrm{Im}(z), \mathrm{Re}(z)\}$ and $\{\mathrm{Im}(z), \mathrm{Im}(z)\}$. By formula (26) and in particular the antisymmetry of $\omega^{\{\mathrm{Re}(z),\mathrm{Re}(z)\}}$, $\omega^{\{\mathrm{Im}(z),\mathrm{Im}(z)\}}$ and $\omega^{\{\mathrm{Re}(z),\mathrm{Im}(z)\}}$ it immediately follows that

$$\begin{aligned}
&\{\mathrm{Re}(z), \mathrm{Re}(z)\} = 0\\
&\{\mathrm{Im}(z), \mathrm{Im}(z)\} = 0\\
&\{\mathrm{Re}(z), \mathrm{Im}(z)\} = -\{\mathrm{Im}(z), \mathrm{Re}(z)\}.
\end{aligned}$$

So it is left to calculate $\{\mathrm{Re}(z), \mathrm{Im}(z)\}$. The most convenient way to do this is expressing the bracket firstly in terms of $\{\bar{z}, z\}$ and secondly to express $\{\bar{z}, z\}$ in terms of the brackets $\{\theta, \phi\}$, $\{\phi, \theta\}$, $\{\theta, \theta\}$ and $\{\phi, \phi\}$. The last four brackets we already calculated and hence $\{\mathrm{Re}(z), \mathrm{Im}(z)\}$ can be written as a function of the variables θ, ϕ. By using the expression of z in terms of θ, ϕ (and R) we can then write $\{\mathrm{Re}(z), \mathrm{Im}(z)\}$ as a function of z, $\bar{z}$ (and R).

From the identities $\mathrm{Re}(z) = \frac{1}{2}(z + \bar{z})$ and $\mathrm{Im}(z) = \frac{1}{2i}(z - \bar{z})$ it follows

$$\begin{aligned}
&\{\mathrm{Re}(z), \mathrm{Im}(z)\} = \tfrac{1}{4i}\{z + \bar{z}, z - \bar{z}\} =\\
&\tfrac{1}{4i}\{z, z\} - \tfrac{1}{4i}\{z, \bar{z}\} + \tfrac{1}{4i}\{\bar{z}, z\} - \tfrac{1}{4i}\{\bar{z}, \bar{z}\} = \tfrac{1}{2i}\{\bar{z}, z\}.
\end{aligned}$$

To find the expression for $\{\bar{z}, z\}$, we in particular use the fact that the Poisson bracket is a derivation in its first argument, giving,

$$\begin{aligned}
&\{\bar{z}, z\} = -\{z, \bar{z}\} =\\
&-4R^2\{\cot(\theta/2)e^{i\phi}, \cot(\theta/2)e^{-i\phi}\} =\\
&-4R^2\cot(\theta/2)\{e^{i\phi}, \cot(\theta/2)e^{-i\phi}\} - 4R^2e^{i\phi}\{\cot(\theta/2), \cot(\theta/2)e^{-i\phi}\} =\\
&4R^2\cot(\theta/2)\{\cot(\theta/2)e^{-i\phi}, e^{i\phi}\} + 4R^2e^{i\phi}\{\cot(\theta/2)e^{-i\phi}, \cot(\theta/2)\} =\\
&4R^2\cot^2(\theta/2)\{e^{-i\phi}, e^{i\phi}\} + 4R^2\cot(\theta/2)e^{-i\phi}\{\cot(\theta/2), e^{i\phi}\}\\
&+ 4R^2e^{i\phi}\cot(\theta/2)\{e^{-i\phi}, \cot(\theta/2)\} + 4R^2\{\cot(\theta/2), \cot(\theta/2)\}.
\end{aligned}$$

We can now work out the four remaining brackets, using $\{f(\theta, \phi), g\} = \frac{\partial f}{\partial \theta}\{\theta, g\} + \frac{\partial f}{\partial \phi}\{\phi, g\}$ and equation (28):

- $\{\cot(\theta/2), \cot(\theta/2)\} = 0$

- $\{e^{-i\phi}, e^{i\phi}\} = -ie^{-i\phi}\{\phi, e^{i\phi}\} = ie^{-i\phi}\{e^{i\phi}, \phi\} = -e^{-i\phi}e^{i\phi}\{\phi, \phi\} = 0$
- $\{\cot(\theta/2), e^{i\phi}\} = -\frac{1}{2}(1+\cot^2(\theta/2))\{\theta, e^{i\phi}\} =$ $\frac{1}{2}(1+\cot^2(\theta/2))ie^{i\phi}\{\phi, \theta\} = \frac{(1+\cot^2(\theta/2))ie^{i\phi}}{R\sin(\theta)}$
- $\{e^{-i\phi}, \cot(\theta/2)\} = \frac{1}{2}(1+\cot^2(\theta/2))ie^{-i\phi}\{\phi, \theta\} = \frac{(1+\cot^2(\theta/2))ie^{-i\phi}}{R\sin(\theta)}$.

As a consequence, the equation for $\{\bar{z}, z\}$ becomes

$\{\bar{z}, z\} =$
$0 + 4R^2\cot(\theta/2)e^{-i\phi}(\frac{(1+\cot^2(\theta/2))ie^{i\phi}}{R\sin(\theta)}) + 4R^2\cot(\theta/2)e^{i\phi}(\frac{(1+\cot^2(\theta/2))ie^{-i\phi}}{R\sin(\theta)}) + 0 =$
$8iR\frac{\cot(\theta/2)}{\sin(\theta)}(1+\cot^2(\theta/2)) = 4iR\left(1+\cot^2(\theta/2)\right)^2$.

Using $\cot^2(\theta/2) = \frac{z\bar{z}}{4R^2}$ it follows that $\{\bar{z}, z\} = 4iR\left(1+\frac{z\bar{z}}{4R^2}\right)^2$. Hence,

$$\{\mathrm{Re}(z), \mathrm{Im}(z)\} = 2R\left(1+\frac{z\bar{z}}{4R^2}\right)^2. \tag{33}$$

Hence, the matrix $\omega^{\mu\nu}$ in (26) is given by

$$\omega^{\mu\nu} = \begin{pmatrix} 0 & 2R\left(1+\frac{z\bar{z}}{4R^2}\right)^2 \\ -2R\left(1+\frac{z\bar{z}}{4R^2}\right)^2 & 0 \end{pmatrix}.$$

It clearly has a nonvanishing determinant, implying that it is invertible. The inverse is given by

$$\omega_{\mu\nu} = \begin{pmatrix} 0 & -\frac{1}{2R}\left(1+\frac{z\bar{z}}{4R^2}\right)^{-2} \\ \frac{1}{2R}\left(1+\frac{z\bar{z}}{4R^2}\right)^{-2} & 0 \end{pmatrix}.$$

Consequently, the symplectic form on U_- equals

$$\begin{aligned} \omega &= -\frac{1}{2R}\left(1+\frac{z\bar{z}}{4R^2}\right)^{-2} d\mathrm{Re}(z)\wedge d\mathrm{Im}(z) + \frac{1}{2R}\left(1+\frac{z\bar{z}}{4R^2}\right)^{-2} d\mathrm{Im}(z)\wedge d\mathrm{Re}(z) \\ &= \frac{1}{R}\left(1+\frac{z\bar{z}}{4R^2}\right)^{-2} d\mathrm{Im}(z)\wedge d\mathrm{Re}(z) \\ &= i\frac{1}{2R}\left(1+\frac{z\bar{z}}{4R^2}\right)^{-2} dz\wedge d\bar{z}. \end{aligned} \tag{34}$$

In subsubsection 7.3.2 will be discussed what the shape of the symplectic form on U_+ is in terms of the coordinates z and $\bar{z}$ and what this implies for the singularity structure of the symplectic form.

Remark 5.2.2.1. Although sometimes z and $\bar{z}$ are treated as independent, they are in fact not. The variables $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$ are independent. This is why we calculated the matrix $\omega^{\mu\nu}$ with respect to the basis $\{\mathrm{Re}(z), \mathrm{Im}(z)\}$ of the complex plane and only at the very end substituted z and $\bar{z}$ in the expression for ω. Notice, however, that the complexifications of z and $\bar{z}$ are independent.

Indeed, ω is closed:

$$d\omega = i\frac{1}{2R}\frac{\partial}{\partial z}\left(1+\frac{z\bar{z}}{4R^2}\right)^{-2}\, dz \wedge dz \wedge d\bar{z} + i\frac{1}{2R}\frac{\partial}{\partial \bar{z}}\left(1+\frac{z\bar{z}}{4R^2}\right)^{-2}\, d\bar{z} \wedge dz \wedge d\bar{z} = 0.$$

Of course, because the form of ω is independent of the choice of coordinates, if ω is closed in the spherical coordinate system, it is closed in any coordinate system. Similarly, when ω is SU(2)-invariant in the spherical coordinate system, it is SU(2)-invariant in any coordinate system. We conclude that equation (33) defines the Kirillov form on the two-dimensional spheres in the stereographic coordinate system.

6 Geometric quantization

-The correspondence between a classical theory and a quantum theory should be based not so much on a coincidence between their predictions in the limit $\hbar \to 0$, as on an analogy between their mathematical structures: the primary role of the classical theory is not in approximating the quantum theory, but in providing a framework for its interpretation.-
P.A.M. Dirac ([14])

Geometric quantization (abbreviated to GQ) refers to a body of ideas pioneered independently by Souriau ([18]) and Kostant ([19]) in the late 60's and early 70's. The aim of the GQ programme is to find a way of formulating the relationship between classical and quantum mechanics in a geometric language, as a relationship between symplectic manifolds (classical phase spaces) and Hilbert spaces (quantum phase spaces). The passage from classical to quantum mechanics depends on the introduction into the classical phase space of an additional geometric structure called a **polarization**.

I will start this section with a formulation of classical mechanics in terms of symplectic geometry, then discuss some criteria (axioms) which a quantization has to satisfy, after which I give a heuristic discussion how to construct a prequantum Hilbert space. This is the space corresponding to the prequantization construction. Prequantization is a 'weak' version of quantization, satisfying only a part of the axioms mentioned above. After this, prequantization is formalized by using the 'natural' mathematical objects on which one has to apply prequantization. After applying this prequantization procedure the Hilbert space is still 'too big' and we need to reduce it. We do this by means of polarizations, which are explained in the subsequent subsection. Finally, equipped with all the necessary mathematical tools, we construct the Hilbert space.

6.1 The mathematical framework of classical mechanics

Symplectic geometry is the adequate mathematical framework for describing the Hamiltonian version of classical mechanics. As such it is also the most suitable starting point for a geometrization of the canonical (Dirac) quantization procedure.

The purpose of subsubsection 6.1.1 is to introduce the formalism of symplectic geometry. Subsubsection 6.1.2 serves to establish the relation of this formalism with that of classical Hamiltonian mechanics.

6.1.1 A crash course in symplectic geometry

To start with, let us repeat definition 5.1,

Definition 5.1. A **symplectic manifold** (M, ω) is a smooth manifold M on which there exists a closed nondegenerate differential 2-form ω (called the symplectic form). In local coordinates:

$$\omega = \omega_{\mu\nu}(x) dx^\mu \wedge dx^\nu, \quad d\omega = 0,$$

where $\omega_{\mu\nu}$ is an anti-symmetric, invertible matrix.

Furthermore, this implies that symplectic manifolds always have even dimension (see section 5.1).

The most important example of a symplectic manifold is the cotangent bundle $M = T^*Q$ of an n-dimensional manifold Q. This is the set of pairs (p, q) where $q \in Q$ and p is a differential 1-form at q. It is made into a vector bundle (and hence into a manifold) by using as coordinates the set of $2n$ funtions $\{p_a, q^b\}$ where the q^bs are coordinates on Q and the p_as are the corresponding components of the differential 1-forms p. $\{p_a, q^b\}$ is called the **extension** to T^*Q of the coordinate system $\{q^b\}$.

The symplectic structure on M is the differential 2-form ω defined by

$$\omega = dp_a \wedge dq^a. \tag{35}$$

Indeed ω defines a symplectic form:

(1) ω is a differential 2-form, because it is of the form (1) for $k = 2$.

(2) ω is closed. Indeed, define the 1-form θ by

$$\theta = p_a dq^a. \tag{36}$$

From this, one sees that $\omega = d\theta$, so it is globally exact and hence closed since $d \circ d = 0$ by definition of the exterior derivative.

From this we can now show that expression (35) is invariant under change of coordinates. Denote the transformed coordinates by $\{\tilde{p}_a, \tilde{q}^b\}$.

Then the transformed and untransformed coordinates are related, using the Einstein summation convention, via, ([20, p.181])

$$p_a = \frac{\partial \tilde{q}^b}{\partial q^a} \tilde{p}_b. \tag{37}$$

It follows that

$\tilde{\omega} = d\tilde{\theta} = d(\tilde{p}_a d\tilde{q}^a) = d(\tilde{p}_a \frac{\partial \tilde{q}^a}{\partial q^b} dq^b) =$
$d(p_b dq^b) = d\theta = \omega.$

Hence (35) is invariant under change of coordinates.

(3) ω is nondegenerate. We write $\omega = \omega_{\mu\nu}(x) dx^\mu \wedge dx^\nu = a_{ij} dp^i \wedge dp^j + b_{ij} dq^i \wedge dq^j + c_{ij} dp^i \wedge dq^j + d_{ij} dq^i \wedge dp^j$. Since $a_{ij} = b_{ij} = 0 \,\forall\, i, j$ and $c_{ij} = \frac{1}{2}\ \forall i, j$, $d_{ij} = -\frac{1}{2}\ \forall i, j$ we get

$$\omega_{\mu\nu} = \begin{pmatrix} 0 & \frac{1}{2} I \\ -\frac{1}{2} I & 0 \end{pmatrix},$$

where I is the $n \times n$ identity matrix. Hence the determinant of $\omega_{\mu\nu}$ is nonzero. We conclude that $\omega_{\mu\nu}$ is an invertible and anti-symmetric matrix.

In typical physical applications, Q is the configuration space of some mechanical system and T^*Q is the phase space of the system. The cotangent bundle is a fundamental example since all symplectic manifolds have this form locally. This follows from Darboux's theorem, which we state without proof (for the proof, see [7, p.7-9]).

Theorem 6.1.1.1. Let (M, ω) be a $2n$-dimensional symplectic manifold and let $m \in M$. Then there is a neighbourhood U of m and a coordinate system $\{p_a, q^b\} (a, b = 1, 2, ..., n)$ on U such that $\omega|_U = dp_a \wedge dq^a$.

Consequently, in any symplectic manifold (M, ω), it is possible to find a 1-form θ in some neighbourhood of each point such that $\omega = d\theta$; such a 1-form is called a **symplectic potential** .

The symplectic form ω gives, at each $m \in M$, an isomorphism between the tangent and cotangent spaces of M at m, $\omega_m : T_m M \to T^*_m M$, expressed in local coordinates as

$$X^i \mapsto X^i \omega_{ik}. \tag{38}$$

Indeed,

- ω_m is a homomorphism. Indeed,
  $\omega_m(X^i + Y^i) =$
  $(X^i + Y^i)\omega_{ik} = X^i\omega_{ik} + Y^i\omega_{ik} = \omega_m(X^i) + \omega_m(Y^i)$.

- ω_m is surjective. Take $Y_k \in T^*_m M$. Then $Y_k = X^i\omega_{ik} = \omega_m(X^i)$ for a certain X^i.

- ω_m is injective. Suppose $X^i\omega_{ik} = Y^j\omega_{jk}$ for certain X^i, Y^j. Then $-\omega_{ki}X^i = -\omega_{kj}Y^j \Leftrightarrow \omega^{lk}\omega_{ki}X^i = \omega^{lk}\omega_{kj}Y^j \Leftrightarrow \delta^l_i X^i = \delta^l_j X^j$
  $\Leftrightarrow X^l = Y^l$.

This extends to an isomorphism between TM and T^*M and between vector fields and one-forms on M,

$$X \mapsto i(X)\omega = 2\omega(X, .),$$

where $i(X)$ denotes the contraction of a differential form with the vector field X, as in (38), i.e. the insertion of X into the first 'slot' of a differential form. The conventions of section 2 are used in applying this contraction. It is sometimes helpful to think of ω as an antisymmetric metric on M; then (38) corresponds to 'lowering the index' on X.

Therefore, the existence of ω on a symplectic manifold allows us to associate a vector field X_f to every function $f \in C^\infty(M, \mathbb{R})$, the space of smooth real-valued functions, via

$$X_f \cong i(X_f)\omega = 2\omega(X_f, .) = -df. \tag{39}$$

In order to justify this statement we have to prove that for every function $f \in C^\infty(M, \mathbb{R})$ we can establish the identity $2\omega(X_f, .) = -df$. This means that if we write $X_f = \xi^k \frac{\partial}{\partial x^k}$ we can find an explicit expression for the coefficients ξ^k. This is the case, as follows from

$2\omega(X_f, .) = 2\omega_{ij}dx^i \wedge dx^j(\xi^k \frac{\partial}{\partial x^k}, .) = df(.) \Leftrightarrow$
$2\xi^k\omega_{kj}dx^j = \partial_j f dx^j \Leftrightarrow$
$\partial_j f = 2\xi^k\omega_{kj} \Leftrightarrow$
$\xi^k = \frac{1}{2}\partial_j f \omega^{jk}$,

where in the last identity the invertibility of ω_{kj} is used. We call a vector field X_f satisfying (39) a **Hamiltonian vector field** of f.

The Lie derivative of ω along X_f is zero:

$\mathcal{L}(X_f)\omega \equiv d(i(X_f)\omega) + i(X_f)d\omega = -ddf = 0,$

where the definition of the Hamiltonian vector field (39), the closedness of the symplectic form ω and the identity $d \circ d = 0$ are used respectively. This implies by property (3) of the Lie derivative (see section 3.1) that

$\frac{\partial}{\partial t}\Big|_{t=0} (\phi_f)^*_t \omega = \mathcal{L}(X_f)\omega = 0 \Leftrightarrow$
$(\phi_f)^*_t \omega = \text{constant},$

where $\phi_f : \text{dom}(X_f) \subset M \times \mathbb{R} \to M$ is the flow generated by X_f and $t \in \mathbb{R}$. It follows that the flow ϕ_f leaves ω invariant.

If (M, ω) is a symplectic manifold and $f, g \in C^\infty(M, \mathbb{R})$ we can globally define the Poisson bracket by

$$\{f, g\} \equiv 2\omega(X_f, X_g) \in C^\infty(M). \tag{40}$$

It describes the change of $g \in C^\infty(M, \mathbb{R})$ along X_f (or vice versa), since

$\{f, g\} = -2\omega(X_g, X_f) = -i(X_f)i(X_g)\omega = i(X_f)dg = dg(X_f) =$
$X_f(g) = L(X_f)g = -L(X_g)f.$

In local coordinates expression (40) matches with expression (27). Indeed,

$\{f, g\} \equiv 2\omega(X_f, X_g) = 2\omega_{ij} dx^i \wedge dx^j (\frac{1}{2}\omega^{kl} \frac{\partial f}{\partial x^l} \frac{\partial}{\partial x^k}, \frac{1}{2}\omega^{mn} \frac{\partial g}{\partial x^n} \frac{\partial}{\partial x^m}) =$
$2\omega_{ij} \frac{1}{2}\omega^{il} \frac{\partial f}{\partial x^l} \frac{1}{2}\omega^{jn} \frac{\partial g}{\partial x^n} = -\frac{1}{2} \frac{\partial f}{\partial x^j} \omega^{jn} \frac{\partial g}{\partial x^n}.$

Here is used the expression for the Hamiltonian vector field in local coordinates (A.1), which is derived in the appendix. As promised, we will now prove that $\{,\}$ in (27) defines a Poisson bracket:

(1) $\{,\}$ is a derivation: already proved at the beginning of subsection 5.2.

(2) $\{,\}$ is anti-symmetric: follows immediately from anti-symmetry of $\omega^{\mu\nu}$, since ω is a symplectic 2-form.

(3) $\{,\}$ satisfies the Jacobi identity: let $f, g, h \in C^\infty(M, \mathbb{R})$ and let x^μ be the local coordinates on M. Using the notation $\partial_\mu g \equiv \frac{\partial}{\partial x^\mu} g$ it follows that

$\{f, \{g, h\}\} = \{f, -\frac{1}{2}\partial_\mu g \omega^{\mu\nu} \partial_\nu h\} =$
$-\frac{1}{2}\partial_\sigma f \omega^{\sigma\pi} \partial_\pi (-\frac{1}{2}\partial_\mu g \omega^{\mu\nu} \partial_\nu h) =$

$$\tfrac{1}{4}\partial_\sigma f\omega^{\sigma\pi}\partial_\pi\partial_\mu g\omega^{\mu\nu}\partial_\nu h + \tfrac{1}{4}\partial_\sigma f\omega^{\sigma\pi}\partial_\mu g\partial_\pi\omega^{\mu\nu}\partial_\nu h + \tfrac{1}{4}\partial_\sigma f\omega^{\sigma\pi}\partial_\mu g\omega^{\mu\nu}\partial_\pi\partial_\nu h.$$

Consequently,

$$\begin{aligned}
&\{f,\{g,h\}\} + \{f,\{g,h\}\} + \{f,\{g,h\}\} = \\
&\tfrac{1}{4}\partial_\sigma f\omega^{\sigma\pi}\partial_\pi\partial_\mu g\omega^{\mu\nu}\partial_\nu h+\tfrac{1}{4}\partial_\sigma f\omega^{\sigma\pi}\partial_\mu g\partial_\pi\omega^{\mu\nu}\partial_\nu h+\tfrac{1}{4}\partial_\sigma f\omega^{\sigma\pi}\partial_\mu g\omega^{\mu\nu}\partial_\pi\partial_\nu h+ \\
&\tfrac{1}{4}\partial_\sigma g\omega^{\sigma\pi}\partial_\pi\partial_\mu h\omega^{\mu\nu}\partial_\nu f+\tfrac{1}{4}\partial_\sigma g\omega^{\sigma\pi}\partial_\mu h\partial_\pi\omega^{\mu\nu}\partial_\nu f+\tfrac{1}{4}\partial_\sigma g\omega^{\sigma\pi}\partial_\mu h\omega^{\mu\nu}\partial_\pi\partial_\nu f+ \\
&\tfrac{1}{4}\partial_\sigma h\omega^{\sigma\pi}\partial_\pi\partial_\mu f\omega^{\mu\nu}\partial_\nu g+\tfrac{1}{4}\partial_\sigma h\omega^{\sigma\pi}\partial_\mu f\partial_\pi\omega^{\mu\nu}\partial_\nu g+\tfrac{1}{4}\partial_\sigma h\omega^{\sigma\pi}\partial_\mu f\omega^{\mu\nu}\partial_\pi\partial_\nu g = \\
&\tfrac{1}{4}\partial_\sigma f\omega^{\sigma\pi}\partial_\mu g\partial_\pi\omega^{\mu\nu}\partial_\nu h+\tfrac{1}{4}\partial_\sigma g\omega^{\sigma\pi}\partial_\mu h\partial_\pi\omega^{\mu\nu}\partial_\nu f+\tfrac{1}{4}\partial_\sigma h\omega^{\sigma\pi}\partial_\mu f\partial_\pi\omega^{\mu\nu}\partial_\nu g = \\
&\tfrac{1}{4}\partial_\mu f\omega^{\mu\pi}\partial_\nu g\partial_\pi\omega^{\nu\sigma}\partial_\sigma h+\tfrac{1}{4}\partial_\nu g\omega^{\nu\pi}\partial_\sigma h\partial_\pi\omega^{\sigma\mu}\partial_\nu f+\tfrac{1}{4}\partial_\sigma h\omega^{\sigma\pi}\partial_\mu f\partial_\pi\omega^{\mu\nu}\partial_\nu g = \\
&\tfrac{1}{4}\partial_\mu f\partial_\nu g\partial_\sigma h(\omega^{\mu\pi}\partial_\pi\omega^{\nu\sigma} + \omega^{\nu\pi}\partial_\pi\omega^{\sigma\mu} + \omega^{\sigma\pi}\partial_\pi\omega^{\mu\nu}) = 0,
\end{aligned}$$

where in the second identity is used that the terms containing second order derivatives of f, g or h cancel each other out by applying suitable interchanges of indices (which is allowed since all indices are summed over). In the third identity a similar interchange of indices is applied. In the last identity the Jacobi identity for Poisson forms is used, (A.3), proved in the appendix.

We have now established that the pair $(C^\infty(M,\mathbb{R}), \{,\})$ forms a Lie algebra, since the space of smooth real-valued functions is a linear space and $\{,\}$ is a bilinear map satisfying anti-symmetry and the Jacobi identity.

A very important identity, relating the Lie algebras of Hamiltonian vector fields and smooth real-valued functions on M, is

$$[X_f, X_g] = X_{\{f,g\}}, \tag{41}$$

where $f, g \in C^\infty(M,\mathbb{R})$ and $X_f, X_g, X_{\{f,g\}}$ are Hamiltonian vector fields on M. On the left hand side, $[,]$ denotes the Lie bracket of vector fields, which is defined by,

$$[X,Y]^b = X^a\partial_a Y^b - Y^a\partial_a X^b, \tag{42}$$

for $X, Y \in \mathrm{Vect}(M)$. In the appendix identity (41) is proved and from this the fact is proved that the space of Hamiltonian vector fields together with commutator bracket forms a Lie algebra.

Finally, we consider certain submanifolds of symplectic manifolds, which will become important in the study of polarizations later in this thesis. Before introducing these submanifolds, let us begin with the following definition.

Definition 6.1.1.1. A **symplectic vector space** is a pair (V, ω) in which V is a $2n$ ($n \in \mathbb{Z}$) dimensional real vector space and $\omega : V \times V \to \mathbb{C}$ is a bilinear form on V satisfying

- $\omega(X, Y) = -\omega(Y, X) \quad X, Y \in V \quad$ (anti-symmetry),
- $\omega(X, .) = 0 \Leftrightarrow X = 0 \quad X \in V \quad$ (non-degeneracy).

With this information we can now define,

Definition 6.1.1.2. A **Lagrangian subspace** $H \subset V$ is a subspace H of a symplectic vector space V with the properties

- $\dim(H) = \frac{1}{2}\dim(V) = n,$
- $\omega(X, Y) = 0 \quad X, Y \in H.$

We are now ready to define the certain submanifolds of symplectic manifolds mentioned previously. Analogously to the linear case,

Definition 6.1.1.3. A **Lagrangian submanifold** $N \subset M$ is an n-dimensional submanifold N of a $2n$-dimensional ($n \in \mathbb{Z}$) symplectic manifold (M, ω) satisfying $\omega|_{TN} = 0$.

To give some examples of Lagrangian submanifolds, let us again consider the cotangent bundle T^*Q of an n-dimensional manifold Q. Then Q is a Lagrangian submanifold of T^*Q. Indeed,

(1) Q is a n-dimensional submanifold of T^*Q: if U are local charts on Q, then $U \times \mathbb{R}^n$ are local charts on T^*Q (the argument is analogous to the one given in section 3 for the tangent bundle). Let ϕ denote the functions $\phi : U \times \mathbb{R}^n \to \mathbb{R}^{2n}$ corresponding to the local charts on T^*Q. It follows that $Q \cap (U \times \mathbb{R}^n) = U \times \{0\} = \phi^{-1}(\mathbb{R}^n \times \{0\})$, which proves the claim.

(2) $\omega|_{TQ} = 0$: the symplectic structure on T^*Q is given by the differential 2-form $\omega = dp_a \wedge dq^a$. A vector field on the tangent space of T^*Q can be written as $v = v_{1,a}\frac{\partial}{\partial p_a} + v_{2,a}\frac{\partial}{\partial q_a}$. On TQ it holds that $v_{1,a} = 0$ (since $p_a = 0$ on Q), so a vector field on TQ takes the form $v = v_a\frac{\partial}{\partial q_a}$. Hence, for $v, w \in TQ$, $\omega(v, w) = (dp_a \wedge dq^a)(v_b\frac{\partial}{\partial q_b})(v_c\frac{\partial}{\partial q_c}) = 0 - 0 = 0$. This implies $\omega|_{TQ} = 0$.

Furthermore, for fixed $q \in Q$, the fibres T_q^*Q are Lagrangian submanifolds of T^*Q. In this case $v_{2,a} = 0$ in the above form for v and $T_q^*Q \cong \mathbb{R}^n$ gives the n-dimensionality of T_q^*Q.

Locally, any Lagrangian submanifold N is given by the vanishing of n smooth real-valued functions f_k on M which are in involution, i.e. satisfying,

$$\{f_k, f_l\} = 0 \quad \forall \ k, l. \tag{43}$$

Indeed, the vanishing of n real-valued functions on M gives n restrictions on M, reducing its dimension by n, leaving a dimension $2n - n = n$ for N. Besides, $\{f_k, f_l\} = 0 \Leftrightarrow X_{f_k}(f_l) = \omega^{qs}\frac{\partial f_k}{\partial x^q}\frac{\partial f_l}{\partial x^s} = 0$. Since the last equation holds for all k, l this implies that the Hamiltonian vector fields X_{f_k} are tangent to $\bigcap_l \{F_l = 0\}$, so that they locally span the tangent bundle TN. Then, noticing that (43) holds if and only if $\omega(X_{f_k}, X_{f_l}) = 0$, this implies $\omega|_{TN} = 0$.

This concludes our crash course in symplectic geometry. Many results of modern symplectic geometry are not mentioned here, and the adventurous reader is referred to the bible of symplectic geometry and classical mechanics, [20], for a detailed account.

6.1.2 Symplectic geometry and classical mechanics

To establish the relation between symplectic geometry and classical mechanics, let us start by recalling the formalism of classical Hamiltonian mechanics. In the Hamiltonian formalism, the arena for classical mechanics in the simplest mechanical systems is the **phase space** . This is a $2n$-dimensional ($n \in \mathbb{Z}$) real linear space, and hence isomorphic to $\mathbb{R}^{2n}$, with coordinates $\{q^1, ..., q^n, p_1, ..., p_n\}$ describing the position and the momentum (velocity) of the n particles involved. These coordinates are called **canonical coordinates** . The dynamics (or time evolution) of the system is gouverned by **Hamilton's equations** , given by

$$\begin{aligned} \frac{d}{dt}q^k &= \frac{\partial H}{\partial p_k} \\ \frac{d}{dt}p_k &= -\frac{\partial H}{\partial q^k}. \end{aligned} \tag{44}$$

Here $H \equiv H(q^k, p_k)$, the **Hamiltonian** , is a real smooth function on phase space describing the energy of the system.

Typically, H is of the form $H = T + V$ with $T \propto p^2 \equiv p^k p_k$ (Einstein summation convention used) the kinetic energy and $V = V(q^k)$ the potential energy, minus whose gradient describes the forces acting on the particles.

As an example, let us consider the harmonic oscillator in one dimension ($n = 1$). It is described by the Hamiltonian $H = \frac{1}{2}(p^2 + q^2)$. Hamilton's equations become

$$\begin{aligned}\frac{d}{dt}q &= \dot{q} = \frac{\partial H}{\partial p} = p \\ \frac{d}{dt}p &= \dot{p} = -\frac{\partial H}{\partial q} = -q.\end{aligned}$$

By substituting $\dot{q} = p$ into $\dot{p} = -q$ it follows that $\ddot{q} = -q$. The solution to this second order ordinary differential equation is a fimiliar one:

$$q(t) = A\cos(t) + B\sin(t) \quad A, B \in \mathbb{R}, \tag{45}$$

leading to the characteristic oscillating behaviour. Since $q(0) = A\cos(0) + B\sin(0) = A$ and $p(0) = \dot{q}(0) = -A\sin(0) + B\cos(0) = B$ we can write (45) as $q(t) = q(0)\cos(t) + p(0)\sin(t)$.

If H does not depend on time explicitly, the equations of motion (44) imply that H is conserved along every trajectory in phase space. Indeed,

$$\begin{aligned}\frac{d}{dt}H =& \frac{\partial H}{\partial q^k}\dot{q}^k + \frac{\partial H}{\partial p_k}\dot{p}_k \\ =& \frac{\partial H}{\partial q^k}\frac{\partial H}{\partial p_k} - \frac{\partial H}{\partial p_k}\frac{\partial H}{\partial q^k} = 0,\end{aligned} \tag{46}$$

where again the Einstein summation convention is used. The evolution of any other smooth real-valued function on phase space, called an **observable**, is given by

$$\begin{aligned}\frac{d}{dt}f =& \frac{\partial f}{\partial q^k}\dot{q}^k + \frac{\partial f}{\partial p_k}\dot{p}_k \\ =& \frac{\partial f}{\partial q^k}\frac{\partial H}{\partial p_k} - \frac{\partial f}{\partial p_k}\frac{\partial H}{\partial q^k}.\end{aligned} \tag{47}$$

In the example of the harmonic oscillator in one dimension, the conservation of the Hamiltonian along any trajectory in phase space implies that $\frac{d}{dt}H = \frac{d}{dt}\left(\frac{1}{2}(p^2 + q^2)\right) = 0 \Leftrightarrow p^2 + q^2 = \text{Constant} \in \mathbb{R}$. Since Constant $\in \mathbb{R}$, we can write Constant $= r^2$ for some constant $r \in \mathbb{R}$. Hence the phase space trajectories of the one-dimensional harmonic oscillator are circles of radius

r. This is in agreement with the explicit solution of the equations of motion (45), since

$$q(t)^2+p(t)^2=(q(0)\cos(t)+p(0)\sin(t))^2+(p(0)\cos(t)-q(0)\sin(t))^2=$$
$$q(0)^2\cos^2(t)+p(0)^2\sin^2(t)+2q(0)p(0)\cos(t)\sin(t)$$
$$-2q(0)p(0)\cos(t)\sin(t)+p(0)^2\cos^2(t)+q(0)^2\sin^2(t)=$$
$$q(0)^2(\cos^2(t)+\sin^2(t))+p(0)^2(\cos^2(t)+\sin^2(t))=$$
$$q(0)^2+p(0)^2=\text{Constant}.$$

The equations (44), (46), (47) characterize Hamiltonian mechanics. They arise naturally in the formalism of symplectic geometry if we look at $\mathbb{R}^{2n}$ as the cotangent bundle $T^*\mathbb{R}^n\cong\mathbb{R}^n\times\mathbb{R}^n$ of the configuration space $\mathbb{R}^n$, equipped with the symplectic form (35). If we write the symplectic form ω from (35) as $\omega_{\mu\nu}dx^\mu\wedge dx^\nu$, then we already derived in subsubsection 6.1.1 that

$$\omega_{\mu\nu}=\begin{pmatrix}0 & \frac{1}{2}I\\ -\frac{1}{2}I & 0\end{pmatrix}.$$

By Cramer's rule, the inverse matrix is given by

$$\omega^{\mu\nu}=\begin{pmatrix}0 & -2I\\ 2I & 0\end{pmatrix}.$$

Hence, using the form of a Hamiltonian vector field in local coordinates, (A.1), derived in the appendix, the Hamiltonian vector field X_f of a function $f(q^1,...,q^n,p_1,...,p_n)$ equals

$$\begin{aligned}X_f=&\frac{1}{2}\omega^{\mu\nu}\frac{\partial f}{\partial x^\mu}\frac{\partial}{\partial x^\nu}\\ =&0\cdot\frac{\partial f}{\partial p_k}\frac{\partial}{\partial p^k}-\frac{\partial f}{\partial q^k}\frac{\partial}{\partial p_k}+\frac{\partial f}{\partial p_k}\frac{\partial}{\partial q^k}+0\cdot\frac{\partial f}{\partial q^k}\frac{\partial}{\partial q_k}\\ =&\frac{\partial f}{\partial p_k}\frac{\partial}{\partial q^k}-\frac{\partial f}{\partial q^k}\frac{\partial}{\partial p_k}\end{aligned}\tag{48}$$

. Therefore the Poisson bracket equals

$$\{f,g\}=X_f(g)=\frac{\partial f}{\partial p_k}\frac{\partial g}{\partial q^k}-\frac{\partial f}{\partial q^k}\frac{\partial g}{\partial p_k}\tag{49}$$

. It follows, that the Poisson brackets between the coordinates and momenta (called the canonical Poisson brackets) are

$\{q^m, q^l\} = \frac{\partial q^m}{\partial p_k}\frac{\partial q^l}{\partial q^k} - \frac{\partial q^m}{\partial q^k}\frac{\partial q^l}{\partial p_k} = 0 - 0 = 0$
$\{p_m, p_l\} = \frac{\partial p_m}{\partial p_k}\frac{\partial p_l}{\partial q^k} - \frac{\partial p_m}{\partial q^k}\frac{\partial p_l}{\partial p_k} = 0 - 0 = 0$
$\{p_m, q^l\} = \frac{\partial p_m}{\partial p_k}\frac{\partial q^l}{\partial q^k} - \frac{\partial p_m}{\partial q^k}\frac{\partial q^l}{\partial p_k} = \delta^k_m \delta^l_k - 0 = \delta^l_m.$

The functions q^k and p_l form a so-called **complete set** of observables. This means that any observable which Poisson commutes (has vanishing Poisson brackets) with all of them is a constant (which means constant in all coordinates and momenta). Indeed, suppose that an observable f Poisson commutes with all coordinates and all momenta. Then

$\{f, q^l\} = \frac{\partial f}{\partial p_k}\frac{\partial q^l}{\partial q^k} - \frac{\partial f}{\partial q^k}\frac{\partial q^l}{\partial p_k} = \frac{\partial f}{\partial p_l} = 0 \ \forall l,$
$\{f, p_l\} = \frac{\partial f}{\partial p_k}\frac{\partial p_l}{\partial q^k} - \frac{\partial f}{\partial q^k}\frac{\partial p_l}{\partial p_k} = -\frac{\partial f}{\partial q^m} = 0 \ \forall l,$
which implies $f(q^1, ..., q^n, p_1, ..., p_n) = \text{Constant}$.

Since $\{H, q^l\} = \frac{\partial H}{\partial p_k}\frac{\partial q^l}{\partial q^k} - \frac{\partial H}{\partial q^k}\frac{\partial q^l}{\partial p_k} = \frac{\partial H}{\partial p_l}$ and $\{H, p_l\} = \frac{\partial H}{\partial p_k}\frac{\partial p_l}{\partial q^k} - \frac{\partial H}{\partial q^k}\frac{\partial p_l}{\partial p_k} = -\frac{\partial H}{\partial q^l}$ in the formalism of symplectic geometry (44) becomes

$$(44) \Leftrightarrow \frac{d}{dt}q^l = \{H, q^l\}, \ \ \frac{d}{dt}p_l = \{H, p_l\}. \tag{50}$$

Similarly, since $\{H, H\} = \frac{\partial H}{\partial p_k}\frac{\partial H}{\partial q^k} - \frac{\partial H}{\partial q^k}\frac{\partial H}{\partial p_k}$, (46) becomes

$$(46) \Leftrightarrow \{H, H\} = 0. \tag{51}$$

Finally, since $\{H, f\} = \frac{\partial H}{\partial p_k}\frac{\partial f}{\partial q^k} - \frac{\partial H}{\partial q^k}\frac{\partial f}{\partial p_k}$, (47) becomes

$$(47) \Leftrightarrow \frac{d}{dt}f = \{H, f\} = X_H(f), \tag{52}$$

so time evolution in Hamiltonian mechanics is determined by the Hamiltonian vector field X_H of the Hamiltonian H.

A first advantage of this new formulation, is that it makes manifest the form invariance of the equations of classical mechanics under symplectomorphisms (diffeomorphisms leaving the symplectic form invariant), since the Poisson bracket is defined in terms of the symplectic form. Because canonical transformations are a particular type of symplectomorphisms, the form-invariance of the equations of classical mechanics under canonical transformations (in particular coordinate transformations) is also manifest.

A second advantage of this formulation is that it generalizes immediately to more complicated systems where the configuration space is some curved or compact symplectic manifold. The necessity to consider such more exotic systems in physics has arisen in recent years in a number of different contexts, e.g. for the description of internal degrees of freedom and in conformal field theory.

6.2 From classical to quantum mechanics

As mentioned earlier, the aim of the geometric quantization (GQ) programme is to formulate the relationship between classical and quantum mechanics in a geometric language. At the classical level the state space is a symplectic manifold (M, ω) and the observables are smooth real-valued functions on M. At the quantum level, the states of a physical system are represented by **rays** in a **Hilbert space** H and the observables by a collection $\tilde{Q}$ of **Hermitian operators** on H. Let us explain all this terminology. To start with,

Definition 6.2.1. A **Hilbert space** H is a **complex inner product space** that is also a **complete metric space** with respect to the distance function induced by the inner product.

Here,

Definition 6.2.2. A **complex inner product space** is a vector space V over the field $\mathbb{C}$ together with an **inner product** , i.e. with a map,
$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ that, for all vectors $x, y, z \in V$ and all scalars $a \in \mathbb{C}$, satisfies

- Conjugate symmetrie: $\langle x, y \rangle = \overline{\langle y, x \rangle}$,
- Antilinearity in the first argument:
  $\langle ax, y \rangle = \bar{a}\langle x, y \rangle$,
  $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$,
- Positive-definiteness: $\langle x, x \rangle \geq 0$ with equality only for $x = 0$.

The inner product in this Hilbert space is necessary for the probabilistic interpretation of the theory and is thus of fundamental importance. The existence of this inner product allows us to define a **norm** , which is the real-valued function $\|x\| = \sqrt{\langle x, x \rangle}$. This function is real-valued since the inner product is conjugate symmetric, from which follows $\langle x, x \rangle = \overline{\langle x, x \rangle}$. The **distance function** between two points $x, y \in V$ is defined in terms of the norm by $d(x, y) = \|x - y\|$. The terminology 'distance function' is justified by the fact that $d(x, y)$ is symmetric ($d(x, y) = d(y, x)$), the distance between points x and y is non-negative ($d(x, x) \geq 0$), only zero if $x = y$, and $d(x, y)$ satisfies the triangle equality ($d(x, z) \leq d(x, y) + d(y, z)$). Now,

Definition 6.2.3. A **complete metric space** is a metric space M, i.e. a non-empty set equipped with a distance function (called a metric), that is **complete**, which means that every Cauchy sequence of points in M has a limit that is also an element of M.

Since it is now clear what a Hilbert space is, we can say what rays in a Hilbert space are, which were the quantum states. Each point (vector) in the Hilbert space (apart from the origin) corresponds by definition to a so-called **pure quantum state** . Since multiplying this vectors by non-zero complex scalars leaves the Schrödinger equation invariant, two vectors that differ only by a non-zero complex scalar are considered to correspond to the same state. In this sense, each pure state is a ray in the Hilbert space.

Finally,

Definition 6.2.4. A **Hermitian operator** on a Hilbert space H is a linear map $\tilde{Q} : H \to H$ for which $\langle \tilde{Q}x, y\rangle = \langle x, \tilde{Q}y\rangle$, for $x, y \in H$ and $\langle \cdot, \cdot \rangle$ the inner product on H.

The first and simplest problem of quantization concerns the kinematic relationship between the classical and quantum domains. The kinematic problem is: *given M and ω, is it possible to construct H and $\tilde{Q}$*? In the rest of section 6, we will shed more light onto this problem. A more subtle problem concerns the dynamical relationship between the classical and quantum domains. This problem will not be treated in this thesis.

Classically, the space $C^\infty(M, \mathbb{R})$ of observables has in addition to the Lie algebra structure provided by the Poisson bracket, the structure of commutative algebra under pointwise multiplication, since then for $f, g \in C^\infty(M, \mathbb{R})$ and $x \in M$, we have:

$$(fg)(x) = f(x)g(x) = (gf)(x). \tag{53}$$

It is this property which has to be sacrificed when moving from the classical to the quantum theory. As already discussed in the introduction, the non-commutative nature of observables is at the heart of phenomena in quantum mechnanics.

Denote by S a subalgebra of $C^\infty(M, \mathbb{R})$ (that is, a subspace of $C^\infty(M, \mathbb{R})$ which is closed under pointwise multiplication of smooth real-valued functions on M). Let us look at quantization as an assignment

$$Q : S \to \text{operators on } H \quad f \mapsto Q(f) \tag{54}$$

of operators $Q(f)$ on some Hilbert space to classical observables f. Abstracting from the analogy between classical mechanics and Schrödinger and Heisenberg quantum mechanics (for example, by using the observations from the experiments described in the introduction) the assignment Q has to satisfy some requirements. Firstly,

(Q1) $\mathbb{R}$-linearity:

$$Q(rf + g) = rQ(f) + Q(g) \quad \forall r \in \mathbb{R}, \;\; f, g \in S. \tag{55}$$

Secondly,

(Q2)

$$[Q(f), Q(g)] = -i\hbar Q(\{f, g\}). \tag{56}$$

Here $[\cdot, \cdot]$ denotes the Lie bracket of vector fields and $\hbar$ is Planck's constant, a constant of nature characteristic of quantum effects. It is a very small number ($\hbar \approx 1.05457162853 \times 10^{-34}$Joule $\cdot$ second), and for most macroscopic considerations the fact that it is not zero can be neglected. Treating $\hbar$ as a parameter, and taking the limit $\hbar \to 0$ (which is considered to be the limit corresponding with classical mechanics), one recovers from (Q2) the commutative structure of classical mechanics. At the microscopic level, however, effects of order $\hbar$ can no longer be neglected and here classical mechanics needs to be replaced by quantum mechanics.

In quantum mechanics, the eigenvalues of operators on the Hilbert space H correspond to the possible results of measurements, and are hence real. This justifies that the operators $\tilde{Q}$ on Hilbert space are taken to be Hermitian, since Hermitian operators have real eigenvalues. Indeed, suppose $\tilde{Q}$ has eigenvalue $q \in \mathbb{C}$, that is $\tilde{Q}x = qx \;\forall x \in H$ and let again $\langle \cdot, \cdot \rangle$ denote the inner product on H. Then for $x, y \in H$,

$$\langle \tilde{Q}x, y \rangle = \langle x, \tilde{Q}y \rangle \Leftrightarrow \langle qx, y \rangle = \langle x, qy \rangle \Leftrightarrow \bar{q}\langle x, y \rangle = \overline{\langle qy, x \rangle} = q\overline{\langle y, x \rangle} = q\langle x, y \rangle$$
$$\Leftrightarrow \bar{q} = q \Leftrightarrow q \in \mathbb{R}.$$

So thirdly,

(Q3)

$$\langle Q(f)x, y \rangle = \langle x, Q(f)y \rangle \quad f \in S, \;\; x, y \in H. \tag{57}$$

The three requirements above can be summarized in one single statement, using definition 3.4.6, with H the complex vector space and $(S, \{\cdot,\cdot\})$ the Lie algebra, where $\{\cdot,\cdot\}$ denotes the Poisson bracket. That is, the operators in Q form a representation of a subalgebra S of the classical observables. This statement is Dirac's **quantum condition** (see [14]) and Dirac formulated it by saying that the Poisson bracket is the classical analogue of the quantum commutator.

However, the quantum condition does not, in itself, uniquely determine the underlying quantum system. Furthermore, not every Hilbert space and operators that satisfy it represent a 'physically reasonable' quantization. There are some guiding principles about when a quantization is 'physically reasonable'. First of all, S must contain the constant functions in $C^\infty(M, \mathbb{R})$ and these must be represented in Q by the corresponding multiples of the identity operator. This is equivalent to:

(Q4)

$$Q(1) = I, \tag{58}$$

with 1 the constant function ($1 : M \to \mathbb{R}\quad m \mapsto 1$) and I the identity operator ($I : H \to H\quad x \mapsto x$). This ensures by (Q2) that $\{p_a, q^a\} = 1$ at the classical level implies $[Q(q), Q(p)] = i\hbar I$ at the quantum level, required by the uncertainty principle (see, for example, [29]).

Secondly, requiring only Dirac's quantum condition and (Q4), the Hilbert space is still 'too big' and we need some irreducibility condition to make it smaller, which appears to be a natural requirement. In general, we cannot have a one-to-one correspondence between $\tilde{Q}$ and the whole of $C^\infty(M)$ without making H 'too large' and the selection of a subalgebra S for which (Q1), (Q2), (Q3) and (Q4) should hold involves picking out some additional structure in M.

One way to phrase this irreducibility condition makes use of the concept of a complete set of observables introduced in subsubsection 6.1.2. In complete analogy, a complete set of operators is defined such that the only operators which commute with all the operators from that set are multiples of the identity operator. The condition is then formulated as:

(Q5) If $\{f_1, ..., f_k\}$ is a complete set of observables, $\{Q(f_1), ..., Q(f_k)\}$ is a complete set of operators.

However, GQ uses another approach to make the Hilbert space irreducible, using so-called polarizations. Let us for a moment put aside the problem of the 'size' of H and look at a construction for H and $\tilde{Q}$ where only the symplectic structure on M is used and where there is an operator in $\tilde{Q}$ for every classical observable in $C^\infty(M,\mathbb{R})$ (so not only for observables in S). So Q works on the whole of $C^\infty(M,\mathbb{R})$ and satisfies an extension of (Q1)-(Q4) to all $f \in C^\infty(M,\mathbb{R})$. Such an assignment is called a **prequantization** and shall be treated next.

6.3 A heuristic discussion of prequantization

The **prequantization** construction is already sophisticated, since the prequantization Hilbert space consisting of 'ordinary' (square-integrable) scalar functions associated with the classical symplectic manifold M, appears not to be the correct prequantum Hilbert space to consider. The (square integrable) scalar functions do not have the correct transformation properties. The objects which do have the correct transformation behaviour are called **sections** (see section 3). However, since most of the readers of this thesis will be more familiar with scalar functions than with sections, to gain intuition for what prequantization actually is about, it pays to first naively consider the wrong prequantum Hilbert space consisting of (square-integrable) scalar functions. Immediately diving into new formalism would only distract us from the main idea of this section.

In the quantization process an important role is played by the identity proved in the appendix,

$$[X_f, X_g] = X_{\{f,g\}}, \tag{39}$$

where $f, g \in C^\infty(M,\mathbb{R})$ and $X_f, X_g, X_{\{f,g\}}$ are Hamiltonian vector fields on M. Using definition 3.4.6, with $C^\infty(M,\mathbb{R})$ the linear space and $(C^\infty(M,\mathbb{R}), \{\cdot,\cdot\})$ the Lie algebra, this identity shows that Hamiltonian vector fields give a representation of the classical Poisson bracket algebra by first order differential operators on M. To see this, notice that a Hamiltonian vector field is an element of $\mathrm{End}(C^\infty(M,\mathbb{R}))$ and that the map $\phi : C^\infty(M,\mathbb{R}) \to \mathrm{End}(C^\infty(M,\mathbb{R}))$ $\ f \mapsto X_f$ is a Lie algebra homomorphism from $(C^\infty(M,\mathbb{R}), \{\cdot,\cdot\})$ into $\mathrm{End}(C^\infty(M,\mathbb{R}))$. Indeed,

$$\phi(\{f,g\}) = X_{\{f,g\}} = [X_f, X_g] = [\phi(f), \phi(g)].$$

The classical phase space M carries a natural measure,

$$\epsilon_\omega \equiv (-1)^{\frac{1}{2}n(n-1)} \frac{1}{n!}\omega^n \equiv (-1)^{\frac{1}{2}n(n-1)} \frac{1}{n!} \underbrace{\omega \wedge ... \wedge \omega}_{n \text{ times}}, \tag{59}$$

for $n \in \mathbb{Z}$ and $\omega = dp \wedge dq$, the symplectic form on T^*Q with Q a 1-dimensional manifold. ϵ_ω is called the **Liouville form** .

Let the Hilbert space associated with M be the space $L^2(M, \mathbb{C})$ of square-integrable, smooth, complex-valued functions on M, with the inner product

$$\langle \phi, \psi \rangle = \left(\frac{1}{2\pi\hbar} \right)^n \int_M \phi \bar{\psi} \epsilon_\omega; \quad \phi, \psi \in L^2(M, \mathbb{C}). \tag{60}$$

Recall that a $\phi \in L^2(M, \mathbb{C})$ is called square-integrable if $\int_M \|\phi\|^2 \epsilon_\omega$ is finite. Here $\|\cdot\|$ is the norm on the Hilbert space $L^2(M, \mathbb{C})$ induced by the inner product $\langle \cdot, \cdot \rangle$.

Let us try to find the assignment Q in (54). Using definition 3.4.1 (and the remark beneath), each classical observable $f \in C^\infty(M, \mathbb{R})$ acts on the set $L^2(M, \mathbb{C})$ according to

$$\phi \mapsto -i\hbar X_f \phi; \quad \phi \in L^2(M, \mathbb{C}). \tag{61}$$

To see this, note that $C^\infty(M, \mathbb{R})$ is a group under pointwise addition and that the map $Q : f \mapsto -i\hbar X_f$ is a group homomorphism. Indeed, using (A.1), it follows that $Q(f + g) = -i\hbar X_{f+g} = -i\hbar X_f - i\hbar X_g = Q(f) + Q(g)$. Furthermore, since the map in (61) is clearly linear, it follows that Q is a representation of $C^\infty(M, \mathbb{R})$ (with pointwise addition as group operation) in $L^2(M, \mathbb{C})$.

The assignment Q in (54) corresponding with this representation is indeed the map Q defined above. For simplicity, let us denote the extension of (Q1)-(Q4) to all $f \in C^\infty(M, \mathbb{R})$ just by the same (Q1)-(Q4). Unfortunately, Q cannot be a correct prequantization, since condition $(Q4)$ is not satisfied. Indeed, $Q(1) = -i\hbar X_1 = \bar{0} \neq I$. Here 1 and I are the constant function, mapping every element in M to $1 \in \mathbb{R}$ and the identity operator defined previously.

Trying to remedy this, we propose a second attempt for the assignment Q,

$$Q : C^\infty(M, \mathbb{R}) \to \mathrm{End}(L^2(M, \mathbb{C})) \quad f \mapsto -i\hbar X_f + fI \tag{62}$$

In this case, we indeed have $Q(1) = -i\hbar X_1 + 1I = 1I = I$. But unfortunately, in this case (Q2) is not satisfied, since

$$
\begin{aligned}
&[Q(f), Q(g)] = [-i\hbar X_f + fI, -i\hbar X_g + gI] = \\
&-\hbar^2[X_f, X_g] - i\hbar[X_f, gI] - i\hbar[fI, X_g] + [fI, gI] = \\
&-\hbar^2 X_{\{f,g\}} - i\hbar X_f(g)I + i\hbar X_g(f)I + [f,g]I = \\
&-\hbar^2 X_{\{f,g\}} - 2i\hbar\{f,g\}I = -i\hbar(-i\hbar X_{\{f,g\}} + 2\{f,g\}I) \neq -i\hbar Q(\{f,g\}).
\end{aligned}
$$

In the third identity, we used the identity (41) and the fact that for $\phi \in L^2(M, \mathbb{C})$ the identity $[X_f, g](\phi) = X_f(g\phi) - gX_f(\phi) = X_f(g)\phi + gX_f(\phi) - gX_f(\phi) = X_f(g)\phi$ holds. In the fourth identity we used equation (49) and the commutatitivity of real-valued smooth scalar function under pointwise multiplication.

Suppose, $\omega = d\theta$ for some 1-form θ, the symplectic potential. Adding another term, the assignment

$$
Q : C^\infty(M, \mathbb{R}) \to \mathrm{End}(L^2(M, \mathbb{C})) \quad f \mapsto -i\hbar X_f + fI - \theta(X_f)I, \tag{63}
$$

does satisfy all conditions (Q1)-(Q4). Indeed,

- (Q1): Let $r \in \mathbb{R}$, $f, g \in C^\infty(M, \mathbb{R})$. Then

  $$
  \begin{aligned}
  &Q(rf + g) = -i\hbar X_{rf+g} + (rf + g)I - \theta(X_{rf+g})I = \\
  &rX_f + X_g + rfI + gI - \theta(rX_f + X_g)I = \\
  &rX_f + X_g + rfI + gI - r\theta(X_f)I - \theta(X_g)I = rQ(f) + Q(g).
  \end{aligned}
  $$

- (Q2): Noting that real-valued smooth scalar functions commute under pointwise multiplication, for $f, g \in C^\infty(M, \mathbb{R})$,

  $$
  \begin{aligned}
  &[Q(f), Q(g)] = [-i\hbar X_f + fI - \theta(X_f)I, -i\hbar X_g + gI - \theta(X_g)I] = \\
  &-\hbar^2 X_{\{f,g\}} - 2i\hbar\{f,g\}I + i\hbar[X_f, \theta(X_g)I] + i\hbar[\theta(X_f)I, X_g] = \\
  &-\hbar^2 X_{\{f,g\}} - 2i\hbar\{f,g\}I + i\hbar[X_f, \theta(X_g)I] + i\hbar[\theta(X_f)I, X_g] = \\
  &-\hbar^2 X_{\{f,g\}} - 2i\hbar\{f,g\}I + i\hbar(X_f(\theta(X_g))I - X_g(\theta(X_f))I.
  \end{aligned}
  $$

  Simplifying further, using

  $$
  \begin{aligned}
  &\theta(X_f) = \theta_i dx^i(X_f^j \tfrac{\partial}{\partial x^j}) = \theta_i X_f^j dx^i(\tfrac{\partial}{\partial x^j}) = \theta_i X_f^i, \text{ it follows,} \\
  &X_f(\theta(X_g)) - X_g(\theta(X_f) = X_f^i \partial_i(\theta_j X_g^j) - X_g^i \partial_i(\theta_j X_f^j) = \\
  &X_f^i(\partial_i\theta_j)X_g^j + X_f^i\theta_j(\partial_i X_g^j) + X_g^i(\partial_i\theta_j)X_f^j + X_g^i\theta_j(\partial_i X_f^j) = \\
  &2\omega_{ij} X_i^f X_j^g + \theta_j(X_f^i \partial_i X_g^j - X_g^i \partial_i X_f^j = \{f,g\} + \theta_j[X_f, X_g]^j =
  \end{aligned}
  $$

$\{f,g\} + \theta_j X^j_{\{f,g\}} = \{f,g\} + \theta(X_{\{f,g\}})$.

This indeed leads to

$[Q(f), Q(g)] = -\hbar^2 X_{\{f,g\}} - 2i\hbar\{f,g\}I + i\hbar I(\{f,g\} + \theta(X_{\{f,g\}}) =$
$-i\hbar(-i\hbar X_{\{f,g\}} - \theta(X_{\{f,g\}}) + \{f,g\}) = -i\hbar Q(\{f,g\})$.

- (Q3): Using the expression for the inner product, (60), on $L^2(M,\mathbb{C})$, for $f \in C^\infty(M,\mathbb{R})$, $\phi,\psi \in L^2(M,\mathbb{C})$, we have

  $\langle Q(f)\phi, \psi\rangle = \left(\frac{1}{2\pi\hbar}\right)^n \int_M (-i\hbar X_f(\phi) + f\phi - \theta(X_f)\phi)\bar{\psi}\epsilon_\omega =$
  $\left(\frac{1}{2\pi\hbar}\right)^n \int_M (\phi i\hbar X^i_f(\frac{\partial}{\partial x^i}\bar{\psi}) + f\phi\bar{\psi} - \theta(X_f)\phi\bar{\psi})\epsilon_\omega =$
  $\left(\frac{1}{2\pi\hbar}\right)^n \int_M (\phi\overline{-i\hbar X^i_f(\frac{\partial}{\partial x^i}\psi)} + \phi\overline{f\psi} - \phi\overline{\theta(X_f)\psi})\epsilon_\omega =$
  $\left(\frac{1}{2\pi\hbar}\right)^n \int_M \phi\overline{Q(f)\psi}\epsilon_\omega = \langle\phi, Q(f)\psi\rangle$.

  Here, in the second equality we used partial integration in the first term of $Q(f)$ (choosing the boundary conditions on $\phi,\psi \in C^\infty(M,\mathbb{R})$ such that the 'boundary term' appearing after partial integration vanishes), the identity $X_f = X^i_f\frac{\partial}{\partial x^i}$ and the invariance of ϵ_ω under X_f. In the third equality is used that f, X^i_f and $\theta(X_f)$ are real-valued. Furthermore, the hermiticity of the derivative operator is used here.

- (Q4): $Q(1) = -i\hbar X_1 + 1I - \theta(X_1)I = 1I - 0I = I$, with 0 the zero map difined previously. Indeed, by linearity of θ it follows that $\theta(X_1) = 0$, since $\theta(\bar{0}) = \theta(r\bar{0}) = r\theta(\bar{0}) \Rightarrow \theta(X_1) = \theta(\bar{0}) = 0$ $(r \in \mathbb{R})$.

However, there is still a difficulty with the last assignment Q. It depends on the choice for the symplectic potential θ. So Q cannot be defined globally unless ω is exact (see Darboux' theorem). At first sight, a way to escape this problem, is by means of so-called **gauge-transformations** : replace θ by $\theta' = \theta + du$ for some $u \in C^\infty(M,\mathbb{R})$. Then, ω is invariant under this transformation, since $\omega' = d\theta' = d(\theta + du) = d\theta + ddu = d\theta = \omega$. Replace $\phi \in L^2(M,\mathbb{C})$ by $\phi' = \exp(iu/\hbar)\phi$. Then the inner-product (and hence the physics involved, since the inner product is the probability measure) is invariant under this transformation. Indeed, for $\phi,\psi \in L^2(M,\mathbb{C})$,

$\langle\exp(iu/\hbar)\phi, \exp(iu/\hbar)\psi\rangle = \left(\frac{1}{2\pi\hbar}\right)^n \int_M \exp(iu/\hbar)\phi\exp(-i\bar{u}/\hbar)\bar{\psi}\epsilon_\omega =$
$\left(\frac{1}{2\pi\hbar}\right)^n \int_M \exp(iu/\hbar)\phi\exp(-iu/\hbar)\bar{\psi}\epsilon_\omega = \langle\phi,\psi\rangle$.

The transformed quantization assignment equals

$-i\hbar X_f\phi' - \theta'(X_f)\phi' + f\phi' =$
$-i\hbar X_f(\exp(iu/\hbar)\phi) - (\theta + du)(X_f)\exp(iu/\hbar)\phi + f\exp(iu/\hbar)\phi =$
$-i\hbar\exp(iu/\hbar)X_f(\phi) - i\hbar(i/\hbar)\exp(iu/\hbar)X_f(u)\phi - \theta(X_f)\exp(iu/\hbar)\phi-$
$X_f(u)\exp(iu/\hbar)\phi + f\exp(iu/\hbar)\phi =$
$-i\hbar\exp(iu/\hbar)X_f(\phi) - \theta(X_f)\exp(iu/\hbar)\phi + f\exp(iu/\hbar)\phi =$
$\exp(iu/\hbar)(-i\hbar X_f(\phi) - \theta(X_f)\phi + f\phi).$

Hence $Q(f)$ becomes independent of the choice of symplectic potential.

Another problem remains: u is determined by θ and θ' only up to a constant, since $d(u + \text{Constant}) = du + d\text{Constant} = du$. So given θ, θ', and ϕ, there is an ambiguity in the overall 'phase' of ϕ':

$$\phi' = \exp(iu/\hbar)\phi = \exp(i(u{+}\text{Constant})/\hbar)\phi = \exp(iu/\hbar)\exp(i\text{Constant}/\hbar)\phi.$$

To take this into account, and to put the prequantization construction on a more rigorous foundation, $Q(f)$ must not act on functions in $L^2(M, \mathbb{C})$ (subject to gauge-transformations), but on the sections of a Hermitian line bundle-with-connection.

6.4 Formalizing the discussion

A natural way to continue is to explain what a Hermitian line bundle with connection is. Furthermore, for a Hermitian line bundle with connection to exist some condition needs to be satisfied. In order to understand this condition one needs to have some knowledge of Čech cohomology. This subsection serves to explain the new mathematical luggage needed in order to put the prequantization construction on a more rigorous foundation.

6.4.1 Putting more structure on vector bundles

To start with, noting that $\mathbb{C}$ is a vector space,

Definition 6.4.1.1. A **line bundle** is a vector bundle with $\mathbb{K} = \mathbb{C}$ and fibre dimension 1.

Let X be a line bundle over the real manifold M. Recall (see section 3) that a **section** over U of a line bundle X is a map $s : U \subset M \to X$ satisfying $\text{pr}(s(m)) = m \ \forall m \in U$. We denote the set of all smooth sections over U of a line bundle X with $\Gamma_X(U)$. It is a vector space, where the scalars are $\mathbb{C}$-valued functions on U.

Denote by $\text{Vect}(M, \mathbb{C})$ the space of smooth complex vector fields (so multiplication by complex numbers is included in the scalar multiplication) on M and by $C^\infty(M, \mathbb{C})$ the set of smooth complex-valued funtions on M. Now define,

Definition 6.4.1.2. A **connection** ∇ on X is a map $\nabla : \text{Vect}(M, \mathbb{C}) \to \text{End}(\Gamma_X(M))$ that assigns to each $X \in \text{Vect}(M, \mathbb{C})$ an operator ∇_X on $\Gamma_X(M)$ such that

(a) $\nabla_{fX+Y} = f\nabla_X + \nabla_Y; \quad f \in C^\infty(M, \mathbb{C}), X, Y \in \mathit{Vect}(M, \mathbb{C}),$

(b) $\nabla_X(fs) = X(f)s + f\nabla_X s; \quad X \in \mathit{Vect}(M, \mathbb{C}), s \in \Gamma_X(M),$

(c) $\nabla_X(s + s') = \nabla_X(s) + \nabla_X(s'); \quad X \in \mathit{Vect}(M, \mathbb{C}), s, s' \in \Gamma_X(M).$

A line bundle equipped with a connection is called a **line bundle-with-connection**. We still need more structure on the line bundle. Let $m \in M$, and let $s, s' \in \Gamma_X(M)$.

Definition 6.4.1.3. A **Hermitian structure** on a line bundle is an inner product
$\langle \cdot, \cdot \rangle : (\{m\} \times \mathbb{C}) \times (\{m\} \times \mathbb{C}) \cong \mathbb{C} \times \mathbb{C} \to \mathbb{R}$
$s(m) \times s'(m) \mapsto \langle s, s' \rangle(m) \equiv \langle s(m), s'(m) \rangle$
in each fibre that is smooth in the sense that
$H : X \to \mathbb{R} \quad (m, s(m)) \mapsto \langle s, s \rangle(m)$ is smooth.

A connection ∇ and a Hermitian structure $\langle \cdot, \cdot \rangle$ are called **compatible** if for every $X \in \text{Vect}(M, \mathbb{C})$ and for every $s, s' \in \Gamma_X(M)$ we have

$$X\langle s, s' \rangle = \langle \nabla_{\bar{X}} s, s' \rangle + \langle s, \nabla_X s' \rangle. \tag{64}$$

Here $\overline{X}$ denotes the **formal complex conjugate** of the complex vector field X. The formal complex conjugate of X satisfies by definition the following rules,

- $\overline{X} + \overline{Y} = \overline{X + Y}$ with X,Y vector fields.
- $\lambda\overline{X} = \overline{\bar{\lambda}X}$ with X a vector field, $\lambda \in \mathbb{C}$ and $\bar{\lambda}$ the complex conjugate of λ.

A line bundle equipped with a connection and a Hermitian structure which are compatible is called a **Hermitian line bundle-with-connection** .

Let X be a line bundle-with-connection ∇ over M.

Definition 6.4.1.4. The **curvature 2-form** of ∇ is the complex differential 2-form Ω on M determined by

$$\Omega(X,Y)s = \frac{i}{2}([\nabla_X, \nabla_Y] - \nabla_{[X,Y]})s; \quad X, Y \in \mathit{Vect}(M), s \in \Gamma_X(M). \tag{65}$$

Here $[\nabla_X, \nabla_Y] = \nabla_X\nabla_Y - \nabla_Y\nabla_X$ is a commutator bracket. Indeed Ω defines a differential 2-form:

(1) Ω is anti-symmetric in $X, Y \in \mathrm{Vect}(M)$, since

$\Omega(X,Y) = \frac{i}{2}([\nabla_X, \nabla_Y] - \nabla_{[X,Y]}) = \frac{i}{2}(-[\nabla_Y, \nabla_X] - \nabla_{-[Y,X]}) = \frac{i}{2}(-[\nabla_Y, \nabla_X] + \nabla_{[Y,X]}) = -\Omega(Y,X)$.

In the third equality property (a) of the connection is used.

(2) Ω is bilinear in $X, Y \in \mathrm{Vect}(M)$. Suppose $X = f_1X_1 + X_2$ for $f_1 \in C^\infty(M,\mathbb{R}), X_1, X_2 \in \mathrm{Vect}(M)$. Then, using properties (a) and (b) of the connection,

$[\nabla_X, \nabla_Y]s = [\nabla_{f_1X_1+X_2}, \nabla_Y]s = [f_1\nabla_{X_1} + \nabla_{X_2}, \nabla_Y]s =$
$[f_1\nabla_{X_1}, \nabla_Y]s + [\nabla_{X_2}, \nabla_Y]s = f_1\nabla_{X_1}(\nabla_Y s) - \nabla_Y(f_1\nabla_{X_1}s) + [\nabla_{X_2}, \nabla_Y]s =$
$f_1\nabla_{X_1}(\nabla_Y s) - f_1\nabla_Y(\nabla_{X_1}s) - Y(f_1)\nabla_{X_1}s + [\nabla_{X_2}, \nabla_Y]s =$
$f_1[\nabla_{X_1}, \nabla_Y]s + [\nabla_{X_2}, \nabla_Y]s - Y(f_1)\nabla_{X_1}s.$

And,

$\nabla_{[X,Y]}s = \nabla_{[f_1X_1+X_2,Y]}s = \nabla_{[f_1X_1,Y]}s + \nabla_{[X_2,Y]}s =$
$\nabla_{f_1[X_1,Y]}s + \nabla_{-Y(f_1)X_1}s + \nabla_{[X_2,Y]}s = f_1\nabla_{[X_1,Y]}s - Y(f_1)\nabla_{X_1}s + \nabla_{[X_2,Y]}s.$

In the third equality is used that $[f_1X_1, Y] = f_1[X_1, Y] - Y(f_1)X_1$ is satisfied by the Lie bracket for vector fields. Consequently,

$\Omega(X,Y)s =$
$\frac{i}{2}(f_1\nabla_{[X_1,Y]}s - Y(f_1)\nabla_{X_1}s + \nabla_{[X_2,Y]}s - f_1\nabla_{[X_1,Y]}s + Y(f_1)\nabla_{X_1}s - \nabla_{[X_2,Y]}s) =$
$f_1\Omega(X_1,Y) + \Omega(X_2,Y).$

Hence, Ω is linear in its first argument. That it is linear in its second argument is proved fully analogously.

Furthermore, Ω is also linear in s over $C^\infty(M,\mathbb{C})$. Firstly, using property (c) of the connection, for $f_1 \in C^\infty(M,\mathbb{C}), s_1, s_2 \in \Gamma_X(M)$,

$[\nabla_X, \nabla_Y](f_1 s_1 + s_2) = \nabla_X(\nabla_Y(f_1 s_1)) - \nabla_Y(\nabla_X(f_1 s_1)) + [\nabla_X, \nabla_Y]s_2 =$
$\nabla_X(Y(f_1)s_1 + f_1\nabla_Y s_1) - \nabla_Y(X(f_1)s_1 + f_1\nabla_X s_1 + [\nabla_X, \nabla_Y]s_2 =$
$X(Y(f_1))s_1 + Y(f_1)\nabla_X s_1 + X(f_1)\nabla_Y s_1 + f_1\nabla_X(\nabla_Y s_1) - Y(X(f_1))s_1$
$- X(f_1)\nabla_Y s_1 - Y(f_1)\nabla_X s_1 - f_1\nabla_Y(\nabla_X s_1) + [\nabla_X, \nabla_Y]s_2 =$
$f_1[\nabla_X, \nabla_Y]s_1 + [\nabla_X, \nabla_Y]s_2 + [X, Y](f_1)s_1.$

And,

$\nabla_{[X,Y]}(f_1 s_1 + s_2) = [X, Y](f_1)s_1 + f_1\nabla_{[X,Y]}s_1 + \nabla_{[X,Y]}s_2.$

It follows that,

$\Omega(X, Y)(f_1 s_1 + s_2) =$
$\frac{i}{2}(f_1[\nabla_X, \nabla_Y]s_1 + [\nabla_X, \nabla_Y]s_2 + [X, Y](f_1)s_1$
$- [X, Y](f_1)s_1 - f_1\nabla_{[X,Y]}s_1 - \nabla_{[X,Y]}s_2) =$
$f_1\Omega(X, Y)s_1 + \Omega(X, Y)s_2.$

Let (U, ψ) be a local trivialization of X. Define the map

$$\hat{1} : U \subset M \to X \;\; m \mapsto \psi(m, 1(m)). \tag{66}$$

where $1 \in C^\infty(U, \mathbb{C})$ denotes the identity map $1 : U \to \mathbb{C} \;\; m \mapsto 1$. The map (68) is clearly a section on $U \subset M$, since the image of $\psi|_m$ is $\mathrm{pr}^{-1}(m)$. It follows that $\mathrm{pr}(s(m)) = m$, as was to be shown. It is called the **unit section**. Let $X \in \mathrm{Vect}(U)$. Let β be the complex 1-form defined on U by

$$\beta(X)\hat{1} = i\nabla_X \hat{1}. \tag{67}$$

Notice that one can always define such a 1-form β in this way, since, first of all, both sides of the equation are linear in X. On the one hand, β is a 1-form and so by definition linear in X. On the other hand, the connection is by definition linear in X by property (a) of the connection. Furthermore, every function can be written as the composition of a 1-form and a vector field (by definition) and every section can be written as a complex-valued function times some other section (since $\Gamma_X(M)$ is a vector space with complex-valued functions as scalars). β is called the **potential 1-form** for ∇. We claim that a general section s on $U \subset M$ can be written as a complex-valued function times the unit section. Indeed a general section s on $U \subset M$ can be written as

$$s : U \subset M \to X \;\; m \mapsto \psi(m, f(m)), \tag{68}$$

where $f \in C^\infty(U, \mathbb{C})$ is a general smooth complex-valued function. Since $1(m) = 1 \in \mathbb{C}\ \forall m \in U$ is a complex basis vector for the one-dimensional vector space $\mathbb{C}$, it follows that $\hat{1}$ is a basis vector for the complex one-dimensional vector space $\Gamma_X(M)$. The claim follows. It follows that we can write the connection of a general section s as,

$$\begin{aligned} \nabla_X s = \nabla_X \left(f\hat{1}\right) = X(f)\hat{1} + f\nabla_X \hat{1} = \\ X(f)\hat{1} + f\left(-i\beta(X)\hat{1}\right) = X(f)\hat{1} - i\beta(X)s. \end{aligned} \tag{69}$$

With expression (69) we are able to show that the curvature 2-form of ∇ is a closed form. Writing out the right-hand side of definition (65) for $X, Y \in \mathrm{Vect}(U)$, $s \in \Gamma_X(U)$, using (69) and (67), gives

$$\begin{aligned}
&\tfrac{i}{2}([\nabla_X, \nabla_Y] - \nabla_{[X,Y]})s = \tfrac{i}{2}(\nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]})s = \\
&\tfrac{i}{2}\nabla_X(Y(f)\hat{1} - i\beta(Y)s) - \tfrac{i}{2}\nabla_Y(X(f)\hat{1} - i\beta(X)s) - \tfrac{i}{2}([X,Y](f)\hat{1} - i\beta([X,Y])s) = \\
&\tfrac{i}{2}X(Y(f))\hat{1} + \tfrac{i}{2}Y(f)\nabla_X\hat{1} + \tfrac{i}{2}(-iX(\beta(Y))s) + \tfrac{i}{2}(-i\beta(Y)\nabla_X s) - \tfrac{i}{2}Y(X(f))\hat{1} \\
&- \tfrac{i}{2}X(f)\nabla_Y\hat{1} - \tfrac{i}{2}(-iY(\beta(X))s) - \tfrac{i}{2}(-i\beta(X)\nabla_Y s) \\
&- \tfrac{i}{2}([X,Y](f)\hat{1} - i\beta([X,Y])s) = \\
&\tfrac{i}{2}Y(f)\nabla_X\hat{1} + \tfrac{i}{2}(-iX(\beta(Y))s) + \tfrac{i}{2}(-i\beta(Y)\nabla_X s) - \tfrac{i}{2}X(f)\nabla_Y\hat{1} \\
&- \tfrac{i}{2}(-iY(\beta(X))s) - \tfrac{i}{2}(-i\beta(X)\nabla_Y s) - \tfrac{i}{2}(-i\beta([X,Y])s) = \\
&\tfrac{1}{2}X(\beta(Y))s - \tfrac{1}{2}Y(\beta(X))s - \tfrac{1}{2}\beta([X,Y])s + \tfrac{i}{2}Y(f)(-i\beta(X)\hat{1}) - \tfrac{i}{2}X(f)(-i\beta(Y)\hat{1}) \\
&+ \tfrac{1}{2}\beta(Y)X(f)\hat{1} - \tfrac{1}{2}\beta(X)Y(f)\hat{1} - \tfrac{i}{2}\beta(Y)\beta(X)s + \tfrac{i}{2}\beta(X)\beta(Y)s = \\
&\tfrac{1}{2}X(\beta(Y))s - \tfrac{1}{2}Y(\beta(X))s - \tfrac{1}{2}\beta([X,Y])s.
\end{aligned}$$

In the fourth equality is used that $X(Y(f)) - Y(X(f)) = [X,Y](f)$ for $X, Y \in \mathrm{Vect}(U), f \in C^\infty(U, \mathbb{C})$. As a next step to establish the closedness of the curvature 2-form Ω, we prove that

$$d\beta(X,Y)s = \tfrac{1}{2}X(\beta(Y))s - \tfrac{1}{2}Y(\beta(X))s - \tfrac{1}{2}\beta([X,Y])s,$$

the expression with which the derivation above ended. Denote the set of local coordinates on the n-dimensional manifold B by $\{x^i\}_{1\leq i\leq n}$. Writing $X = X^i\frac{\partial}{\partial x^i}$ for $X \in \mathrm{Vect}(U)$ and $\beta = \beta_j dx^j$ for the potential 1-form β, it indeed follows that

$$\begin{aligned}
&d\beta(X,Y) = d(\beta_j dx^j)(X,Y) = (d\beta_j \wedge dx^j)(X,Y) = \\
&(\tfrac{\partial\beta_j}{\partial x^i}dx^i \wedge dx^j)(X^i\tfrac{\partial}{\partial x^k}, Y^m\tfrac{\partial}{\partial x^m}) = \tfrac{1}{2}\tfrac{\partial\beta_j}{\partial x^i}(X^iY^j - Y^iX^j) = \\
&\tfrac{1}{2}(X^i\tfrac{\partial\beta_j}{\partial x^i}Y^j - Y^i\tfrac{\partial\beta_j}{\partial x^i}X^j) = \\
&\tfrac{1}{2}(X^i\tfrac{\partial}{\partial x^i}(\beta_jY^j) - \beta_jX^i\tfrac{\partial}{\partial x^i}Y^j - Y^i\tfrac{\partial}{\partial x^i}(\beta_jX^j) + \beta_jY^i\tfrac{\partial}{\partial x^i}X^j) = \\
&\tfrac{1}{2}(X\beta(Y) - Y\beta(X) - \beta_j[X,Y]^j) = \tfrac{1}{2}X(\beta(Y)) - \tfrac{1}{2}Y(\beta(X)) - \tfrac{1}{2}\beta([X,Y]).
\end{aligned}$$

It follows that on every neighbourhood $U \subset M$ we have $\Omega = d\beta$. Hence on every $U \subset M$ the identity $d\Omega = (d \circ d)\beta = 0$ holds, hence Ω is closed.

Another 'tool' we will need in the next section are so-called transition functions. Let (U_1, ψ_1) and (U_2, ψ_2) be two local trivializations of M and take the intersection $U_1 \cap U_2$ non-empty.

Definition 6.4.1.5. A **transition function** between (U_1, ψ_1) and (U_2, ψ_2) is the function $c_{12} : U_1 \cap U_2 \to \mathbb{C}$ defined by

$$\psi_2|_{(U_1 \cap U_2)\times\mathbb{C}}(m, z) = c_{12}(m)\psi_1|_{(U_1 \cap U_2)\times\mathbb{C}}(m, z). \tag{70}$$

Notice that this function is well defined, since $\psi_1(m, z)$ and $\psi_2(m, z)$ both take values in the fibre over m. Indeed, by definition
$\psi_{1,2}|_m : \{m\} \times \mathbb{C} \to X_m$, so that $\psi_{1,2}(m, z) = \psi_{1,2}|_m(z)$ are values in X_m. Since the $X_m \cong \mathbb{C}$ are by definition vector spaces over $\mathbb{C}$, it follows that its elements can be multiplied by complex numbers, such as $c_{12}(m)$.

We quote without proof, that it is always possible to find a collection $\{(U_j, \psi_j)\}$ of local trivializations of X such that $\{U_j\}$ is a contractible open cover of M (which means that each U_j and each finite intersection of U_js is contractible to a point). Denote the unit section determined by (U_j, ψ_j) by $\hat{1}_j$. Then we know from previous discussions that any $s \in \Gamma(M)$ can be written as $s_j = f\hat{1}_j$ on U_j for some $f \in C^\infty(U_j, \mathbb{C})$. In particular, on each nonempty $U_j \cap U_i \equiv U_{ji}$,

$$\hat{1}_i|_{U_{ji}}(m) = c_{ji}(m)\hat{1}_j|_{U_{ji}}(m), \tag{71}$$

where c_{ji} is the transition function between (U_j, ψ_j) and (U_i, ψ_i). Indeed, c_{ji} being a transition function means by definition that it satisfies the equation $\psi_i|_{U_{ji}\times\mathbb{C}}(m, z) = c_{ji}(m)\psi_j|_{U_{ji}\times\mathbb{C}}(m, z)$. It follows that

$$\hat{1}_i|_{U_{ji}} = \psi_i|_{U_{ji}\times\mathbb{C}}(m, 1) = c_{ji}(m)\psi_j|_{U_{ji}\times\mathbb{C}}(m, 1) = c_{ji}(m)\hat{1}_j|_{U_{ji}}(m),$$

as stated by (71).

Whenever $U_k \cap U_j \cap U_i \equiv U_{kji}$ is nonempty, the transitions functions satisfy the relations

$$c_{ij} = (c_{ji})^{-1}, \tag{72}$$

$$c_{kj}|_{U_{ikj}}(m)c_{ji}|_{U_{ikj}}(m)c_{ik}|_{U_{ikj}}(m) = 1. \tag{73}$$

Notice that $(c_{ji})^{-1}$ denotes the inverse of (c_{ji}) with respect to pointwise multiplication. First we prove property (72). Note that c_{ji} satisfies $\psi_i|_{U_{ji}\times\mathbb{C}}(m,z) = c_{ji}(m)\psi_j|_{U_{ji}\times\mathbb{C}}(m,z)$ and c_{ij} satisfies $\psi_j|_{U_{ij}\times\mathbb{C}}(m,z) = c_{ij}(m)\psi_i|_{U_{ij}\times\mathbb{C}}(m,z)$. Consequently, $c_{ij} = (c_{ji})^{-1}$ is smooth (since ψ_i, ψ_j are smooth) and

$\psi_i|_{U_{ji}\times\mathbb{C}}(m,z)\psi_j|_{U_{ij}\times\mathbb{C}}(m,z) = c_{ji}(m)c_{ij}(m)\psi_i|_{U_{ij}\times\mathbb{C}}(m,z)\psi_j|_{U_{ji}\times\mathbb{C}}(m,z) \Leftrightarrow$
$c_{ji}(m)c_{ij}(m) = 1 \Leftrightarrow c_{ij} = (c_{ji})^{-1}$.

Here is used that $U_{ij} = U_{ji}$. To prove relation (73) we first construct functions $\tilde{c_{ji}} : U_{ji}\times\mathbb{C} \to U_{ji}\times\mathbb{C}$ from the original complex-valued scalar functions $c_{ji} : U_{ji} \to \mathbb{C}$. This can be done by means of the map

$$\tilde{c_{ji}} : U_{ji}\times\mathbb{C} \to U_{ji}\times\mathbb{C}, \;\; (m,z) \mapsto (m, c_{ji}(m)z) \tag{74}$$

which is well-defined, since $m \in U_{ji}$ and $c_{ji}(m)z \in \mathbb{C}$. Notice that this map is a diffeomorphism,

- $\tilde{c_{ji}}$ is smooth: This follows directly from the definition, since c_{ji} is smooth.
- The inverse $\tilde{c_{ji}}^{-1}$ exists with respect to composition and is smooth: the inverse map is given by
  $\tilde{c_{ji}}^{-1} : U_{ji}\times\mathbb{C} \to U_{ji}\times\mathbb{C} \;\; (m,z) \mapsto (m, c_{ji}^{-1}(m)z)$. Indeed,

  $(\tilde{c_{ji}} \circ \tilde{c_{ji}}^{-1})(m,z) = \tilde{c_{ji}}(m, c_{ji}^{-1}(m)z) = (m, c_{ji}(m)c_{ji}^{-1}(m)z) = (m,z)$.

  Smoothness follows from the fact that $c_{ji}^{-1} = c_{ij}$ is smooth.
- $\tilde{c_{ji}}$ is bijective: it is surjective, since the expression $c_{ji}(m)z$ spans $\mathbb{C}$. For injectivity take $\tilde{c_{ji}}(m',z') = \tilde{c_{ji}}(m,z) \Leftrightarrow (m', c_{ji}(m')z') = (m, c_{ji}(m)z)$. This implies that $m = m'$ and so $c_{ji}(m)z' = c_{ji}(m)z \Leftrightarrow z = z'$. Injectivity follows.

Furthermore, the restriction
$\tilde{c_{ji}}|_{\{m\}\times\mathbb{C}} : \{m\}\times\mathbb{C} \cong \mathbb{C} \to \{m\}\times\mathbb{C} \cong \mathbb{C} \;\; z \mapsto c_{ji}(m)z$ is a linear isomorphism,

- $\tilde{c_{ji}}|_{\{m\}\times\mathbb{C}}$ is linear: all linear maps from $\mathbb{C}$ to $\mathbb{C}$ are classified by maps of the form $f : \mathbb{C} \to \mathbb{C} \;\; z \mapsto cz$ where $c \in \mathbb{C}$. Indeed $c_{ji}(m) \in \mathbb{C}$, so linearity follows.
- $\tilde{c_{ji}}|_{\{m\}\times\mathbb{C}}$ is bijective: This follows directly from the discussion above.

- $\tilde{c_{ji}}|_{\{m\}\times\mathbb{C}}$ is a homomorphism under addition, since $\tilde{c_{ji}}|_{\{m\}\times\mathbb{C}}(z_1+z_2) = c_{ji}(z_1+z_2) = c_{ji}z_1 + c_{ji}z_2 = \tilde{c_{ji}}|_{\{m\}\times\mathbb{C}}(z_1) + \tilde{c_{ji}}|_{\{m\}\times\mathbb{C}}(z_2)$ for $z_1, z_2 \in \mathbb{C}$.

From this discussion it follows that we can write
$\tilde{c_{ji}} = (\psi_i^{-1} \circ \psi_j)|_{U_{ij}\times\mathbb{C}} : U_{ij} \times \mathbb{C} \xrightarrow{\psi_j|_{U_{ij}\times\mathbb{C}}} \mathrm{pr}^{-1}(U_{ij}) \xrightarrow{\psi_i^{-1}|\mathrm{pr}^{-1}(U_{ij})} U_{ij} \times \mathbb{C}$.
Consequently,

$(\tilde{c_{kj}} \circ \tilde{c_{ji}} \circ \tilde{c_{ik}})|_{U_{ijk}\times\mathbb{C}} = (\psi_j^{-1} \circ \psi_k)|_{U_{ijk}\times\mathbb{C}} \circ (\psi_i^{-1} \circ \psi_j)|_{U_{ijk}\times\mathbb{C}} \circ (\psi_k^{-1} \circ \psi_i)|_{U_{ijk}\times\mathbb{C}} = (\psi_k^{-1} \circ \psi_i \circ \psi_i^{-1} \circ \psi_j \circ \psi_j^{-1} \circ \psi_k)|_{U_{ijk}\times\mathbb{C}} = \tilde{1_{ij}}|_{U_{ijk}\times\mathbb{C}}$,

where $\tilde{1_{ij}} : U_{ij} \times \mathbb{C} \to U_{ij} \times \mathbb{C}\ \ (m,z) \mapsto (m,z)$. Using (74) we derive

$(\tilde{c_{kj}} \circ \tilde{c_{ji}} \circ \tilde{c_{ik}})|_{U_{ijk}\times\mathbb{C}}(m,z) = (m, c_{kj}|_{U_{ijk}}(m) c_{ji}|_{U_{ijk}}(m) c_{ik}|_{U_{ijk}}(m) z) = (m,z) \Leftrightarrow c_{kj}|_{U_{ijk}}(m) c_{ji}|_{U_{ijk}}(m) c_{ik}|_{U_{ijk}}(m) = 1$,

which proves property (73).

6.4.2 Čech cohomology and the de Rham isomorphism

As stated before, for a Hermitian line bundle-with-connection to exist a condition needs to be satisfied, the so-called integrality condition. This is formulated in terms of elements of the Čech cohomology group. In this subsubsection it is therefore explained what Čech cohomology is and the related de Rham isomorphism will be discussed.

Let M be a manifold and let $U = \{U_i\}$ be an open cover of M, where $i \in I$, some index set.

Definition 6.4.2.1. A ***p*-cochain** Cochain relative to U is a collection $f = \{f_{i_1 i_2 \ldots i_{p+1}}\}$ of functions, of some specified type (i.e. smooth, locally constant, holomorphic) with the properties

(CO1) $f_{i_1 i_2 \ldots i_{p+1}}$ is defined on $U_{i_1} \cap U_{i_2} \cap \ldots \cap U_{i_{p+1}}$,

(CO2) f contains just one function $f_{i_1 i_2 \ldots i_{p+1}}$ for each ordered set of $p+1$ indices for which $U_{i_1} \cap U_{i_2} \cap \ldots \cap U_{i_{p+1}}$ is nonempty,

(CO3) $f_{i_1 i_2 \ldots i_{p+1}} = f_{[i_1 i_2 \ldots i_{p+1}]} \equiv \frac{1}{(p+1)!} \sum_\sigma f_{\sigma(i_1)\sigma(i_2)\ldots\sigma(i_{p+1})}$. Here σ denotes a permutation of the set of indices $\{i_1 i_2 \ldots i_{p+1}\}$.

Furthermore,

Definition 6.4.2.2. Let $U_{i_0} \in U$ and let $U_{i_0} \cap U_{i_1} \cap ... \cap U_{i_{p+1}}$ be nonempty. Then the **coboundary operator** Coboundary operator δ is defined on cochains by

$$\delta f = \{(\delta f)_{i_0 ... i_{p+1}}\}$$
$$(\delta f)_{i_0 ... i_{p+1}} = (p+2) \sum_{l=0}^{p+1} (-1)^l (f_{i_0 ... \hat{i_l} ... i_{p+1}})|_{U_{i_0} \cap ... \cap U_{i_{p+1}}} \tag{75}$$

Here $\hat{i_l}$ denotes omitting the l-th index of $f_{i_0 ... i_l ... i_{p+1}}$.

The coboundary operator δ takes p-cochains into $(p+1)$-cochains. Indeed,

- $(\delta f)_{i_0 i_1 ... i_{p+1}}$ is defined on $U_{i_0} \cap U_{i_1} \cap ... \cap U_{i_{p+1}}$: follows directly from the definition (75),
- δf contains just one function $(\delta f)_{i_0 i_1 ... i_{p+1}}$ for each ordered set of $p+2$ indices for which $U_{i_0} \cap U_{i_1} \cap ... \cap U_{i_{p+1}}$ is nonempty: follows from definition (75) and the fact that f contains just one function $f_{i_1 i_2 ... i_{p+1}}$ whenever $U_{i_1} \cap U_{i_2} \cap ... \cap U_{i_{p+1}}$ is nonempty,
- $(\delta f)_{i_0 i_1 ... i_{p+1}} = (\delta f)_{[i_0 i_1 ... i_{p+1}]}$: Indeed,
  $(\delta f)_{[i_0 i_1 ... i_{p+1}]} \equiv \frac{1}{(p+2)!} \sum_\sigma \text{sign}(\sigma)(\delta f)_{\sigma(i_0)\sigma(i_1)...\sigma(i_{p+1})} =$
  $\frac{1}{(p+2)!}(p+2)!(\delta f)_{i_0 i_1 ... i_{p+1}} = (\delta f)_{i_0 i_1 ... i_{p+1}}$.

Definition 6.4.2.3. A p-cochain is called a **p-cocycle** if $\delta f = 0$ and a **p-coboundary** if $f = \delta g$ for some $(p-1)$-cochain g.

One can prove that $\delta^2 = 0$, which implies that every p-coboundary is also a p-cocycle. The only important case we consider later on in this thesis will be when $p = 2$ (i.e. that 2-coboundaries are also 2-cocycles). So for the purposes of this thesis it suffices to show $\delta^2 = 0$ for $p = 2$. A 2-cochain is a collection of functions $\{f_{i_1 i_2 i_3}\}$ of some specified type satisfying (CO1),(CO2) and (CO3). Let $U_{i_0} \cap U_{i_1} \cap ... \cap U_{i_4}$ be nonempty. Indeed,

$$\begin{aligned}
&(\delta(\delta f))_{i_0 i_1 i_2 i_3 i_4} \equiv \\
&5((\delta f)_{i_1 i_2 i_3 i_4} - (\delta f)_{i_0 i_2 i_3 i_4} + (\delta f)_{i_0 i_1 i_3 i_4} - (\delta f)_{i_0 i_1 i_2 i_4} + (\delta f)_{i_0 i_1 i_2 i_3})|_{U_{i_0 i_1 ... i_4}} = \\
&5\{4(f_{i_2 i_3 i_4} - f_{i_1 i_3 i_4} + f_{i_1 i_2 i_4} - f_{i_1 i_2 i_3}) - 4(f_{i_2 i_3 i_4} - f_{i_0 i_3 i_4} + f_{i_0 i_2 i_4} - f_{i_0 i_2 i_3}) + \\
&4(f_{i_1 i_3 i_4} - f_{i_0 i_3 i_4} + f_{i_0 i_1 i_4} - f_{i_0 i_1 i_3}) - 4(f_{i_1 i_2 i_4} - f_{i_0 i_2 i_4} + f_{i_0 i_1 i_4} - f_{i_0 i_1 i_2}) + \\
&4(f_{i_1 i_2 i_3} - f_{i_0 i_2 i_3} + f_{i_0 i_1 i_3} - f_{i_0 i_1 i_2})\}|_{U_{i_0 i_1 ... i_4}} = 0.
\end{aligned}$$

Consequently, $\delta^2 = 0$ when $p = 2$.

In the rest of this subsubsection let us consider a specified type of functions which is referred to in the definition for p-cocycles and that will also be used later on in this thesis. These are the locally constant functions, functions which are constant in a neighbourhood of each point and take values in $\mathbb{R}$. For these functions we denote the set of p-cochains by $C^p(U, \mathbb{R})$ and the set of p-cocycles by $Z^p(U, \mathbb{R})$. We can now define,

Definition 6.4.2.4. The **$\check{C}$ech cohomology group** with coefficients in $\mathbb{R}$, denoted by $H^p(U, \mathbb{R})$, is the quotient group $Z^p(U, \mathbb{R})/\delta C^{p-1}(U, \mathbb{R})$.

The equivalence class of $f \in Z^p(U, \mathbb{R})$ in $H^p(U, \mathbb{R})$ is denoted by $[f]$. Two cocycles f in $H^p(U, \mathbb{R})$ that are equivalent (that is, differ by a coboundary) are said to be **($\check{C}$ech) cohomologous**. Indeed the above described equivalence defines an equivalence relation $\sim$,

- $f \sim f$: follows from $f = f + \delta h$ for $h = 0$. Indeed when $h = 0$ we have that $\delta h = 0$ is a p-coboundary.
- $f \sim g \Rightarrow g \sim f$: since $f \sim g \Leftrightarrow f = g + \delta h$ for some coboundary δh, it follows that $g = f - \delta h = f + \delta(-h) \Rightarrow g \sim f$.
- $f \sim g, g \sim h \Rightarrow f \sim h$: if $f \sim g \Leftrightarrow f = g + \delta i$ and
  $g \sim h \Leftrightarrow g = h + \delta j$ then
  $f = h + \delta j + \delta i = h + \delta(i + j) \Rightarrow f \sim h$.

Let $U = \{U_i\}$ be a locally finite contractible cover of M. That is, each point in M has a neighbourhood intersecting only a finite number of U_is $\in U$ (locally finite cover of M) and each $U_i \in U$ and each finite intersection of U_is $\in U$ is contractible to a point (contractible cover of M). Now define,

Definition 6.4.2.5. A **partition of unity** subordinate to U is a collection of functions $\{h_i : M \to \mathbb{R}\}$ with the properties

(PU1) $0 \leq h_i(m) \leq 1$,

(PU2) $\mathrm{supp}(h_i) \subset U_i$,

(PU3) $\sum_i h_i(m) = 1$.

We quote without proof that for every manifold it is always possible to find a locally finite contractible cover U of M and a partition of unity subordinate to U (for the proof, see for example [21]).

For $f = \{f_{i_1 i_2 \ldots i_{p+1}}\} \in C^p(U, \mathbb{R})$, define $\tilde{f}_{i_1 i_2 \ldots i_{p+1}} : M \to \mathbb{R}$ by

$$\tilde{f}_{i_1 i_2 \ldots i_{p+1}} \equiv \begin{cases} f_{i_1 i_2 \ldots i_{p+1}} & x \in U_{i_1 i_2 \ldots i_{p+1}} \\ 0 & x \notin U_{i_1 i_2 \ldots i_{p+1}} \end{cases}$$

Subsequently, define $\alpha : C^p(U, \mathbb{R}) \to \{p\text{-forms on } M\}$ by

$$\alpha(f) \equiv \alpha_f = \sum_{i_1, i_2, \ldots, i_{p+1}} \tilde{f}_{i_1 i_2 \ldots i_{p+1}} h_{i_1} dh_{i_2} \wedge \ldots \wedge dh_{i_{p+1}}. \tag{76}$$

From (76) it follows that

$$d(\alpha_f) = \alpha_{\delta f}. \tag{77}$$

Again, since the only important case we consider later on in this thesis will be when $p = 2$, we prove (77) for this case only. In this case, $\alpha_f = \sum_{i_1, i_2, i_3} \tilde{f}_{i_1 i_2 i_3} h_{i_1} dh_{i_2} \wedge dh_{i_3}$. The left-hand side of (77) gives

$d\alpha_f = d\left(\sum_{i_1, i_2, i_3} \tilde{f}_{i_1 i_2 i_3} h_{i_1} dh_{i_2} \wedge dh_{i_3}\right) =$
$\sum_{i_1, i_2, i_3} h_{i_1} d\tilde{f}_{i_1 i_2 i_3} \wedge dh_{i_2} \wedge dh_{i_3} + \sum_{i_1, i_2, i_3} \tilde{f}_{i_1 i_2 i_3} dh_{i_1} \wedge dh_{i_2} \wedge dh_{i_3} =$
$\sum_{i_1, i_2, i_3} \tilde{f}_{i_1 i_2 i_3} dh_{i_1} \wedge dh_{i_2} \wedge dh_{i_3}$

In the last identity is used that $d\tilde{f}_{i_1 i_2 i_3} = 0$, since $\tilde{f}_{i_1 i_2 i_3}$ is locally constant. Furthermore, it follows from $\sum_i h_i = 1$ that $d(\sum_i h_i) = \sum_i dh_i = 0$. Using this, the right-hand side of (77) gives

$\alpha_{\delta f} = \sum_{i_0, i_1, i_2, i_3} (\delta \tilde{f})_{i_0 i_1 i_2 i_3} h_{i_0} dh_{i_1} \wedge dh_{i_2} \wedge dh_{i_3} =$
$\sum_{i_0, i_1, i_2, i_3} (f_{i_1 i_2 i_3} - f_{i_0 i_2 i_3} + f_{i_0 i_1 i_3} - f_{i_0 i_1 i_2})|_{U_{i_0 i_1 i_2 i_3}} h_{i_0} dh_{i_1} \wedge dh_{i_2} \wedge dh_{i_3} =$
$\sum_{i_0, i_1, i_2, i_3} (\tilde{f}_{i_1 i_2 i_3} - \tilde{f}_{i_0 i_2 i_3} + \tilde{f}_{i_0 i_1 i_3} - \tilde{f}_{i_0 i_1 i_2}) h_{i_0} dh_{i_1} \wedge dh_{i_2} \wedge dh_{i_3} =$
$\sum_{i_0, i_1, i_2, i_3} \tilde{f}_{i_1 i_2 i_3} h_{i_0} dh_{i_1} \wedge dh_{i_2} \wedge dh_{i_3} = \sum_{i_1, i_2, i_3} \tilde{f}_{i_1 i_2 i_3} dh_{i_1} \wedge dh_{i_2} \wedge dh_{i_3}.$

In the third identity is used that $h_{i_0} dh_{i_1} \wedge dh_{i_2} \wedge dh_{i_3}$ equals zero outside $U_{i_0 i_1 i_2 i_3} \equiv U_{i_0} \cap U_{i_1} \cap U_{i_2} \cap U_{i_3}$. In the fourth identity is used that $\sum_{i_0} h_{i_0} = 1, \sum_{i_1} dh_{i_1} = \sum_{i_2} dh_{i_2} = \sum_{i_3} dh_{i_3} = 0$. Identity (77) now follows.

The map (76) and the identity (77) tell us that associated with any $[f] \in H^p(U, \mathbb{R})$ we have a set of closed p-forms, any two of which differ by an exact p-form. Indeed, take $[f] \in H^p(U, \mathbb{R})$. Then, using (76) and (77) respectively, it follows that $\delta f = 0 \Rightarrow \alpha_{\delta f} = d\alpha_f = 0$.

So α_f is indeed a closed p-form. The fact that any two of these differ by

an exact p-form follows from the following argumentation. In the deduction above the only freedom we have is in choosing the cocycle f satisfying $\delta f = 0$. We could also have chosen the cocycle $f+\delta h$, where δh is a coboundary, since $\delta(f + \delta h) = \delta f + \delta^2 h = \delta f$. So another closed p-form can be given by
$\alpha_{f+\delta h} = \alpha_f + \alpha_{\delta h} = \alpha_f + d\alpha_h$,
which indeed differs from the original α_f by an exact p-form $d\alpha_h$.

Now define,

Definition 6.4.2.6. The **pth de Rham cohomology group** of M , denoted by $\tilde{H}(M, \mathbb{R})$, is the space of equivalence classes of p-forms, α_f, with two forms regarded as equivalent if they differ by an exact p-form.

Indeed, the equivalence described in the definition above defines an equivalence relation $\sim$, since,

- $\alpha_f \sim \alpha_f$: follows from $\alpha_f = \alpha_f + d\alpha_h$ for $\alpha_h = 0$. Indeed when $\alpha_h = 0$ we have that $d\alpha_h = 0$ is an exact p-form.
- $\alpha_f \sim \alpha_g \Rightarrow \alpha_g \sim \alpha_f$: since $\alpha_f \sim \alpha_g \Leftrightarrow \alpha_f = \alpha_g + d\alpha_h$ for some exact form $d\alpha_h$, it follows that $\alpha_g = \alpha_f - d\alpha_h = \alpha_f + d\alpha_{-h} \Rightarrow \alpha_g \sim \alpha_f$.
- $\alpha_f \sim \alpha_g, \alpha_g \sim \alpha_h \Rightarrow \alpha_f \sim \alpha_h$: if $\alpha_f \sim \alpha_g \Leftrightarrow \alpha_f = \alpha_g + d\alpha_i$ and
  $\alpha_g \sim \alpha_h \Leftrightarrow \alpha_g = \alpha_h + d\alpha_j$ then
  $\alpha_f = \alpha_h + d\alpha_j + d\alpha_i = \alpha_h + d\alpha_{i+j} \Rightarrow \alpha_f \sim \alpha_h$.

In the rest of this subsection we prove that the map α in (76) determines an isomorphism between the pth Čech cohomology group and the pth de Rham cohomology group. This isomorphism is called the **de Rham isomorphism**.
. First define $\iota : H^p(U, \mathbb{R}) \to Z^p(U, \mathbb{R})\ \ [f] \mapsto f$ and
$\pi : \{\text{closed } p\text{-forms on } M\} \to \tilde{H}^p(U, \mathbb{R})\ \ \alpha_f \mapsto \langle\alpha_f\rangle$. Subsequently, define
$\tilde{\alpha} = \pi \circ \alpha|_{Z^p(U,\mathbb{R})} \circ \iota :$
$H^p(U, \mathbb{R}) \xrightarrow{\iota} Z^p(U, \mathbb{R}) \xrightarrow{\alpha|_{Z^p(U,\mathbb{R})}} \{\text{closed } p\text{-forms on } M\} \xrightarrow{\pi} \tilde{H}^p(U, \mathbb{R})$.
To prove the claim we have to show that $\tilde{\alpha}$ is an isomorphism. Indeed,

(1) $\tilde{\alpha}$ is a homomorphism: Indeed,
$\tilde{\alpha}([f] + [g]) = (\pi \circ \alpha|_{Z^p(U,\mathbb{R})} \circ \iota)([f] + [g])$
$= (\pi \circ \alpha|_{Z^p(U,\mathbb{R})} \circ \iota)([f + g]) = (\pi \circ \alpha_{Z^p(U,\mathbb{R})})(f + g) = \pi \circ \alpha_{Z^p(U,\mathbb{R})}(f + g)$
$= \pi \circ (\alpha_{Z^p(U,\mathbb{R})}(f) + \alpha_{Z^p(U,\mathbb{R})}(g)) = \langle\alpha_{Z^p(U,\mathbb{R})}(f) + \alpha_{Z^p(U,\mathbb{R})}(g)\rangle =$
$\langle\alpha_{Z^p(U,\mathbb{R})}(f)\rangle + \langle\alpha_{Z^p(U,\mathbb{R})}(g)\rangle = \tilde{\alpha}([f]) + \tilde{\alpha}([g])$.

(2) $\tilde{\alpha}$ is bijective: clearly ι and π are bijective. Furthermore, $\alpha|_{Z^p(U,\mathbb{R})}$ is bijective. Surjectivity of $\alpha|_{Z^p(U,\mathbb{R})}$ follows directly from (77), whereas injectivity follows directly from (76). This implies that the composition $\tilde{\alpha} = \pi \circ \alpha|_{Z^p(U,\mathbb{R})} \circ \iota$ is bijective too.

6.5 The integrality condition

To make the prequantization construction considered in subsection 6.3 rigourous one needs a Hermitian line bundle-with-connection over the classical phase space (the symplectic manifold M). Existence of such a line bundle-with-connection is however not guaranteed. For such a line bundle-with-connection to exist a certain condition should be satisfied by the curvature 2-form Ω defined on it. This is the so-called integrality condition (see [22],[19]). Let us formulate it:

(IC) Integrality Condition: Let M be a symplectic manifold and let Ω be a closed real 2-form on M (the symplectic form). There exists an open cover $U = \{U_j\}$ of M such that the equivalence class defined by $\frac{1}{2\pi}\Omega$ in $H^2(U, \mathbb{R})$ contains a cocycle z in which all the z_{ijk} are integers.

In the sequel of this section I shall prove that this integrality condition is both a necessary and sufficient condition for a Hermitian line bundle-with-connection to exist. The construction presented below is called the $\check{C}$**ech construction** .

Theorem 6.5.1. The Integrality Condition is both a necessary and sufficient condition for a Hermitian line bundle-with-connection to exist.

Proof. We proceed by showing seperately sufficience and necessarity of IC for a Hermitian line bundle-with-connection to exist.

- The Integrality Condition is a sufficient condition for a Hermitian line bundle-with-connection to exist:

  Let a closed real 2-form Ω on M (the symplectic form) and a contractible open cover $U = \{U_j\}$ of M satisfying (IC) be given. By Darboux's theorem there exist real 1-forms β_j on U_j (the symplectic potentials) such that on U_j,

  $$\Omega = d\beta_j. \tag{78}$$

  Furthermore, on each non-empty intersection $U_i \cap U_j$, there exists a smooth real function $\tilde{f}_{ij} = -\tilde{f}_{ji}$ such that

  $$d\tilde{f}_{ij} = \beta_i - \beta_j. \tag{79}$$

  Indeed $\beta_i - \beta_j$ is only well-defined on $U_i \cap U_j$, as is $\tilde{f}_{ij}$. Furthermore, since $d\tilde{f}_{ij} = \beta_i - \beta_j = -(\beta_j - \beta_i) = -d\tilde{f}_{ji}$ we can choose the functions such that they satisfy $\tilde{f}_{ij} = -\tilde{f}_{ji}$. Subsequently, on each non-empty triple intersection $U_i \cap U_j \cap U_k$ there exists a locally constant smooth real function $\tilde{a}_{ijk}$ defined by

$$\tilde{a}_{ijk} \equiv \tilde{f}_{ij} + \tilde{f}_{jk} + \tilde{f}_{ki}. \tag{80}$$

Indeed $\tilde{a}_{ijk}$ is a locally constant function, since
$d\tilde{a}_{ijk} = d\tilde{f}_{ij} + d\tilde{f}_{jk} + d\tilde{f}_{ki} = \beta_i - \beta_j + \beta_j - \beta_k + \beta_k - \beta_i = 0.$
Moreover, the constant $\tilde{a}_{ijk}$ define a 2-cochain: properties (CO1) and (CO2) follow directly from its definition. Property (CO3) follows from
$\tilde{a}_{[ijk]} \equiv \frac{1}{6}\{\tilde{a}_{ijk} - \tilde{a}_{ikj} + \tilde{a}_{jki} - \tilde{a}_{jik} + \tilde{a}_{kij} - \tilde{a}_{kji}\} =$
$\frac{1}{6}\{\tilde{f}_{ij} + \tilde{f}_{jk} + \tilde{f}_{ki} - \tilde{f}_{ik} - \tilde{f}_{kj} - \tilde{f}_{ji} + \tilde{f}_{jk} + \tilde{f}_{ki} + \tilde{f}_{ij}$
$- \tilde{f}_{ji} - \tilde{f}_{ik} - \tilde{f}_{kj} + \tilde{f}_{ki} + \tilde{f}_{ij} + \tilde{f}_{jk} - \tilde{f}_{kj} - \tilde{f}_{ji} - \tilde{f}_{ik}\} =$
$\tilde{f}_{ij} + \tilde{f}_{jk} + \tilde{f}_{ki} = \tilde{a}_{ijk},$
where the anti-symmetry of f_{ij}, f_{jk} and f_{ki} is used in the third identity. The constant $\tilde{a}_{ijk}$'s even define a cocycle $a \in Z^2(U, \mathbb{R})$, since, using (75),
$(\delta\tilde{a})_{ijkl} = \tilde{a}_{jkl} - \tilde{a}_{ikl} + \tilde{a}_{ijl} - \tilde{a}_{ijk} =$
$\tilde{f}_{jk} + \tilde{f}_{kl} + \tilde{f}_{lj} - \tilde{f}_{ik} - \tilde{f}_{kl} - \tilde{f}_{li} + \tilde{f}_{ij} + \tilde{f}_{jl} + \tilde{f}_{li} - \tilde{f}_{ij} - \tilde{f}_{jk} - \tilde{f}_{ki} = 0.$
In the last identity the anti-symmetry of the functions $\tilde{f}_{lj}$ and $\tilde{f}_{ik}$ is used. The equivalence class $[a] \in H^2(U, \mathbb{R})$ determined by $a \in Z^2(U, \mathbb{R})$ depends only on Ω and not on the choices made for the β_i, f_{ij} and a_{ijk}. To see this, choose a general $\hat{\beta}_i = \beta_i + dh_i$ for some exact 1-form dh_i. Then $\hat{\beta}_i$ and β_i give rise to the same closed real 2-form Ω, since
$d\hat{\beta}_i = d(\beta_i + dh_i) = d\beta_i = \Omega.$
For this new $\hat{\beta}_i$ we get the new
$d\hat{f}_{ij} = \hat{\beta}_i - \hat{\beta}_j = \beta_i + dh_i - \beta_j - dh_j = d\tilde{f}_{ij} + dh_i - dh_j$
$\Leftrightarrow d(\hat{f}_{ij} - \tilde{f}_{ij} - h_i + h_j) = 0$
$\Leftrightarrow \hat{f}_{ij} = \tilde{f}_{ij} + h_i - h_j + d_{ij},$
where $d_{ij} = -d_{ji}$ are constant functions on nonempty $U_i \cap U_j$. Hence the new $\hat{a}_{ijk}$ become
$\hat{a}_{ijk} = \hat{f}_{ij} + \hat{f}_{jk} + \hat{f}_{ki} =$
$\tilde{f}_{ij} + h_i - h_j + d_{ij} + \tilde{f}_{jk} + h_j - h_k + d_{jk} + \tilde{f}_{ki} + h_k - h_i + d_{ki} =$
$\tilde{a}_{ijk} + d_{ij} + d_{jk} + d_{ki}.$
Now the $e_{ijk} \equiv d_{ij} + d_{jk} + d_{ki}$ define a 2-coboundary $e \equiv \{e_{ijk}\}$. It is clear that the e_{ijk} form a 2-cochain, since the properties (CO1),(CO2) and (CO3) can be checked completely analogous as in the a_{ijk} case. The fact that they in particular form a 2-coboundary follows from
$(\delta d)_{ijk} = d_{ij} - d_{ik} + d_{jk} = d_{ij} + d_{jk} + d_{ki} = e_{ijk}.$
But this implies that the cocycles $\hat{a} \equiv \{\hat{a}_{ijk}\}$ and $\tilde{a} \equiv \{\tilde{a}_{ijk}\}$ only differ by the coboundary e. This in turn implies that $\hat{a}$ and $\tilde{a}$ define the same equivalence class $[a] \in H^2(U, \mathbb{R})$. As a consequence, the equivalence class $[a]$ only depends on the choice for Ω, not on the choices for the

β_i, f_{ij} and a_{ijk}.
Next, define $z_{ijk} \equiv \frac{1}{2\pi}(\tilde{a}_{ijk} + d_{ij} + d_{jk} + d_{ki})$ on nonempty $U_i \cap U_j \cap U_k$. Then it follows from the argument given above that $z \equiv \{z_{ijk}\}$ is an element of the class defined by $\frac{1}{2\pi}\Omega$ in $H^2(U, \mathbb{R})$. Let $m \in U_{ijk}$. By (IC), when $U_i \cap U_j$ is nonempty, we can define the constant d_{ij} such that all $z_{ijk}(m)$ are integers when $U_i \cap U_j \cap U_k$ is nonempty. Notice that we wouldn't be able to choose the d_{ij} such that the $z_{ijk}(m)$ are integers without (IC), since in general the $\tilde{a}_{ijk}$ are only locally constant and the intersections $U_i \cap U_j \cap U_k$ do not need to be connected.
We are now ready to define on each nonempty $U_i \cap U_j$,

$$c_{ij} \equiv \exp(-i(\tilde{f}_{ij} + d_{ij})) \tag{81}$$

Then, on each nonempty U_{ijk} it holds that
$c_{ij}(m)c_{jk}(m)c_{ki}(m) =$
$\exp(-i(\tilde{f}_{ij}(m) + \tilde{f}_{jk}(m) + \tilde{f}_{ki}(m) + d_{ij}(m) + d_{jk}(m) + d_{ki}(m))) =$
$\exp(-2\pi i z_{ijk}) = \cos(-2\pi z_{ijk}) + i\sin(-2\pi z_{ijk}) = 1$.
Furthermore, $(c_{ij})^{-1} = c_{ji}$, since
$c_{ji}(m)c_{ij}(m) = \exp(-i(\tilde{f}_{ji}(m) + d_{ji}(m) + \tilde{f}_{ij}(m) + d_{ij}(m))) =$
$\exp(0) = 1$.
In other words, the complex functions c_{ij} defined above satisfy the same properties (72) and (73) as transitions functions. We claim that the smooth complex functions c_{ij} defined on nonempty U_{ij} allow us to construct a line bundle X over M of which the c_{ij} are transition functions. This is the line bundle whose total space is $\coprod_i U_i \times \mathbb{C}/\sim$ with $\sim$ the equivalence relation defined by: $(j, m, z) \sim (k, \tilde{m}, \tilde{z})$ whenever $(j, m, z) \in I \times U_j \times \mathbb{C}$, $(k, \tilde{m}, \tilde{z}) \in I \times U_k \times \mathbb{C}$ and $m = \tilde{m}$ and $\tilde{z} = c_{jk}(m)z$, and with the projection map $\mathrm{pr} : \coprod_i U_i \times \mathbb{C}/\sim \to M \quad [(j, m, z)] \mapsto m$.
Here the points in $\coprod_i U_i \times \mathbb{C}$ are denoted as triples $(j, m, z) \in I \times U_j \times \mathbb{C}$ with I some index set. Notice that the projection map pr is well-defined since all points $(j, m, z) \sim (k, \tilde{m}, \tilde{z})$ are mapped to the same point $m = \tilde{m} \in M$ by pr. Furthermore, notice that $\sim$ does indeed define an equivalence relation, since

- $(j, m, z) \sim (j, m, z)$: follows from $m = m$ and $z = c_{jj}(m)z$. The last identity holds, because $c_{jj}(m) = \exp(-i(\tilde{f}_{jj}(m) + d_{jj}(m))) = \exp(0) = 1$, using the anti-symmetry of f_{jj} and d_{jj}.

– $(j,m,z) \sim (k,\tilde{m},\tilde{z}) \Rightarrow (k,\tilde{m},\tilde{z}) \sim (j,m,z)$: suppose $(j,m,z) \sim (k,\tilde{m},\tilde{z})$, then $m = \tilde{m}$ and $z = c_{jk}(m)\tilde{z}$. This implies $\tilde{m} = m$ and $\tilde{z} = c_{kj}(m)z$ using (72), i.e. $(k,\tilde{m},\tilde{z}) \sim (j,m,z)$.
– $(j,m,z) \sim (k,\tilde{m},\tilde{z})$, $(k,\tilde{m},\tilde{z}) \sim (l,\hat{m},\hat{z}) \Rightarrow (j,m,z) \sim (l,\hat{m},\hat{z})$: suppose $(j,m,z) \sim (k,\tilde{m},\tilde{z})$ and $(k,\tilde{m},\tilde{z}) \sim (l,\hat{m},\hat{z})$, that is, $m = \tilde{m} = \hat{m}$ and $z = c_{jk}\tilde{z} = c_{jk}c_{kl}\hat{z}$. This implies $m = \hat{m}$ and $z = c_{jl}\hat{z}$ using (73) and (72) respectively ($c_{jk}c_{kl}c_{lj} = 1 \Rightarrow c_{jk}c_{kl} = c_{jl}$). It thus follows that $(j,m,z) \sim (l,\hat{m},\hat{z})$.

We define addition and scalar multiplication on the fibers ($m \in M$ fixed) of $X \equiv \coprod_i U_i \times \mathbb{C}/\sim$ by $[(j,m,z_1)] + [(j,m,z_2)] = [(j,m,z_1+z_2)]$ and $\lambda[(j,m,z_1)] = [(j,m,\lambda z_1)]$ ($\lambda \in \mathbb{C}$) respectively. Let us now prove the claim:

(1) X satisfies (VB1) for $\mathbb{K} = \mathbb{C}$: First of all, $\mathrm{pr}^{-1}(m) = \{(j,m,z) \in I \times U_j \times \mathbb{C}\}/\sim$. Since m is fixed and $j \in I$ only takes those values such that $m \in U_j$ (let us say that there are a such values), we can view $\{(j,m,z) \in I \times U_j \times \mathbb{C}\}$ as the set consisting of a copies of $\mathbb{C}$. Then $\mathrm{pr}^{-1}(m)$ can be viewed as just one copy of $\mathbb{C}$, since in this case points (k,m,z) and $(l,m,\tilde{z})$ (m fixed) are 'identified' if $z = c_{kl}(m)\tilde{z}$ and the map $f : \mathbb{C} \to \mathbb{C}$ $\tilde{z} \mapsto c_{kl}(m)\tilde{z}$ is clearly an isomorphism (with respect to addition on $\mathbb{C}$). Our goal is indeed to show that $\mathrm{pr}^{-1}(m)$ is isomorphic to $\mathbb{C}$. To this end, define $\alpha_k : \mathbb{C} \to \mathrm{pr}^{-1}(m)$ $z \mapsto [(k,m,z)]$. That α_k is surjective follows directly from the discussion above. It is also injective. Indeed,
$\alpha_k(z) = \alpha_k(\tilde{z}) \Leftrightarrow [(k,m,z)] = [(k,m,\tilde{z})]$
$\Leftrightarrow (k,m,z) \sim (k,m,\tilde{z}) \Leftrightarrow z = c_{kk}(m)\tilde{z} = \tilde{z}$.
The last identity indeed holds, since $c_{kk}(m) = 1$. Finally, it is a homomorphism, because for $z_1, z_2 \in \mathbb{C}$ it holds that
$\alpha_k(z_1 + z_2) = [(k,m,z_1+z_2)] =$
$[(k,m,z_1)] + [(k,m,z_2)] = \alpha_k(z_1) + \alpha_k(z_2)$.

(2) X satisfies (VB2) for $\mathbb{K} = \mathbb{C}$: take $m \in M$. If we pick a neighbourhood U of m we can always choose it such that it is contained in the intersection of the U_i which contain m, since this intersection is an open set. With this in mind, $\mathrm{pr}^{-1}(U) = \{(j,m,z) \in I \times U_i \times \mathbb{C} | m \in U \subset U_i\}/\sim$. Now define the map $\Psi_k : U \times \mathbb{C} \to \mathrm{pr}^{-1}(U)$ $(m,z) \mapsto [(k,m,z)]$. Its surjectivity follows directly from surjectivity of α_k and the fact that Ψ_k sends m to $m = \tilde{m}$ (Ψ_k is the 'identity map in m'). Injectivity

follows from,
$\Psi_k(m,z) = \Psi_k(\tilde{m},\tilde{z}) \Leftrightarrow [(k,m,z)] = [(k,\tilde{m},\tilde{z})]$
$\Leftrightarrow (k,m,z) \sim (k,\tilde{m},\tilde{z}) \Leftrightarrow m = \tilde{m}, z = c_{kk}(m)\tilde{z} = \tilde{z}$.
It follows that $(m,z) = (\tilde{m},\tilde{z})$. Furthermore Ψ_k is smooth: it sends (m,z) to $[(k,m,z)]$ where $(k,m,z) \sim (k,\tilde{m},\tilde{z})$ if $m = \tilde{m}$ and $z = c_{kk}(m)\tilde{z} = \tilde{z}$, so it sends m to $m = \tilde{m}$ in a smooth fashion ('identity map in m') and it sends z to $z = \tilde{z}$ in a smooth fashion ('identity map in z). It follows that Ψ_k is smooth. The smoothness of Ψ_k on $U_k \cap U_j$ should not depend on the choice k. Indeed it does not, since $(k,m,z) \sim (j,\tilde{m},\tilde{z})$ if $m = \tilde{m}$, $z = c_{kj}(m)\tilde{z}$ and c_{kj} is smooth. For the inverse map, define
$(\Psi_k)^{-1} : \mathrm{pr}^{-1}(U) \to U \times \mathbb{C}$ $[(k,m,z)] \mapsto (m,z)$. Indeed it is an inverse, since $(\Psi_k^{-1} \circ \Psi_k)(m,z) = \Psi_k^{-1}([(k,m,z)]) = (m,z)$. That the inverse map is well-defined follows from $c_{kk}(m) = 1$. Its smoothness is proved in the same way the smoothness of Ψ_k is proved.

The fact that the restriction map,
$\Psi_k|_m : \{m\} \times \mathbb{C} \to \mathrm{pr}^{-1}(m)$ $(m,z) \mapsto [(k,m,z)]$ is an isomorphism follows directly from $\{m\} \times \mathbb{C} \cong \mathbb{C}$ and the fact that α_k is an isomorphism. Linearity follows from
$\Psi_k|_m(\lambda z_1 + z_2) = [(k,m,\lambda z_1 + z_2)] = [(k,m,\lambda z_1)] + [(k,m,z_2)] =$
$\lambda[(k,m,z_1)] + [(k,m,z_2)] = \lambda\Psi_k|_m(z_1) + \Psi_k|_m(z_2)$.

(3) The c_{ij} are the transition functions of the line bundle X: in order to verify this, we have to check if the c_{ij}s satisfy (70). Indeed, on $(U_i \cap U_j) \times \mathbb{C}$, we have
$c_{jk}(m)\Psi_j(m,z) = c_{jk}(m)[(j,m,z)] = [(j,m,c_{jk}(m)z)] = [(k,m,z)] =$
$\Psi_k(m,z)$.

What is left to check is that this line bundle X is equipped with a Hermitian structure and a connection which are compatible. For the Hermitian structure, define, for $m \in U_i$ the inner product in each fibre by
$\langle \cdot, \cdot \rangle : \mathrm{pr}^{-1}(m) \cong \mathbb{C} \times \mathrm{pr}^{-1}(m) \cong \mathbb{C} \to \mathbb{R}$
$([(i,m,z_1)], [(i,m,z_2)]) \mapsto \langle [(i,m,z_1)], [(i,m,z_2)] \rangle = \bar{z}_1 z_2$.
In order to show that this inner product is well-defined for $m \in U_i \cap U_j$, note that the $\tilde{f}_{ij}, d_{ij} \in \mathbb{R}$. It follows that
$|c_{ij}|^2 = c_{ij}\bar{c_{ij}} = \exp(-i(\tilde{f}_{ij} + d_{ij}))\exp(i(\tilde{f}_{ij} + d_{ij})) = \exp(0) = 1$.
Hence, on $U_i \cap U_j$ we have
$\langle [(i,m,z_1)], [(i,m,z_2)] \rangle = \langle [(i,m,c_{ij}z_1)], [(i,m,c_{ij}z_2)] \rangle =$
$\bar{c_{ij}z_1}c_{ij}z_2 = \bar{c_{ij}}\bar{z}_1 c_{ij} z_2 = |c_{ij}|^2 \bar{z}_1 z_2 = \bar{z}_1 z_2$,

as should be the case. Furthermore, for $m \in U_i$ the map $H : X \to \mathbb{R}\ (m,z) \mapsto \langle [(i,m,z)], [(i,m,z)] \rangle = \bar{z}z$ is smooth and on $U_i \cap U_j$ (where $(i,m,z) \sim (j,\tilde{m},\tilde{z})$ if $m=\tilde{m}$ and $z = c_{ij}\tilde{z}$) the smoothness does not depend on the choice i, since c_{ij} is smooth. Hence $\langle \cdot,\cdot \rangle$ defines a Hermitian structure on X.

The line bundle X is also equipped with a connection. From the real β_i we can construct complex $\tilde{\beta}_i \equiv i\beta_i$. Pick a collection $\{(U_i,\Psi_i)\}$ of local trivializations of X, a unit section $\hat{1}_i$ on U_i and a vector field $X \in \mathrm{Vect}(U_i)$. From the complex $\tilde{\beta}_i$ we can define a connection ∇_X via

$$\tilde{\beta}_i \hat{1}_i = i\nabla_X \hat{1}_i \tag{82}$$

Since a general section on U_i can be written as $s_i = f\hat{1}_i$ for $f \in C^\infty(U_i,\mathbb{C})$ a connection of a general section on U_i can be written as

$$\nabla_X s_i = X(f)\hat{1}_i - i\beta_i(X)s_i, \tag{83}$$

as derived previously on a general neighbourhood U. From (71) it follows that on U_{ij} in general sections s_i,s_j are related by

$$s_i(m) = c_{ji}(m)s_j(m). \tag{84}$$

Since $\nabla_X s_i$ should be a section, in order for ∇_X to be a well-defined connection the relation $(\nabla_X s_j)(m) = c_{ij}(m)(\nabla_X s_i)(m)$ should be satisfied on U_{ij}. Indeed, on U_{ij},
$(\nabla_X s_j)(m) = (\nabla_X c_{ij}s_i)(m) = (\nabla_X c_{ij}f\hat{1}_i)(m) =$
$(X(c_{ij}f)\hat{1}_i)(m) - i(\beta_j(X)c_{ij}s_i)(m) =$
$(X(c_{ij})s_i)(m) + (c_{ij}X(f)\hat{1}_i)(m) - i(\beta_j(X)c_{ij}s_i)(m) =$
$-i(c_{ij}(\beta_i(X) - \beta_j(X))s_i)(m) + (c_{ij}X(f)\hat{1}_i)(m) - i(\beta_j(X)c_{ij}s_i)(m) =$
$(c_{ij}(X(f)\hat{1}_i - ic_{ij}\beta_i(X)s_i))(m) = c_{ij}(m)(\nabla_X s_i)(m).$
In the fifth identity we used
$X(c_{ij}) = dc_{ij}(X) = d(\exp(-i(\tilde{f}_{ij} + d_{ij})))(X) =$
$(-id(\tilde{f}_{ij} + d_{ij})\exp(-i(\tilde{f}_{ij} + d_{ij})))(X) = (-id\tilde{f}_{ij}c_{ij})(X) =$
$(-i(\beta_i - \beta_j)c_{ij})(X).$
Finally, for X to be a Hermitian line bundle with connection we have to check that the constructed connection and Hermitian structure are compatible. Let $X \in \mathrm{Vect}(U_i)$, $s_i, s'_i \in \Gamma_X(U_i)$. Then we know that $s_i = f\hat{1}_i$ for $f \in C^\infty(U_i,\mathbb{C})$ and $s'_i = g\hat{1}_i$ for another $g \in C^\infty(U_i,\mathbb{C})$. Hence, on the one hand,
$X\langle s_i, s'_i \rangle \equiv X\langle f\hat{1}_i, g\hat{1}_i \rangle \equiv X(\bar{f}g)\langle \hat{1}_i, \hat{1}_i \rangle =$
$X(\bar{f})g\langle \hat{1}_i, \hat{1}_i \rangle + \bar{f}X(g)\langle \hat{1}_i, \hat{1}_i \rangle.$

On the other hand,
$\langle \nabla_{\bar{X}} s_i, s_i' \rangle + \langle s_i, \nabla_X s_i' \rangle =$
$\langle \bar{X}(f)\hat{1}_i - i\beta_i(\bar{X})s_i, s_i' \rangle + \langle s, X(g)\hat{1}_i - i\beta_i(X)s_i' \rangle =$
$X(\bar{f})\langle \hat{1}_i, s_i' \rangle + i\beta_i(X)\langle s_i, s_i' \rangle + X(g)\langle s_i, \hat{1}_i \rangle - i\beta_i(X)\langle s_i, s_i' \rangle =$
$X(\bar{f})\langle \hat{1}_i, s_i' \rangle + X(g)\langle s_i, \hat{1}_i \rangle = X(\bar{f})g\langle \hat{1}_i, \hat{1}_i \rangle + X(g)\bar{f}\langle \hat{1}_i, \hat{1}_i \rangle.$
This establishes the compatability condition (64) on some U_i. The fact that the compatibility condition continues to hold on the overlaps $U_i \cap U_j$ in a well-defined manner follows from the well-definedness of the inner product on the overlaps (since $|c_{ij}|^2 = 1$) and from $\nabla_X s_j = \nabla_X(c_{ij}s_i) = c_{ij}\nabla_X s_i$, so we have proved the 'sufficient'-part of the theorem (since the open cover of M in (IC) can always be chosen such that it is contractible for a collection $\{U_i, \Psi_i\}$ of local trivializations of X).

Remark 6.5.1. Note that we started with a symplectic 2-form Ω on M. As we constructed the Hermitian line bundle-with-connection we in particular constructed the connection ∇ defined by (83) on general sections $s_i = f\hat{1}_i$ for $f \in C^\infty(U_i, \mathbb{C})$. It now follows from the argument presented in section 6.4.1 that the curvature 2-form of ∇ on M equals this symplectic form on M, once M is the base manifold of a line bundle-with-connection.

- The Integrality Condition is a necessary condition for a Hermitian line bundle-with-connection to exist:

  Suppose we are given a line bundle X over M with connection ∇. Pick a selection of local trivializations $\{U_i, \Psi_i\}$ of B such that $U = \{U_i\}$ is a contractible open cover of M. Pick the diffeomorphisms $\Psi_i : U_i \times \mathbb{C} \to \mathrm{pr}^{-1}(U_i)$ such that the restriction maps $\Psi_i|_m : \{m\} \times \mathbb{C} \cong \mathbb{C} \to \mathrm{pr}^{-1}(m)$ are compatible with the inner product, meaning that for $z_1, z_2 \in \mathbb{C}$ we have $\langle \Psi_i|_m(z_1), \Psi_i|_m(z_2) \rangle \equiv \langle z_1, z_2 \rangle$. We can always choose the $\Psi_i|_m$ this way. To see this, take $z_1 \in \mathbb{C}$ and suppose that $\langle \Psi_i|_m(z_1), \Psi_i|_m(z_1) \rangle \neq \langle z_1, z_1 \rangle$. Then, since both $\langle \Psi_i|_m(z_1), \Psi_i|_m(z_1) \rangle$ and $\langle z_1, z_1 \rangle$ belong to $\mathbb{R}$ they should be related by some real positive function $\lambda : U_i \to \mathbb{R}_{>0}$ where $\mathbb{R}_{>0}$ is the set of positive real numbers. So we can choose a rescaled $\tilde{\Psi}_i|_m(z_1) \equiv \left(\sqrt{\lambda(m)}\right)^{-1} \Psi_i|_m(z_1)$ which then gives

$$\langle \tilde{\Psi}_i|_m(z_1), \tilde{\Psi}_i|_m(z_1) \rangle = \left(\sqrt{\lambda(m)}\right)^{-1} \left(\sqrt{\lambda(m)}\right)^{-1} \langle \Psi_i|_m(z_1), \Psi_i|_m(z_1) \rangle =$$

$$\left(\sqrt{\lambda(m)}\right)^{-1} \left(\sqrt{\lambda(m)}\right)^{-1} \lambda(m)\langle z_1, z_1 \rangle = \langle z_1, z_1 \rangle.$$

Since $\tilde{\Psi}_i|_m$ is linear it follows that this map does define a linear isomorphism compatible with the inner product. Hence we can always construct such a map, because the $\Psi_i|_m$ we started with was arbitrary. Since we have a well known inner product on $\mathbb{C}$, namely $\langle z_1, z_2\rangle \equiv \bar{z_1}z_2$ we now have an inner product $\langle \Psi_i|_m(z_1), \Psi_i|_m(z_2)\rangle = \langle z_1, z_2\rangle \equiv \bar{z_1}z_2$ on $\mathrm{pr}^{-1}(m)$ for $m \in U_i$ too. Subsequently, consider the functions
$\tilde{c_{ji}} = (\psi_i^{-1} \circ \psi_j)|_{U_{ij}\times\mathbb{C}} : U_{ij} \times \mathbb{C} \xrightarrow{\psi_j|_{U_{ij}\times\mathbb{C}}} \mathrm{pr}^{-1}(U_{ij}) \xrightarrow{\psi_i^{-1}|\mathrm{pr}^{-1}(U_{ij})} U_{ij} \times \mathbb{C}$
introduced in subsection 6.4.1 for $m \in U_{ij}$. Since $\Psi_j|_m$ is compatible with the inner product, so is $\Psi_i^{-1}|_{m\times\mathbb{C}}$ and so is
$\tilde{c_{ij}}|_{m\times\mathbb{C}} \equiv \Psi_j|_m \circ \Psi_i^{-1}|_{m\times\mathbb{C}}$. Using (74) this implies for the norm of the corresponding transition functions c_{ij} that
$\langle \tilde{c_{ij}}|_{m\times\mathbb{C}}(z), \tilde{c_{ij}}|_{m\times\mathbb{C}}(z)\rangle = \langle z, z\rangle \Leftrightarrow \langle c_{ij}(m)z, c_{ij}(m)z\rangle = \langle z, z\rangle$
$\Leftrightarrow |c_{ij}(m)|^2\langle z, z\rangle \Leftrightarrow |c_{ij}(m)| = 1.$
From this it follows that the inner product $\langle\cdot,\cdot\rangle$ on $\mathrm{pr}^{-1}(m)$ is well-defined for $m \in U_i \cap U_j$ as well. Indeed, take $z_1, z_2 \in \mathrm{pr}^{-1}(m) \cong \mathbb{C}$. We know that the sections $s_i \in \Gamma_X(U_i)$, $s_j \in \Gamma_X(U_j)$ are related by $s_j = c_{ij}s_i$ on U_{ij}. Hence, writing $z_1 \equiv \tilde{s}_i(m)$, $z_2 \equiv \hat{s}_i(m)$, on U_{ij} we have
$\langle \tilde{s}_j, \hat{s}_j\rangle(m) = \langle c_{ij}\tilde{s}_i, c_{ij}\hat{s}_i\rangle(m) = |c_{ij}(m)|^2\langle \tilde{s}_i, \hat{s}_i\rangle(m) = \langle z_1, z_2\rangle,$
as should be the case. The line bundle X is now provided globally with a Hermitian structure, since for $m \in U_i$ the map
$H : X \to \mathbb{R}\ \ (m, z) \mapsto \langle z, z\rangle = \bar{z}z$ is smooth and this smoothness does not depend on the choice i, since c_{ij} is smooth.

The existence of the connection ∇ allows us to define a complex 1-form β_i on U_i via $\hat{1}_i\beta_i(X) = \nabla_X(\hat{1}_i)$ for $X \in \mathrm{Vect}(U_i)$, and for general sections $s_i = f\hat{1}_i$ on U_i ($f \in C^\infty(U_i, \mathbb{C})$) via the relation $\nabla_X s_i = X(f)\hat{1}_i - i\beta_i(X)s_i$. The compatibility of the connection ∇ and the Hermitian structure $\langle\cdot,\cdot\rangle$ follows now by the same arguments as in the 'sufficient'-part, since $|c_{ij}(m)|^2 = 1$ for $m \in U_{ij}$ in this case too. Hence we have a Hermitian line bundle-with-connection.

Next, define $f_{ij} \equiv \frac{-1}{i}\log(c_{ij})$ on U_{ij} where log is the complex logarithm whose values are only determined upto integer multiples $n \in \mathbb{Z}$ of $2\pi i$. And from this define, $z_{ijk} \equiv f_{ij} + f_{jk} + f_{ki}$ on U_{ijk}. In subsubsection 6.4.1 we proved that the transition functions c_{ij} satisfy
$c_{ij}(m)c_{jk}(m)c_{ki}(m) = 1$ on U_{ijk}. It follows that
$z_{ijk}(m) \equiv f_{ij}(m) + f_{jk}(m) + f_{ki}(m) =$
$\frac{-1}{i}\log(c_{ij}(m)) + \frac{-1}{i}\log(c_{jk}(m)) + \frac{-1}{i}\log(c_{jk}(m)) =$
$\frac{-1}{i}\log(c_{ij}(m)c_{jk}(m)c_{ki}(m)) = \frac{-\log(1)}{i} + \frac{n_{ijk}2\pi i}{i} = 2\pi n_{ijk} \in 2\pi\mathbb{Z}$
for $n_{ijk} : U_{ijk} \to \mathbb{Z}$ a constant integer. Hence the z_{ijk}s are constants.

The fact that these z_{ijk} form a cochain follows from the same arguments as for the a_{ijk} in the 'sufficient'-part. In particular, the z_{ijk} determine a cocycle $z \equiv \{z_{ijk}\}$. Indeed,
$(\delta\tilde{z})_{ijkl} = \tilde{z}_{jkl} - \tilde{z}_{ikl} + \tilde{z}_{ijl} - \tilde{z}_{ijk} =$
$\tilde{f}_{jk} + \tilde{f}_{kl} + \tilde{f}_{lj} - \tilde{f}_{ik} - \tilde{f}_{kl} - \tilde{f}_{li} + \tilde{f}_{ij} + \tilde{f}_{jl} + \tilde{f}_{li} - \tilde{f}_{ij} - \tilde{f}_{jk} - \tilde{f}_{ki} = 0.$
Here is used the anti-symmetry of the f_{ij}. Indeed, since we proved in subsubsection 6.4.1 that the transition functions c_{ij} satisfy
$c_{ij}(m) = (c_{ji}(m))^{-1}$, it follows that
$f_{ij} = \frac{-1}{i}\log(c_{ij}) = \frac{-1}{i}\log((c_{ji})^{-1}) = \frac{1}{i}\log(c_{ji}) = -f_{ji}$.

Our independently constructed β_i and f_{ij} satisfy the relation
$df_{ij} = \beta_i - \beta_j$. Indeed, since sections transform as $s_j = c_{ij}s_i$ on $U_i \cap U_j$ we know that $\nabla_X s_j$ for $X \in \mathrm{Vect}(U_j)$ satisfies $\nabla_X s_j = c_{ij}\nabla_X s_i$. It follows, writing $s_i = f\hat{1}_i$ for $f \in C^\infty(U_i, \mathbb{C})$, that
$\nabla_X(c_{ij}s_i) = c_{ij}\nabla_X s_i \Leftrightarrow$
$X(c_{ij}f)\hat{1}_i - i\beta_j(X)c_{ij}s_i = c_{ij}(X(f)\hat{1}_i - i\beta_i(X)s_i) \Leftrightarrow$
$X(c_{ij})f\hat{1}_i + c_{ij}X(f)\hat{1}_i - i\beta_j(X)c_{ij}s_i = c_{ij}(X(f)\hat{1}_i - i\beta_i(X)s_i) \Leftrightarrow$
$X(c_{ij})s_i - ic_{ij}s_i(\beta_j(X) - \beta_i(X)) = 0 \Leftrightarrow X(c_{ij}) = ic_{ij}(\beta_j - \beta_i)(X) \Leftrightarrow$
$dc_{ij}(X) = ic_{ij}(\beta_j - \beta_i)(X) \Leftrightarrow \frac{dc_{ij}}{c_{ij}} = i(\beta_j - \beta_i) \Leftrightarrow$
$\frac{-1}{i}d\log(c_{ij}) = \beta_i - \beta_j \Leftrightarrow df_{ij} = \beta_i - \beta_j$.
Subsequently, define $\Omega = d\beta_i$ on U_i. Since for general sections
$s_i = f\hat{1}_i$ ($f \in C^\infty(U_i, \mathbb{C})$) the β_i are defined via (83) it follows by the argument given in subsubsection 6.4.1 that Ω equals the curvature 2-form corresponding to X. Hence, by remark 6.5.1 it equals the symplectic 2-form on M (since M is the base manifold of the (Hermitian) line bundle-with-connection X).

Let $[z] \in H^2(U, 2\pi\mathbb{Z})$ be the equivalence class determined by the cocycle $z \in Z^2(U, 2\pi\mathbb{Z})$. This equivalence class depends only on Ω and not on the choices made for the β_i, f_{ij} (in particular not on the ambiguity in defining them) and z_{ijk}s. The argument is fully equivalent to the one presented in the 'sufficient'-part, except that in this case the fact that the values of $f_{ij}(m)$ ($m \in U_{ij}$) are only determined up to integer multiples of $2\pi i$ has to be taken into account. This ambiguity is absent for the equivalence classes $[z] \in H^2(U, 2\pi\mathbb{Z})$. Indeed, take
$\tilde{f}_{ij} = f_{ij} - 2\pi i n_{ij}$ on U_{ij} for $n_{ij} : U_i \cap U_j \to \mathbb{Z}$ an integer. These new $\tilde{f}_{ij}$ give rise to the same 2-form Ω, since the n_{ij} are constant functions. Furthermore,
$\tilde{z}_{ijk} \equiv \tilde{f}_{ij} + \tilde{f}_{jk} + \tilde{f}_{ki} = f_{ij} - 2\pi i n_{ij} + f_{jk} - 2\pi i n_{jk} + f_{ki} - 2\pi i n_{ki} =$
$z_{ijk} - 2\pi i(n_{ij} + n_{jk} + n_{ki}) = z_{ijk} - 2\pi(\delta n)_{ijk}$
on U_{ijk}. So the z differ at most by a coboundary and hence give rise to the same equivalence class $[z]$.

As a consequence we have found an open cover U of M such that the class $[z]$ defined by Ω in $H^2(U, \mathbb{R})$ contains a cocycle z in which all the z_{ijk} are 2π multiples of integers. The integrality condition is clearly an equivalent statement and the proof of the 'necessary'-part is complete.

$\square$

6.6 The prequantum Hilbert space

In subsection 6.3 we already mentioned that a prequantum Hilbert space consisting of 'ordinary' (square-integrable) complex-valued scalar functions $f : M \to \mathbb{C}$ is not the correct Hilbert space to consider. Now that we have treated the right objects to consider, sections of a Hermitian line bundle-with-connection, and formulated the condition for the classical phase space M to be quantizable we are ready to define the correct prequantum Hilbert space and the prequantum operators acting on it. This will be the purpose of this subsection.

Suppose that the symplectic manifold (M, ω) is quantizable in the sense that $\frac{1}{\hbar}\omega$ satisfies the integrality condition (IC). Then we know that it is possible to construct a Hermitian line bundle-with-connection X with curvature 2-form $\frac{1}{\hbar}\omega$, called the **prequantum line bundle** Prequantum line bundle. Now let $m \in M$ and $s_1, s_2 \in \Gamma_X(M)$. Define

$$(s_1, s_2)(m) \equiv (\frac{1}{2\pi\hbar})^n \int_M \langle s_1, s_2 \rangle(m)\epsilon_\omega. \tag{85}$$

Here ϵ_ω denotes the Liouville form as defined in (59) and $\langle \cdot, \cdot \rangle$ denotes the Hermitian structure on X as defined in the previous subsection by $\langle s_1, s_2 \rangle(m) \equiv \langle s_1(m), s_2(m) \rangle \equiv \langle z_1, z_2 \rangle \equiv \bar{z}_1 z_2$ using the notation $z_1 \equiv s_1(m) \in \mathbb{C}$, $z_2 \equiv s_1(m) \in \mathbb{C}$. We claim that $(\cdot, \cdot)$ defines an inner product. Note that the space $\Gamma_X(M)$ is a vector space with complex-valued functions as scalars. Let us now verify the claim,

- $(\cdot, \cdot)$ is conjugate symmetric: first of all, $\langle \cdot, \cdot \rangle$ is conjugate symmetric, since
  $\langle s_1, s_2 \rangle(m) \equiv \bar{z}_1 z_2 = \overline{z_1 \bar{z}_2} = \overline{\langle s_1, s_2 \rangle}(m)$.
  Conjugate symmetry of $(\cdot, \cdot)$ now follows from
  $(s_1, s_2)(m) \equiv (\frac{1}{2\pi\hbar})^n \int_M \langle s_1, s_2 \rangle(m)\epsilon_\omega = (\frac{1}{2\pi\hbar})^n \int_M \overline{\langle s_2, s_1 \rangle}(m)\epsilon_\omega =$
  $\overline{(\frac{1}{2\pi\hbar})^n \int_M \langle s_1, s_2 \rangle(m)\epsilon_\omega}$,
  where in the last identity $\bar{\epsilon_\omega} = \epsilon_\omega$ and the fact that integration is a linear operation are used.

- $(\cdot,\cdot)$ is antilinear in the first argument: Let f be a complex-valued function on M. To begin with, $\langle\cdot,\cdot\rangle$ is antilinear in its first argument, since
$\langle fs_1, s_2\rangle(m) = \langle f(m)s_1(m), s_2(m)\rangle = \bar{f(m)z_1}z_2 =$
$\bar{f(m)}\bar{z_1}z_2 = \bar{f(m)}\langle s_1, s_2\rangle(m)$
and for $s_3 \in \Gamma_X(M)$ we have
$\langle s_1 + s_2, s_3\rangle(m) = \bar{z_1 + z_2}z_3 = \bar{z_1}z_3 + \bar{z_2}z_3 = \langle s_1, s_3\rangle(m) + \langle s_2, s_3\rangle(m)$,
denoting $s_3(m) \equiv z_3$.
Antilinearity of $(\cdot,\cdot)$ in its first argument now follows directly from integration being a linear operation.

- $(\cdot,\cdot)$ is positive definite: denote $z_1 \equiv a + ib$ for $a, b \in \mathbb{R}$. Firstly, $\langle\cdot,\cdot\rangle$ is positive definite, because
$\langle s_1, s_1\rangle(m) = \bar{z_1}z_1 = (a+ib)(a-ib) = a^2 + b^2 \geq 0$,
and
$\langle s_1, s_1\rangle(m) = 0 \Leftrightarrow a^2 + b^2 = 0 \Leftrightarrow a = 0 \text{ and } b = 0 \Leftrightarrow$
$z_1 = 0 \Leftrightarrow s_1(m) = 0$.
Positive-definiteness of $(\cdot,\cdot)$ now follows directly from the properties of integration and ϵ_ω and the positive value of the prefactor $\frac{1}{2\pi\hbar}$.

We are now ready to define the (correct) prequantum Hilbert space.

Definition 6.6.1. The **prequantum Hilbert space** H is the space of all $s \in \Gamma_X(M)$ equipped with the inner product $(\cdot,\cdot)$ defined in (85) for which the integral of $(s,s)\epsilon_\omega$ over M exists and is finite.

Furthermore, each classical observable $f \in S \subset C^\infty(M,\mathbb{R})$ acts on an appropriate subset of H according to

$$s \mapsto -i\hbar\nabla_{X_f}s + fs, \tag{86}$$

where X_f denotes the Hamiltonian vector field corresponding to the classical observable $f \in S \subset C^\infty(M,\mathbb{R})$. The corresponding quantization assignment Q defined in (54) is

$$Q(f) = -i\hbar\nabla_{X_f} + f. \tag{87}$$

Our goal is now to show that this assigment satisfies the conditions (Q1)-(Q4) introduced in subsection 6.2 and hence is a suitable prequantum assignment of operators on the prequantum Hilbert space (actually, some subset of it, which is argued below) to classical observables f. Indeed Q satisfies these conditions,

(1) Q satisfies (Q1): let $f, g \in S \subset C^\infty(M, \mathbb{R})$ and $\lambda \in \mathbb{R}$. Then,
$Q(\lambda f + g) = -i\hbar\nabla_{X_{\lambda f+g}} + (\lambda f + g) = -i\hbar\nabla_{\lambda X_f + X_g} + \lambda f + g =$
$-\, i\hbar\lambda\nabla_{X_f} - i\hbar\nabla_{X_g} + \lambda f + g = \lambda Q(f) + Q(g),$
which establishes the claim.

(2) Q satisfies (Q2): let $f, g \in S \subset C^\infty(M, \mathbb{R})$ and let $\frac{1}{\hbar}\omega$ be the symplectic form on M (as mentioned before). Then
$[Q(f), Q(g)] = [-i\hbar\nabla_{X_f} + f, -i\hbar\nabla_{X_g} + g] =$
$-\, i\hbar[\nabla_{X_f}, g] - i\hbar[f, \nabla_{X_g}] - \hbar^2[\nabla_{X_f}, \nabla_{X_g}] =$
$-\, i\hbar[\nabla_{X_f}, g] - i\hbar[f, \nabla_{X_g}] - \hbar^2\left(\nabla_{[X_f, X_g]} + \frac{2\omega(X_f, X_g)}{i\hbar}\right) =$
$-\, i\hbar[\nabla_{X_f, g} - i\hbar[f, \nabla_{X_g}] - \hbar^2\left(\nabla_{X_{\{f,g\}}} + \frac{\{f,g\}}{i\hbar}\right) =$
$-\, i\hbar[\nabla_{X_f}, g] - i\hbar[f, \nabla_{X_g}] - \hbar^2\nabla_{X_{\{f,g\}}} + i\hbar\{f, g\}.$
In the second identity the commutativity of scalar functions under pointwise multiplication is used. In the third identity we used the fact that the symplectic form $\frac{1}{\hbar}\omega$ (which equals the curvature 2-form as a base manifold of X) satisfies by definition the equality
$\frac{\omega(X_f, X_g)}{\hbar} \equiv \frac{i}{2}\left([\nabla_{X_f}, \nabla_{X_g}] - \nabla_{[X_f, X_g]}\right)$. In the fourth identity the definition of the Poisson bracket (40) and the fundamental identity (41) are used. Furthermore,taking $s \in H$ we have
$[\nabla_{X_f}, g]s = \nabla_{X_f}(gs) - g\nabla_{X_f}s = g\nabla_{X_f}s + X_f(g)s - g\nabla_{X_f}s =$
$X_f(g)s = \{f, g\}s,$
and hence $[\nabla_{X_f}, g] = \{f, g\}$. And
$[f, \nabla_{X_g}] = -[\nabla_{X_g}, f] = -\{g, f\} = \{f, g\}.$
It thus follows that,
$[Q(f), Q(g)] = -2i\hbar\{f, g\} + i\hbar\{f, g\} - \hbar^2\nabla_{X_{\{f,g\}}} =$
$-\, i\hbar(\{f, g\} - i\hbar\nabla_{X_{\{f,g\}}} = -i\hbar Q(\{f, g\}),$
as was to be shown.

(3) Q satisfies (Q3): let $s_1, s_2 \in H$ and $f \in S \subset C^\infty(M, \mathbb{R})$. Then
$(Q(f)s_1, s_2) = (-i\hbar\nabla_{X_f}s_1 + fs_1, s_2) =$
$i\hbar(\nabla_{X_f}s_1, s_2) + f(s_1, s_2) = i\hbar(\frac{1}{2\pi\hbar})^n \int_M \langle\nabla_{X_f}s_1, s_2\rangle\epsilon_\omega + f(s_1, s_2) =$
$-\, i\hbar(\frac{1}{2\pi\hbar})^n \int_M \langle s_1, \nabla_{X_f}s_2\rangle\epsilon_\omega + i\hbar(\frac{1}{2\pi\hbar})^n \int_M (X_f(s_1, s_2))\epsilon_\omega + f(s_1, s_2) =$
$-\, i\hbar(\frac{1}{2\pi\hbar})^n \int_M \langle s_1, \nabla_{X_f}s_2\rangle\epsilon_\omega + (s_1, fs_2) = (s_1, Q(f)s_2).$
In the fourth identity the compatibility of the Hermitian structure $\langle\cdot,\cdot\rangle$ on X and the connection ∇ on X is used. In the fifth identity we used that on the appropriate subset of the prequantum Hilbert space we are considering the term $(\frac{1}{2\pi\hbar})^n \int_M (X_f(s_1, s_2))\epsilon_\omega$ vanishes. For the technical proof of this fact we refer the reader to ([7, p.129]).

(4) Q satisfies (Q4): let 1 denote the map $1 : M \to \mathbb{R}$ $m \mapsto 1$ and 0 the zero vector field. Then

$Q(1) = -i\hbar\nabla_{X_1} + 1 = -i\hbar\nabla_0 + 1 = 1,$
where linearity of ∇ is used in the sense that $\nabla_0 = \tilde{0}$ with $\tilde{0}$ the zero operator on $\Gamma_X(M)$ and so we are done.

6.7 Polarizations

As mentioned in the introduction of this section the prequantum Hilbert space is considered 'too large', in the sense that it cannot represent the phase space of a physically reasonable quantum system. One way to reduce it is by putting an additional geometric structure on the classical phase space (as a base manifold of the prequantum line bundle), called a polarization. As metioned by Woodhouse (see [7, p.133]), it should be stressed that the justification is not based on general mathematical results (such as the irreducibility of representations) but on the examination of particular examples: the construction unifies and generalizes a number of familiar techniques that, in the past, have not appeared to have had any obvious connection with each other and that have sometimes seemed to be overspecialized with applications only to particular physical systems.

In this subsection we will first treat polarizations in general and then look at a specific kind of polarization, the so-called Kähler polarization.

6.7.1 Polarizations in general

The prequantum Hilbert space consists of sections s of a prequantum line bundle X which depend on all the $2n$ coordinates of the symplectic manifold (M, ω) on which they are defined. We will abstract the idea of a polarized section by first considering complex scalar functions depending on $2n$ coordinates. In subsubsection 6.1.1 we already proved that the configuration space Q, depending on the n position coordinates $\{q^1, ..., q^n\}$, is a Lagrangian submanifold of the cotangent bundle T^*Q, depending on the the $2n$ coordinates $\{q^1, ..., q^n, p_1, ..., p_n\}$, so complex scalar functions on T^*Q depend on $2n$ coordinates, whereas complex scalar functions on Q depend only on n coordinates. More generally, if we consider complex scalar functions f on a general symplectic manifold M, a way of eliminating half of the $2n$ coordinates they depend on is to demand that the functions f are constant along n independent vector fields on M. That is, for $k \in \{1, ..., n\}$, $X_k \in \mathrm{Vect}(M, \mathbb{R})$, we have

$$X_k(f) = 0. \tag{88}$$

Indeed, this gives n restrictions on M, reducing its dimension by n. Now define,

Definition 6.7.1.1. Let M be a manifold. For all $m \in M$ consider the subspaces $P_m \subset T_m M$. If m has an open neighborhood $U \subset M$ such that for all $u \in U$ a set of r independent vector fields $\{X_j\}$ exists with $\mathrm{Span}\{X_j\} = P_u$, then

$$P \equiv \coprod_{m \in M} P_m \subset TM \tag{89}$$

is called a **distribution** of dimension r on M.

If we denote the space of vector fields $X \in \mathrm{Vect}(M, \mathbb{R})$ such that $X|_{\{m\}} \in P_m$ for every $m \in M$ by $\mathrm{Vect}(M, \mathbb{R}; P)$ it follows that the n vector fields X_k considered above definition 6.7.1.1 are the base elements of the (in this case) n-dimensional space of vector fields $\mathrm{Vect}(M, \mathbb{R}; P)$.

Let us generalize this story to sections s on M. Since the directional derivative does not map sections to sections it has no invariant meaning for sections of X. However, the connection ∇ does map sections to sections. Consider the n-dimensional distribution P of the tangent bundle TM of M. A natural generalization of (88) to sections is

$$\nabla_X s = 0, \tag{90}$$

for all $X \in \mathrm{Vect}(M, \mathbb{R}; P)$. Let $X, Y \in \mathrm{Vect}(M, \mathbb{R}; P)$ be arbitrary and $s \in \Gamma_X(M)$. Then (90) implies

$$[\nabla_X, \nabla_Y] s = \nabla_X(\nabla_Y s) - \nabla_Y(\nabla_X s) = 0, \tag{91}$$

for all $X, Y \in \mathrm{Vect}(M, \mathbb{R}; P)$. Since the curvature 2-form $\frac{\omega}{\hbar}$ of ∇ on X satisfies $\frac{\omega(X,Y)}{\hbar} \equiv \frac{i}{2}\left([\nabla_X, \nabla_Y] - \nabla_{[X,Y]}\right)$ equation (91) implies

$$\nabla_{[X,Y]} s - \frac{2i}{\hbar}\omega(X, Y)s = 0 \ \forall X, Y \in Vect(M, \mathbb{R}; P). \tag{92}$$

Condition (92) is automatically satisfied, provided that

(i) $[X, Y] = 0 \ \forall X, Y \in \mathrm{Vect}(M, \mathbb{R}; P)$,

(ii) $\omega(X, Y) = 0 \ \forall X, Y \in \mathrm{Vect}(M, \mathbb{R}; P)$.

Let us first consider condition (i). To this end,

Definition 6.7.1.2. Let P be a distribution on a manifold M. If $[X, Y] \in Vect(M, \mathbb{R}; P)$ $\forall X, Y \in \mathrm{Vect}(M, \mathbb{R}; P)$, the distribution is called **involutive** .

In other words, condition (i) translates to the statement that P on M must be an involutive distribution. To continue,

Definition 6.7.1.3. Let M be a manifold. If for every point on M, a coordinate chart U can be found such that a distribution P on M is spanned by just the derivatives with respect to the coordinates on U, the distribution is called **completely integrable** .

The following important theorem, known as Frobenius theorem, is quoted without proof. For the proof, the interested reader is referred to ([23, p.170]).

Theorem 6.7.1.1. For any distribution P, P is completely integrable if and only if P is involutive.

It follows that condition (i) is equivalent to saying that P on M must be completely integrable. Furthermore,

Definition 6.7.1.4. Let M be a manifold and P a distribution on M. If $N \subset M$ is an n-dimensional submanifold of M with $T_mN \subset P_m$ $\forall m \in M$, N is called an **integral manifold** of the distribution P.

We state without proof, that given a rank r completely integrable distribution P on M, there exists an integral submanifold $N \subset M$ of dimension $n = r$ for all $m \in M$ such that $T_mN = P_m$ (for the proof, see ([23])). Consequently, from condition (i) it follows that for all $m \in M$ there exist integral manifolds $N \subset M$ through $P_m = T_mN$ of dimension n, since the rank of P on the symplectic manifold M is $n(= \dim(P))$.

Now looking at condition (ii) and definition 6.1.1.3, we see that condition (ii) translates to the statement that the integral manifolds $N \subset M$ are Lagrangian submanifolds of M. So, in first instance, one would define a polarization of (M, ω) to be a maximally integrable distribution P of TM such that P_m is a Lagrangian subspace of T_mM for every $m \in M$. However, this definition is far too restrictive. For instance, on a two-dimensional surface a polarization corresponds to a nowhere-vanishing vector field. But the two-dimensional sphere, S^2, does not have a globally defined nowhere-vanishing vector field and hence a polarization in the above sense does not exist for S^2. The solution to this problem is to complexify the tangent bundle of M. That is, TM will be replaced by $TM^{\mathbb{C}} \equiv \coprod_{m \in M} T_mM^{\mathbb{C}} \equiv \coprod_{m \in M} (T_mM \otimes \mathbb{C})$.

Maximally integrable distributions P of $TM^{\mathbb{C}}$, such that $P_m = (T_m N)^{\mathbb{C}}$ for $N \subset M$ a Lagrangian submanifold of M, are more likely to exist. We thus arrive at

Definition 6.7.1.5. Let (M, ω) be a symplectic manifold. A **polarization** P of (M, ω) is a maximally integrable distribution of $TM^{\mathbb{C}}$ such that P_m is a Lagrangian subspace of $T_m M^{\mathbb{C}}$ for all $m \in M$.

To conclude this subsubsection, we define the corresponding polarized sections as follows,

Definition 6.7.1.6. A **polarized section** of the prequantum line bundle X is a section s over M satisfying

$$\nabla_{\bar{X}} s = 0 \;\; \forall X \in Vect(M, \mathbb{C}; P). \tag{93}$$

6.7.2 Kähler polarizations

A particular class of symplectic manifolds for which geometric quantization is fairly well understood and works with comparative ease are Kähler manifolds. These manifolds have natural and well-behaved polarizations. They are the subject of this subsubsection.

To understand what a Kähler manifold is we first need to define what an almost complex manifold and complex manifold are respectively.

Definition 6.7.2.1. An **almost complex manifold** (M, J) is a real manifold M that is equipped with a smooth real tensor field J such that at every point $m \in M$ the linear endomorphisms $J_m : T_m M \to T_m M$ satisfy $J_m \circ J_m = -I_m$ with $I_m : T_m M \to T_m M \;\; v \mapsto v$ the identity operator on $T_m M$.

We need one more ingredient before we come to the definition of a complex manifold,

Definition 6.7.2.2. Let (M, J) be an almost complex manifold. The smooth real tensor field J is called **integrable** if the Nijenhuis tensor $N(X, Y) \equiv [X, Y] + J[JX, Y] + J[X, JY] - [JX, JY]$ vanishes for all $X, Y \in \mathrm{Vect}(M, \mathbb{R})$.

Subsequently,

Definition 6.7.2.3. A **complex manifold** (M, J) is an almost complex manifold such that J is integrable. In this case, the smooth tensor field J is called the **complex structure** .

We are now ready to define what a Kähler manifold is.

Definition 6.7.2.4. A **Kähler manifold** (M, J, ω) is a complex manifold (M, J) which at the same time is a symplectic manifold (M, ω) such that the symplectic form ω and complex structure J are compatible, that is,
$\omega(JX, JY) = \omega(X, Y)\ \forall X, Y \in \mathrm{Vect}(M, \mathbb{R})$.

Now suppose we are given a Kähler manifold (M, J, ω). Let $m \in M$ be arbitrary. Consider the complexified tangent space $(T_mM)^{\mathbb{C}} \equiv T_mM \otimes \mathbb{C}$. We can diagonalize the linear endomorphisms $J_m : T_mM \to T_mM$ in $(T_mM)^{\mathbb{C}}$. To this end, let $\tilde{v}$ be a vector in $(T_mM)^{\mathbb{C}}$ and set $J_m\tilde{v} = \lambda\tilde{v}$ with $\lambda \in \mathbb{C}$ an eigenvalue of J_m. Since we know that $J_m \circ J_m = -I_m$, it follows for the eigenvalues λ of J_m that
$(J_m \circ J_m)(\tilde{v}) = -\tilde{v} \Leftrightarrow J_m(\lambda\tilde{v}) = \lambda^2\tilde{v} = -\tilde{v} \Leftrightarrow \lambda^2 = -1 \Leftrightarrow \lambda = \pm i$.
The eigenspace $(T_mM)^{(1,0)}$ corresponding to the eigenvalue i of J_m is spanned by vectors of the form $v - iJ_mv$ for $v \in T_mM$. Indeed, denote such an element in $\mathrm{Span}(v - iJ_mv)$ by $\tilde{v} \equiv \sum_i c_i(v_i - iJ_mv_i)$ for $v_i \in T_mM$ and $c_i \in \mathbb{C}$. Then
$J_m\tilde{v} = J_m \sum_i c_i(v_i - iJ_mv_i) = \sum_i c_i(J_mv_i) - \sum_i i(c_iJ_m(J_mv_i) =$
$\sum_i c_i(J_mv_i) + i\sum_i c_iv_i = i\sum_i c_i(v_i - iJ_mv_i) = i\tilde{v}$.
Similarly, the eigenspace $(T_mM)^{(0,1)}$ corresponding to the eigenvalue $-i$ of J_m is spanned by vectors of the form $v + iJ_mv$ for $v \in T_mM$.

Furthermore, since (M, J, ω) is in particular a symplectic manifold it has real dimension $2n$ and so does its tangent space T_mM for $m \in M$. It follows that $(T_mM)^{\mathbb{C}}$ has complex dimension $2n$. The spaces $(T_mM)^{(1,0)}$ and $(T_mM)^{(0,1)}$ are formal complex conjugates of each other, since the formal complex conjugate of an element in $\overline{(T_mM)^{(1,0)}}$ is in an element in $(T_mM)^{(0,1)}$. Indeed,
$\overline{\sum_i c_i(v_i - iJ_mv_i)} = \sum_i c_i\overline{(v_i - iJ_mv_i)} =$
$\sum_i c_i\overline{v_i} + i\sum_i c_i\overline{J_mv_i} = \sum_i c_i(v_i + iJ_mv_i)$.
Hence, the spaces $(T_mM)^{(1,0)}$ and $(T_mM)^{(0,1)}$ have the same dimension n, since $(T_mM)^{(1,0)} \cup (T_mM)^{(0,1)} = T_mM^{\mathbb{C}}$. Furthermore,
$(T_mM)^{(1,0)} \cap (T_mM)^{(0,1)} = \{0\}$ with 0 the zero vector in $T_mM^{\mathbb{C}}$.

As a next step we show that $(T_mM)^{(0,1)}$ is a Lagrangian subspace of $(T_mM)^{\mathbb{C}}$. We see that $(T_mM)^{(0,1)}$ is a linear subspace of $(T_mM)^{\mathbb{C}}$, since the zero vector is contained in it (take $v_i = 0$) and it is clearly closed under vector addition and scalar multiplication. That it is even a Lagrangian subspace of $(T_mM)^{\mathbb{C}}$ follows from
$\omega_m\left(\sum_i c_i(v_i - iJ_mv_i), \sum_i d_i(v_i - iJ_mv_i)\right) =$
$\omega_m\left(J_m\sum_i c_i(v_i - iJ_mv_i), J_m\sum_i d_i(v_i - iJ_mv_i)\right) =$
$\omega_m\left(i\sum_i c_i(v_i - iJ_mv_i), i\sum_i d_i(v_i - iJ_mv_i)\right) =$
$-\,\omega_m\left(\sum_i c_i(v_i - iJ_mv_i), \sum_i d_i(v_i - iJ_mv_i)\right) \Leftrightarrow$

$\omega_m(v,w) = 0 \ \forall v, w \in (T_mM)^{(0,1)}$,
where $\omega_m : (T_mM)^{\mathbb{C}} \times (T_mM)^{\mathbb{C}} \to \mathbb{C}$ is the alternating multilinear map defined by the symplectic 2-form ω for $m \in M$.

Now construct from this the distribution $TM^{(0,1)} \equiv \coprod_m (T_mM)^{(1,0)}$ on $TM^{\mathbb{C}}$. The distribution $TM^{(0,1)}$ is maximally integrable. To see this, first note that this is equivalent to showing that the distribution is involutive. Secondly, because M is a complex manifold the complex structure J defined on it is integrable. It follows that for any $X, Y \in \mathrm{Vect}(M, \mathbb{R})$ it holds that
$N(X,Y) \equiv [X,Y] + J[JX,Y] + J[X,JY] - [JX,JY] = 0 \Leftrightarrow$
$i[JX,Y] + i[X,JY] = iJ[X,Y] - iJ[JX,JY]$.
Take two vector fiels $\tilde{X} \equiv \sum_i c_i(X_i - iJX_i) \in TM^{(0,1)}$ and
$\tilde{Y} \equiv \sum_j d_j(Y_j - iJY_j) \in TM^{(0,1)}$ for $X_i, Y_j \in \mathrm{Vect}(M,\mathbb{R})$ and $c_i, d_j \in \mathbb{C}$. That $TM^{(0,1)}$ is involutive now follows from
$\Big[\sum_i c_i(X_i - iJX_i), \sum_j d_j(Y_j - iJY_j)\Big] = \sum_{i,j} c_i d_j \, [X_i - iJX_i, Y_j - iJY_j] =$
$\sum_{i,j} c_i d_j \, [X_i, Y_j] - i[JX_i, Y_j] - i[X_i, JY_j] - [JX_i, JY_j] =$
$\sum_{i,j} c_i d_j \, [X_i, Y_j] - iJ[X_i, Y_j] + iJ[JX_i, JY_j] - [JX_i, JY_j] =$
$\sum_{i,j} c_i d_j \, (I - iJ)[X_i, Y_j] - (I - iJ)[JX_i, JY_j] =$
$\sum_{i,j} c_i d_j (I - iJ)([X_i, Y_j] - [JX_i, JY_j]) \in TM^{(0,1)}$.
In the third identity equation the vanishing of the Nijenhuis tensor is used. As a consequence we have shown that the space $TM^{(0,1)}$ is a polarization of (M, J, ω)! We call this polarization a **Kähler polarization** since it is naturally constructed on a Kähler manifold. Similarly, one can check that $TM^{(1,0)}$ is a Kähler polarization of (M, J, ω).

Is there an easy way we can recognize whether a symplectic manifold is a Kähler manifold, so that we can construct the Kähler polarization on it? That there indeed is such a way can be seen by going to local coordinates. If we introduce local coordinates z and $\bar{z}$ on the patches of a Kähler manifold (M, J, ω), in first instance the Kähler form (the symplectic form compatible with the complex structure) can locally be written as
$\omega = i\omega_{ij}(z,\bar{z})dz^i \wedge dz^j + i\omega_{i\bar{j}}(z,\bar{z})dz^i \wedge d\bar{z}^{\bar{j}} +$
$i\omega_{\bar{i}j}(z,\bar{z})d\bar{z}^{\bar{i}} \wedge dz^j + i\omega_{\bar{i}\bar{j}}(z,\bar{z})d\bar{z}^{\bar{i}} \wedge d\bar{z}^{\bar{j}}$.
From the additional structure on a Kähler manifold one can deduce that the ω_{ij}-component and $\omega_{\bar{i}\bar{j}}$-component of ω vanish and more particularly, that ω takes the form

$$\omega = 2i\omega_{i\bar{j}}(z,\bar{z})dz^i \wedge d\bar{z}^{\bar{j}}. \tag{94}$$

The proof of this will not be provided in this thesis and we refer the reader

to ([6]) for a much more detailed derivation. The fact that the Kähler form is closed has important consequences. Indeed,
$d\omega = 0 \Leftrightarrow 2id(\omega_{i\bar{j}}(z,\bar{z})dz^i \wedge d\bar{z}^{\bar{j}}) = 0 \Leftrightarrow$
$2i\frac{\partial\omega_{i\bar{j}}(z,\bar{z})}{\partial z^k}dz^k \wedge dz^i \wedge d\bar{z}^{\bar{j}} + 2i\frac{\partial\omega_{i\bar{j}}(z,\bar{z})}{\partial \bar{z}^k}d\bar{z}^{\bar{k}} \wedge dz^i \wedge d\bar{z}^{\bar{j}} = 0.$
Both parts of the expression above must separately vanish. Denoting $\partial_i = \frac{\partial}{\partial z^i}$ and $\partial_{\bar{i}} = \frac{\partial}{\partial \bar{z}^i}$, this implies that

- $\partial_k \omega_{i\bar{j}} dz^k \wedge dz^i \wedge d\bar{z}^{\bar{j}} = 0 \Leftrightarrow$
  $\partial_k \omega_{i\bar{j}} dz^k \wedge dz^i \wedge d\bar{z}^{\bar{j}} + \partial_i \omega_{k\bar{j}} dz^i \wedge dz^k \wedge d\bar{z}^{\bar{j}} = 0 \Leftrightarrow$
  $\partial_k \omega_{i\bar{j}} - \partial_i \omega_{k\bar{j}} = 0,$
- $\partial_{\bar{k}} \omega_{i\bar{j}} d\bar{z}^{\bar{k}} \wedge dz^i \wedge d\bar{z}^{\bar{j}} = 0 \Leftrightarrow$
  $\partial_{\bar{k}} \omega_{i\bar{j}} d\bar{z}^{\bar{k}} \wedge dz^i \wedge d\bar{z}^{\bar{j}} + \partial_{\bar{j}} \omega_{i\bar{k}} d\bar{z}^{\bar{j}} \wedge dz^i \wedge d\bar{z}^{\bar{k}} = 0 \Leftrightarrow$
  $\partial_{\bar{k}} \omega_{i\bar{j}} - \partial_{\bar{j}} \omega_{i\bar{k}} = 0.$

This implies that locally in a patch U_a of the Kähler manifold M there exists a function $K_i(z,\bar{z})$, called the **Kähler potential** , such that $\omega_{i\bar{j}} = \partial_i \partial_{\bar{j}} K_a(z,\bar{z})$. Hence the Kähler form can locally be written as

$$\omega = 2i\partial_i \partial_{\bar{j}} K_a(z,\bar{z}) dz^i \wedge d\bar{z}^{\bar{j}} = 2i\partial\bar{\partial} K_a, \tag{95}$$

with $\partial \equiv dz^k \wedge \partial_k$ and $\bar{\partial} \equiv d\bar{z}^k \wedge \partial_{\bar{k}}$ such that $d = \partial + \bar{\partial}$. As a consequence, two natural symplectic potentials on a Kähler manifold are $2i\bar{\partial}K_a$ and $-2i\partial K_a$. Indeed,
$d(i\bar{\partial}K_a(z,\bar{z})) = 2i\frac{\partial}{\partial z^i}\left(\frac{\partial}{\partial \bar{z}^j}K_a(z,\bar{z})dz^i \wedge d\bar{z}^j\right) + 2i\frac{\partial}{\partial \bar{z}^i}\left(\frac{\partial}{\partial \bar{z}^j}K_a(z,\bar{z})d\bar{z}^i \wedge d\bar{z}^j\right) =$
$2i\frac{\partial}{\partial z^i}\left(\frac{\partial}{\partial \bar{z}^j}K_a(z,\bar{z})dz^i \wedge d\bar{z}^j\right) = 2i\frac{\partial}{\partial z^i}\frac{\partial}{\partial \bar{z}^j}K_a(z,\bar{z})dz^i \wedge d\bar{z}^j = \omega$
and
$d(-i\partial K_a(z,\bar{z})) = -2i\frac{\partial}{\partial z^i}\left(\frac{\partial}{\partial z^j}K_a(z,\bar{z})dz^i \wedge dz^j\right) - 2i\frac{\partial}{\partial \bar{z}^i}\left(\frac{\partial}{\partial z^j}K_a(z,\bar{z})d\bar{z}^i \wedge dz^j\right)$
$= -2i\frac{\partial}{\partial \bar{z}^i}\left(\frac{\partial}{\partial z^j}K_a(z,\bar{z})d\bar{z}^i \wedge dz^j\right) = -2i\frac{\partial}{\partial \bar{z}^i}\frac{\partial}{\partial z^j}K_a(z,\bar{z})d\bar{z}^i \wedge dz^j = \omega.$
This concludes the subsection on polarizations.

6.8 The Hilbert space

Equipped with all the necessary mathematical tools, we are now ready to construct the quantum phase space (Hilbert space). At this point it is important to notice that the general theory about quantization is far from completely understood at the moment. We restrict our attention to quantum phase spaces which arise from symplectic manifolds equipped with a Kähler polarization, because here the least difficulties arise. In this case, there are no obstructions in constructing the Hilbert space and defining a natural measure on it. We will determine the quantum Hilbert space corresponding to a

Kähler polarization and discuss the operators acting on it.

Let us denote the space of polarized sections of the prequantum line bundle X over the Kähler manifold (M, J, ω) with respect to a Kähler polarization P by $\Gamma_X(M; P)$. As previously, denote the prequantum Hilbert space by H.

Definition 6.8.1. The **Hilbert space $\mathcal{H}_P$** is the space defined by $\mathcal{H}_P \equiv H \cap \Gamma_X(M; P)$. In other words, it consists of all square-integrable polarized sections of X.

Since a Kähler polarization just picks out a particular subspace of the prequantum Hilbert space, the Hilbert space is well-defined, since the intersection $H \cap \Gamma_X(M; P)$ is nonempty. The technical proof that $\mathcal{H}_P$ does indeed define a Hilbert space (so in particular has a natural inner product defined on it) will not be provided in this thesis. For a detailed proof we refer the reader to ([7, p.136-137]).

The next question to ask is which classical observables define operators on $\mathcal{H}_P$. The requirement is that the prequantization assignment $Q(f)$ for $f \in C^\infty(M, \mathbb{R})$ maps polarized (square-integrable) sections to polarized (square integrable) sections. Let $f \in C^\infty(M, \mathbb{R})$, $X \in \mathrm{Vect}(M, \mathbb{C}; P)$ and $s \in \mathcal{H}_P$. Then
$\nabla_{\bar{X}}(Q(f)s) = \nabla_{\bar{X}}(-i\hbar\nabla_{X_f}s + fs) = -i\hbar\nabla_{\bar{X}}\nabla_{X_f}s + f\nabla_{\bar{X}}s + \bar{X}(f)s =$
$-i\hbar\nabla_{X_f}\nabla_{\bar{X}}s - i\hbar[\nabla_{\bar{X}}, \nabla_{X_f}]s + f\nabla_{\bar{X}}s + \bar{X}(f)s =$
$-i\hbar\nabla_{X_f}\nabla_{\bar{X}}s - 2\omega(\bar{X}, X_f) - i\hbar\nabla_{[\bar{X}, X_f]}s + f\nabla_{\bar{X}}s + \bar{X}(f)s =$
$-i\hbar\nabla_{X_f}\nabla_{\bar{X}}s - i\hbar\nabla_{[\bar{X}, X_f]}s + f\nabla_{\bar{X}}s = Q(f)(\nabla_{\bar{X}}s) - i\hbar\nabla_{[\bar{X}, X_f]}s$.
In the fourth identity we used that the curvature 2-form $\omega/\hbar$ of ∇ on M satisfies $\omega(\bar{X}, X_f)/\hbar = \frac{i}{2}([\nabla_{\bar{X}}, \nabla_{X_f}] - \nabla_{[\bar{X}, X_f]})$. In the fifth identity we used that ω is non-degenerate and vanishes on the polarization P and $\bar{X}(f) = 0$ on P. That $\bar{X} = 0$ follows from the fact that the polarized sections satisfy $\nabla_{\bar{X}}s = 0$. Writing $s = g\hat{1}$ for $g \in C^\infty(M, \mathbb{C})$ and β for the potential 1-form corresponding to ∇ this indeed implies that
$\bar{X}(g)\hat{1} - i\beta(\bar{X})s = 0 \Leftrightarrow \bar{X}(g)\hat{1} = 0 \Leftrightarrow \bar{X} = 0$.
The first equivalence follows from the fact that under the regularity conditions assumed in this subsection one can locally always find potential 1-forms corresponding to ∇ which vanish on P (which means $\beta(\bar{X}) = 0$). Such potential 1-forms are called locally **adapted** to P. As a consequence of the above discussion we see that polarized sections are indeed mapped to polarized sections by $Q(f)$ provided that

$$[X_f, \bar{X}] \in \mathrm{Vect}(M, \mathbb{C}; P) \ \forall \bar{X} \in \mathrm{Vect}(M, \mathbb{C}; P). \tag{96}$$

This condition (96) tells us that a classical observable f on M defines an operator on $\mathcal{H}_P$ via the assignment $Q(f)$ provided that its flow leaves the polarization invariant. We denote the functions in $C^\infty(M,\mathbb{R})$ whose flow leaves the polarization invariant by $C^\infty(M,\mathbb{R};P)$.

Now take the specific Kähler polarization $P = TM^{(0,1)}$, as defined in subsubsection 6.7.1, for a 2-dimensional Kähler manifold M. Locally, introducing the complex coordinates z and $\bar{z}$ on M, it is spanned by the vector $\frac{\partial}{\partial \bar{z}}$. The observables preserving this polarization are of the form

$$f(z,\bar{z}) = f_0 + f_1 z + \bar{f}_1 \bar{z} + f_2 z\bar{z} \tag{97}$$

with $f_0, f_2 \in \mathbb{R}$ and $f_1 \in \mathbb{C}$. Indeed, using (A.1) we see that the Hamiltonian vector field on M is given by

$$X_f = \frac{1}{2}\omega^{z\bar{z}}\frac{\partial f}{\partial \bar{z}}\frac{\partial}{\partial z} + \frac{1}{2}\omega^{\bar{z}z}\frac{\partial f}{\partial z}\frac{\partial}{\partial \bar{z}}. \tag{98}$$

Hence condition (96) gives
$f \in C^\infty(M,\mathbb{R};P) \Leftrightarrow [X_f, \frac{\partial}{\partial \bar{z}}] \in \mathrm{Vect}(M,\mathbb{C};P) \Leftrightarrow$
$-\frac{1}{2}\omega^{z\bar{z}}\frac{\partial^2 f}{\partial \bar{z}\partial \bar{z}}\frac{\partial}{\partial z} - \frac{1}{2}\omega^{\bar{z}z}\frac{\partial^2 f}{\partial \bar{z}\partial z}\frac{\partial}{\partial \bar{z}} \in \mathrm{Vect}(M,\mathbb{C};P) \Leftrightarrow \frac{\partial^2 f}{\partial \bar{z}\partial \bar{z}} = 0 \Leftrightarrow$
$f(z,\bar{z}) = g(\bar{z}) + h(\bar{z})z$,
with $g(\bar{z})$ and $h(\bar{z})$ smooth complex valued functions only depending on $\bar{z}$. But since observables are real and hence satisfy $\overline{f(z,\bar{z})} = f(z,\bar{z})$, equation (97) follows.

We will conclude this section by showing that for the harmonic oscillator, a physically relevant and easy application, the quantization construction described above already fails to give correct results. The required modification which is needed to give the correct results is called metaplectic quantization. We will not discuss metaplectic quantization in this thesis, but will show that it gives the correct results for the harmonic oscillator. The Hamiltonian of the harmonic oscillator in 1 dimension is given by

$$H(q,p) = \frac{1}{2}(p^2 + q^2) = z\bar{z}, \tag{99}$$

where q, p are the position and momentum coordinate on the phase space of the harmonic oscillator, which is $T^*\mathbb{R}^2$, and $z \equiv \frac{1}{\sqrt{2}}(p + iq)$. Note that $T^*\mathbb{R}^2$ is a 2-dimensional Kähler manifold, with complex structure defined by
$J(\frac{\partial}{\partial q}) = -\frac{\partial}{\partial p},\ J(\frac{\partial}{\partial p}) = \frac{\partial}{\partial q}$
and symplectic form $\omega = idz \wedge d\bar{z}$. Hence $H(q,p)$ is an observable which preserves the polarization $P = TM^{(0,1)}$, since it is a specific case of (97). In

terms of z and $\bar{z}$ the symplectic form ω equals
$\omega = dp \wedge dq = \frac{-i}{2}(dz + d\bar{z}) \wedge (dz - d\bar{z}) = \frac{i}{2}(dz \wedge d\bar{z}) - \frac{i}{2}(d\bar{z} \wedge dz)$.
Hence the Hamiltonian vector field associated with the observable H equals, using (A.1),
$X_H = i\frac{\partial H}{\partial \bar{z}}\frac{\partial}{\partial z} - i\frac{\partial H}{\partial z}\frac{\partial}{\partial \bar{z}} = iz\frac{\partial}{\partial z} - i\bar{z}\frac{\partial}{\partial \bar{z}}$.
Furthermore, the polarized sections with respect to $P = TM^{(0,1)}$ are holomorphic scalar functions (scalar functions depending on the z coordinate only). Indeed, the connection along directions in P takes the form of an 'ordinary' directional derivative, since the potential 1-form β corresponding to ∇ vanishes on P. This implies, for $s \in \Gamma_X(M; P)$, $X \in \mathrm{Vect}(M, \mathbb{C}; P)$ and ψ the scalar function associated with s, that
$\nabla_{\bar{X}} s = 0 \Leftrightarrow \frac{d}{d\bar{z}}\psi(z, \bar{z}) = 0 \Leftrightarrow \psi(z, \bar{z}) = \psi(z)$.
As a consequence, the assignment $Q(H)$ acting on holomorphic functions, equals,
$Q(H)\psi(z) = -i\hbar(\nabla_{X_H} - i\beta(X_H))\psi(z) + H\psi(z) =$
$-i\hbar X_H(\psi(z)) - i\hbar(-i\beta(X_H)) + H\psi(z) =$
$\hbar z\frac{\partial}{\partial z}\psi(z) - i\hbar(-i)\left(i\frac{zd\bar{z}}{\hbar}\right)\left(iz\frac{\partial}{\partial z} - i\bar{z}\frac{\partial}{\partial \bar{z}}\right)\psi(z) + z\bar{z}\psi(z) =$
$\hbar z\frac{\partial}{\partial z}\psi(z) - z\bar{z} + z\bar{z} = \hbar z\frac{\partial}{\partial z}\psi(z)$.
Hence the quantized Hamiltonian corresponds with the operator $\hbar z\frac{\partial}{\partial z}$. The eigenfunctions of this operator are the monomials z^n $(n \geq 0)$ with eigenvalues $n\hbar$. Indeed,
$\hbar z\frac{\partial}{\partial z}z^n = \hbar z n z^{n-1} = n\hbar z^n$.
But as the reader might know from introductory quantum mechanics this spectrum of the Hamiltonian is incorrect and should be shifted by the ground state energy $\frac{1}{2}\hbar$. The problem arises from the fact that viewed as an operator the Hamiltonian $\hat{H} = \hat{z}\hat{\bar{z}}$ is not Hermitian, whereas the symmetrized Hamiltonian $\hat{H} = \frac{1}{2}(\hat{z}\hat{\bar{z}} + \hat{\bar{z}}\hat{z})$ is. This symmetrization process can be incorporated by a quantization procedure called metaplectic quantization.

A general rule of thumb for including the **metaplectic correction** (in this case the term $\frac{1}{2}\hbar$) to the operator corresponding to a polarization preserving observable is the following (see [24]). Let the polarization P be spanned by the n complex vector fields X_k on a Kähler manifold M, $k \in \{1, ..., n\}$. If f preserves the polarization P, there is a matrix $\mathbf{a}(f) \equiv (a_{kl}(f))$ of functions on M satisfying

$$[X_f, X_k] = a_k^l(f)X_l, \tag{100}$$

as follows from (96). In terms of this matrix, the corrected quantum operator is (see for example [9]),

$$\tilde{Q}(f) = Q(f) - \frac{1}{2}i\hbar\ \mathrm{tr}(\mathbf{a}(f)). \tag{101}$$

Furthermore,
$[X_H, \frac{\partial}{\partial \bar{z}}] = -i \left(\frac{\partial}{\partial \bar{z}} z\right) \frac{\partial}{\partial z} + i \left(\frac{\partial}{\partial \bar{z}} \bar{z}\right) \frac{\partial}{\partial \bar{z}} = i \frac{\partial}{\partial \bar{z}}$.
It follows that the matrix $\mathbf{a}(f)$ in this case equals

$$\mathbf{a}(f) = \begin{pmatrix} 0 & 0 \\ 0 & i \end{pmatrix}.$$

Hence $\mathrm{tr}(\mathbf{a}(f)) = i$. Thus the corrected quantum operator becomes
$Q(H) = \hbar z \frac{\partial}{\partial z} - \frac{1}{2} i\hbar(i) = \hbar \left(z \frac{\partial}{\partial z} + \frac{1}{2}\right)$.
The eigenfunctions of this operators are again the monomials z^n $(n \geq 0)$ with the correct eigenvalues $(n + \frac{1}{2})\hbar$. Indeed,
$\hbar \left(z \frac{\partial}{\partial z} z^n + \frac{1}{2} z^n\right) = ((n + \frac{1}{2})\hbar) z^n$.

7 The orbit method and the representations of SU(2)

-It is the harmony of the diverse parts, their symmetry, their happy balance; in a word it is all that introduces order, all that gives unity, that permits us to see clearly and to comprehend at once both the ensemble and the details.- Henri Poincare about symmetry ([25])

The goal of the orbit method is to establish a correspondence between irreducible unitary reprensentations of a Lie group and its coadjoint orbits. It thus provides a general framework for finding all irreducible unitary representations of a Lie group by finding its coadjoint orbits. Representation theory remains the method of choice for symplifying the physical analysis of systems possesing (a high degree of) symmetry and its appearance is ubiquitous in, for example, high energy physics.

In this section the representation theory of SU(2) will be treated by means of the orbit method. The method of investigation is less formal then the one used in sections 4, 5 and 6. In order to determine the irreducible representations of SU(2) by means of the orbit method we proceed as follows. We construct the quantum phase space on the coadjoint orbits of SU(2) determined in section 4 (the spheres) which satisfy the integrality condition. To this end, we first have to check what the integrality means in the case of a sphere. Subsequently, we quantize the observables on these spheres by geometric quantization, as described in section 6. Since SU(2) is a compact manifold, its representations were already well-known to mathematicians long before the orbit method was invented. In order to finally understand the representation theory of SU(2) by means of the orbit method, it is worthwile to first determine the representations of SU(2) by the old-fashioned method.

7.1 The integrality condition for the coadjoint orbits of SU(2)

Denote by S^2 the sphere of some arbitrary radius $R \in \mathbb{R}$ and by ω the symplectic form defined on it. In order to understand what the integrality condition means in the case of a sphere of radius R we proceed by proving that the statements beneath are equivalent.

(I_1) There exists an open cover $U = \{U_j\}$ of S^2 such that the class defined by ω in $H^2(U, \mathbb{R})$ contains a cocycle $\tilde{a}$ in which all the $\tilde{a}_{ijk}$ are integral multiples of 2π.

(I_2) The integral of ω over S^2 is an integral multiple of 2π.

Theorem 7.1.1. The formulations (I_1) and (I_2) of the Integrality Condition for S^2 are equivalent.

Proof. We proceed by showing the implications $(\Leftarrow)$ and $(\Rightarrow)$ seperately.

- $(I_1) \Leftarrow (I_2)$: suppose that $\int_{S^2} \omega \in 2\pi\mathbb{Z}$ is given. Now consider the pyramid cover $U = \{U_i\}$ $i \in 1, 2, 3, 4$ of the sphere depicted in figure 4 below.

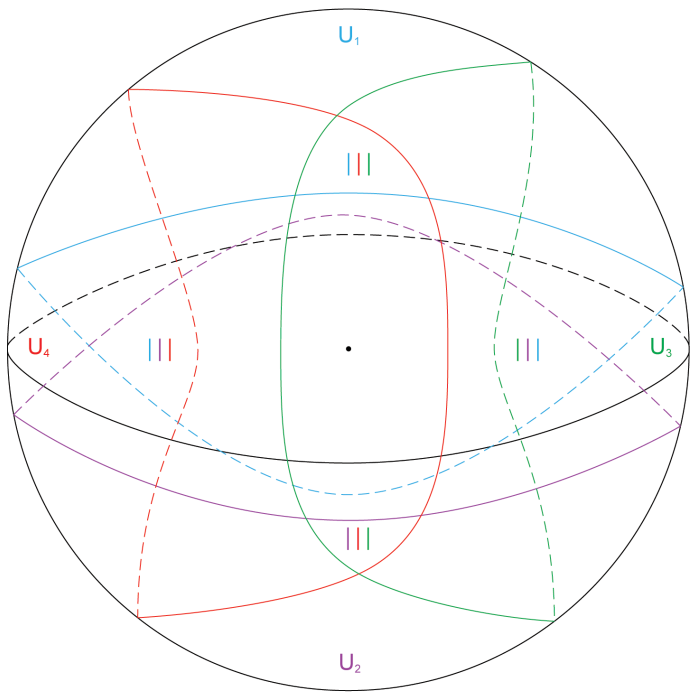

Figure 4: The pyramid cover of the sphere. The sphere is covered by the four patches U_1 (blue), U_2 (violet), U_3 (green) and U_4 (red) as indicated in the figure. Considering the intersections U_{123}, U_{124}, U_{134} and U_{234} and picking a point in the center of each of them, drawing all the geodesics between these points gives the edges of a regular tetrahedron.

In subsection 6.5 it is already proved that the existence of the closed real 2-form ω on a symplectic manifold M (so in particular a closed real 2-form on the sphere) gives the existence of real 1-form β_j on U_j such that $\omega = d\beta_j$ on U_j, the existence of anti-symmetric smooth real functions $f_{ij} = -f_{ji}$ on U_{ij} such that $df_{ij} = \beta_i - \beta_j$ on U_{ij} and the

existence of locally constant smooth real functions $a_{ijk} \equiv f_{ij} + f_{jk} + f_{ki}$ on U_{ijk}. When considering the pyramid cover of the sphere this gives the globally constant functions (since the U_{ijk} are connected)
$a_{123} \equiv f_{12} + f_{23} + f_{31}$ on U_{123}
$a_{124} \equiv f_{12} + f_{24} + f_{41}$ on U_{124}
$a_{134} \equiv f_{13} + f_{34} + f_{41}$ on U_{134}
$a_{234} \equiv f_{23} + f_{34} + f_{42}$ on U_{234}.
We can define new $\tilde{f}_{ij}$ from the old f_{ij} by
$\tilde{f}_{12} \equiv f_{12} - a_{124}$ on U_{124}
$\tilde{f}_{13} \equiv f_{13} - a_{134}$ on U_{134}
$\tilde{f}_{23} \equiv f_{23} - a_{234}$ on U_{234}
and leaving the other f_{ij} unchanged, since $da_{ijk} = 0$ and so ω remains the same. In terms of the new $\tilde{f}_{ij}$ the new $\tilde{a}_{ijk} \equiv \tilde{f}_{ij} + \tilde{f}_{jk} + \tilde{f}_{ki}$ become
$\tilde{a}_{124} \equiv \tilde{f}_{12} + \tilde{f}_{24} + \tilde{f}_{41} = f_{12} - a_{124} + f_{24} + f_{41} =$
$f_{12} - (f_{12} + f_{24} + f_{41}) + f_{24} + f_{41} = 0$
$\tilde{a}_{134} \equiv \tilde{f}_{13} + \tilde{f}_{34} + \tilde{f}_{41} = f_{13} - a_{134} + f_{34} + f_{41} =$
$f_{13} - (f_{13} + f_{34} + f_{41}) + f_{34} + f_{41} = 0$
$\tilde{a}_{234} \equiv \tilde{f}_{23} + \tilde{f}_{34} + \tilde{f}_{42} = f_{23} - a_{234} + f_{34} + f_{42} =$
$f_{23} - (f_{23} + f_{34} + f_{42}) + f_{34} + f_{42} = 0$
$\tilde{a}_{123} \neq 0.$
Hence $\tilde{a}_{124}, \tilde{a}_{134}, \tilde{a}_{234} \in 2\pi\mathbb{Z}$. If we can prove that $\int_{S^2} \omega = -\tilde{a}_{123}$ it follows that $\tilde{a}_{123} \in 2\pi\mathbb{Z}$ and consequently, using the results of subsection 6.5, we have found an open cover U of S^2 such that the class defined by ω in $H^2(U, \mathbb{R})$ contains a cocycle $\tilde{a}$ in which all the $\tilde{a}_{ijk}$ are integral multiples of 2π. Indeed, for a consistent choice of orientations on the boundaries of the U_i,
$\int_{S^2} \omega = \int_{U_1} \omega + \int_{U_2} \omega + \int_{U_3} \omega + \int_{U_4} \omega =$
$\int_{U_1} d\beta_1 + \int_{U_2} d\beta_2 + \int_{U_3} d\beta_3 + \int_{U_4} d\beta_4 =$
$\int_{\partial U_1} \beta_1 + \int_{\partial U_2} \beta_2 + \int_{\partial U_3} \beta_3 + \int_{\partial U_4} \beta_4 =$
$- \int_{U_{12}} \beta_1 + \int_{U_{13}} \beta_1 - \int_{U_{14}} \beta_1 + \int_{U_{21}} \beta_2 - \int_{U_{23}} \beta_2 + \int_{U_{24}} \beta_2 -$
$\int_{U_{31}} \beta_3 + \int_{U_{32}} \beta_3 - \int_{U_{34}} \beta_3 + \int_{U_{41}} \beta_4 - \int_{U_{42}} \beta_4 + \int_{U_{43}} \beta_4 =$
$\int_{U_{12}} (-\beta_1 + \beta_2) + \int_{U_{13}} (\beta_1 - \beta_3) + \int_{U_{14}} (-\beta_1 + \beta_4) +$
$\int_{U_{23}} (-\beta_2 + \beta_3) + \int_{U_{24}} (\beta_2 - \beta_4) + \int_{U_{34}} (-\beta_3 + \beta_4) =$
$\int_{U_{12}} d\tilde{f}_{21} + \int_{U_{13}} d\tilde{f}_{13} + \int_{U_{14}} d\tilde{f}_{41} + \int_{U_{23}} d\tilde{f}_{32} + \int_{U_{24}} d\tilde{f}_{24} + \int_{U_{34}} d\tilde{f}_{43} =$
$\int_{\partial U_{12}} \tilde{f}_{21} + \int_{\partial U_{13}} \tilde{f}_{13} + \int_{\partial U_{14}} \tilde{f}_{41} + \int_{\partial U_{23}} \tilde{f}_{32} + \int_{\partial U_{24}} \tilde{f}_{24} + \int_{\partial U_{34}} \tilde{f}_{43} =$
$(\tilde{f}_{21}(u_{123}) - \tilde{f}_{21}(u_{124})) + (\tilde{f}_{13}(u_{123}) - \tilde{f}_{13}(u_{134})) + (\tilde{f}_{41}(u_{124}) - \tilde{f}_{41}(u_{134})) +$
$(\tilde{f}_{32}(u_{123}) - \tilde{f}_{32}(u_{234})) + (\tilde{f}_{24}(u_{124}) - \tilde{f}_{24}(u_{234})) + (\tilde{f}_{43}(u_{134}) - \tilde{f}_{43}(u_{234})) =$
$-\tilde{a}_{123}(u_{123}) + \tilde{a}_{124}(u_{124}) + \tilde{a}_{134}(u_{134}) + \tilde{a}_{234}(u_{234}) = -\tilde{a}_{123}(u_{123}) = -\tilde{a}_{123}(\cdot).$
In the second identity we used Stokes' theorem. In the third iden-

tity we used that $\partial U_1 = U_{12} \cup U_{13} \cup U_{14}$, et cetera. In the sixth identity again Stokes' theorem is used. In the seventh identity the $u_{123}, u_{124}, u_{234}, u_{134}$ denote points in $U_{123}, U_{124}, U_{234}, U_{134}$ respectively and is used that $\partial V_{12} = V_{123} \cup V_{124}$, et cetera. In the last identity is used that $\tilde{a}_{123}$ is constant.

- $(I_1) \Rightarrow (I_2)$: this implication is now easy to prove. Suppose we are given the pyramid cover considered above such that the class defined by ω in $H^2(U, \mathbb{R})$ contains a cocycle $\tilde{a}$ in which all the $\tilde{a}_{ijk}$s are integral multiples of 2π. Such $\tilde{a}_{ijk}$ always exist, since the intersections U_{ijk} are all connected for the pyramid cover, from which it follows that the $\tilde{a}_{ijk}$ are globally constant. We thus know that $\tilde{a}_{123}$, $\tilde{a}_{124}$, $\tilde{a}_{134}$ and $\tilde{a}_{124}$ are integral multiples of 2π. It follows that
$-\tilde{a}_{123}(u_{123}) + \tilde{a}_{124}(u_{124}) + \tilde{a}_{134}(u_{134}) + \tilde{a}_{234}(u_{234}) = \int_{S^2} \omega$
and hence that $\int_{S^2} \omega \in 2\pi\mathbb{Z}$, since $\tilde{a}_{123}$, $\tilde{a}_{124}$, $\tilde{a}_{134}$ and $\tilde{a}_{124}$ are integral multiples of 2π. The implication follows.

□

As a consequence, for a Hermitian line bundle-with-connection over S^2 to exist (the prequantum line bundle over S^2) condition (I_2) should be satisfied. We already know from subsubsection 5.2.1 that $\omega = R\sin(\theta) d\theta \wedge d\phi$. Hence condition (I_2) gives
$\int_{S^2} \omega = \int_0^{\pi} \int_0^{2\pi} R \sin(\theta) d\theta d\phi = 2\pi R \int_0^{\pi} \sin(\theta) d\theta = 4\pi R \in 2\pi\mathbb{Z} \Leftrightarrow R \in \mathbb{Z}/2$.
In other words, we can construct the prequantum line bundle for all spheres of half-integer radius. We are now ready to construct the quantum phase space on the spheres of half-integer radius. This shall be done in the next subsection.

7.2 Constructing the Hilbert space on the spheres of half-integer radius

Consider a sphere S^2 of half-integer radius $k \in \mathbb{Z}/\not{\not{\mathbb{Z}}}$. We cover the sphere by the two charts already introduced in subsubsection 5.2.1, i.e. the cover $U \equiv \{U_\pm\}$. As explained in subsubsection 5.2.2 these charts are isomorphic to $\mathbb{C}$. We can write
$S^2 = \mathbb{C} \coprod \mathbb{C}/\sim$ with $\sim$ the equivalence relation defined by $(1, z_1) \sim (1, z_2)$ when $z_1 = z_2$, $(2, z_1) \sim (2, z_2)$ when $z_1 = z_2$ and $(1, z_1) \sim (2, z_2)$ when $z_2 = \frac{1}{z_1}$ and $z_1 \neq 0$. Here $(1, z)$ for $z \in \mathbb{C}$ denotes an element in the first copy of $\mathbb{C}$ in the definition for S^2 and $(2, z)$ for $z \in \mathbb{C}$ denotes an element in the second copy of $\mathbb{C}$ in the definition for S^2. Indeed $\sim$ defines an equivalence relation, since

- $(i,z) \sim (i,z)$: we have $(i,z) \sim (i,z) \Leftrightarrow z = z$ for both $i = 1$ and $i = 2$, which is indeed true.

- $(i,z_1) \sim (j,z_2) \Rightarrow (j,z_2) \sim (i,z_1)$: we will prove this relation for the case $i = 1$, $j = 2$. Suppose $(1,z_1) \sim (2,z_2)$, which means $z_2 = \frac{1}{z_1}$ and $z_1 \neq 0$. Then $z_1 = \frac{1}{z_2}$ and $z_2 \neq 0$. Hence $(2,z_2) \sim (1,z_1)$.

- $(i,z_1) \sim (j,z_2), (j,z_2) \sim (k,z_3) \Rightarrow (i,z_1) \sim (k,z_3)$: we will prove this relation for the case $i = 1$, $j = 2$, $k = 1$. Suppose $(1,z_1) \sim (2,z_2)$ and $(2,z_2) \sim (1,z_3)$. Then, by definition, $z_2 = \frac{1}{z_1}$, $z_1 \neq 0$, $z_3 = \frac{1}{z_2}$ and $z_2 \neq 0$. Hence $z_3 = \frac{1}{\left(\frac{1}{z_1}\right)} = z_1$. It follows that $(1,z_1) \sim (1,z_3)$.

On the overlap of the first copy of $\mathbb{C}$ and the second copy of $\mathbb{C}$ in the definition for S^2 points are indeed identified by $\sim$ if one views the first copy of $\mathbb{C}$ (minus the point $(1,0)$, but the equivalence relation does act trivially on this point anyway) as the image under the map $\alpha : \mathbb{C}/\{0\} \to \mathbb{C}\ z \mapsto z$ and the second copy of $\mathbb{C}$ (minus the point $(2,0)$) as the image under the map $\beta : \mathbb{C}/\{0\} \to \mathbb{C}\ z \mapsto \frac{1}{z}$. The fact that one can let the north pole of S^2 correspond to the point $(1,0)$ and the southpole to the point $(2,0)$ follows from the fact that $(1,\epsilon) \sim (2,\tilde{\epsilon})$ when $\tilde{\epsilon} = \frac{1}{\epsilon}$ for $\epsilon \neq 0$ infinitesimally small.

In subsection 6.5 we have seen that on a symplectic manifold M (with cover $U \equiv \{U_i\}$) one can construct a prequantum line bundle X whose total space is $\coprod_i U_i \times \mathbb{C}/\sim$ with $\sim$ the equivalence relation defined by: $(j,m,z) \sim (k,\tilde{m},\tilde{z})$ whenever $(j,m,z) \in I \times U_j \times \mathbb{C}$, $(k,\tilde{m},\tilde{z}) \in I \times U_k \times \mathbb{C}$ and $m = \tilde{m}$ and $\tilde{z} = c_{jk}(m)z$. In case $M = S^2$ with cover $U \equiv \{U_\pm\}$, the total space becomes $X = \coprod_i U_i \times \mathbb{C}/\sim = (U_1 \times \mathbb{C}) \coprod (U_1 \times \mathbb{C})/\sim = (\mathbb{C} \times \mathbb{C}) \coprod (\mathbb{C} \times \mathbb{C})/\sim$ with $\sim$ the equivalence relation defined by $(1,m,z) \sim (1,\tilde{m},\tilde{z})$ when $m = \tilde{m}$ and $z = \tilde{z}$, $(2,m,z) \sim (2,\tilde{m},\tilde{z})$ when $m = \tilde{m}$ and $z = \tilde{z}$, $(1,m,z) \sim (2,\tilde{m},\tilde{z})$ when $\tilde{m} = \frac{1}{m}$ and $m \neq 0$ and $\tilde{z} = \frac{1}{z^k}z$. Here is used that the transition functions on S^2 of radius $k \in \mathbb{Z}$ with cover U satisfy $c_{11}(m) = 1$, $c_{22}(m) = 1$ and $c_{12}(m) = \frac{1}{z^k}$.

At any point $m \in S^2$ the tangent space of S^2 equals the plane $\mathbb{R}^2 \cong \mathbb{C}$. It follows that $TU_+ \cong U_1 \times \mathbb{C} \cong \mathbb{C} \times \mathbb{C}$ and similarly $TU_- \cong \mathbb{C} \times \mathbb{C}$. Consequently, $(TU_+)^{\mathbb{C}} \equiv \coprod_{m \in U_+} T_m U_+ \cong \mathbb{C} \times \mathbb{C}^2$ and $(TU_-)^{\mathbb{C}} \cong \mathbb{C} \times \mathbb{C}^2$. Together the $(TU_+)^{\mathbb{C}}$ and $(TU_+)^{\mathbb{C}}$ determine $(TS^2)^{\mathbb{C}}$. A point in $(TU_+)^{\mathbb{C}}$ can be denoted by $\left(z, a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y}\right)$ for $z \in \mathbb{C}$, $(x,y) \in \mathbb{R}^2 \cong \mathbb{C}$ and $a,b \in \mathbb{C}$.

As a next step we show that S^2 is a Kähler manifold. We know from subsection 5.2.2 that the symplectic form ω on S^2 equals
$\omega = \frac{i}{2R}\left(1+\frac{z\bar{z}}{4R^2}\right)^{-2} dz \wedge d\bar{z}$. We claim that ω can be written as $\omega = 2i\partial\bar{\partial}K$ for $K = R\log(1+z\bar{z}/4R^2)$, which shows that S^2 is a Kähler manifold. Here $\partial \equiv dz \wedge \frac{\partial}{\partial z}$ and $\bar{\partial} \equiv d\bar{z} \wedge \frac{\partial}{\partial \bar{z}}$. Indeed,
$\bar{\partial}K = \frac{\partial}{\partial \bar{z}}K \wedge d\bar{z} = \frac{\partial}{\partial \bar{z}}(R\log(1+z\bar{z}/4R^2)) \wedge d\bar{z} = \frac{z/4R}{1+z\bar{z}/4R^2} \wedge d\bar{z}$.
Subsequently,
$2i\partial\bar{\partial}K = i\partial\frac{z/2R}{1+z\bar{z}/4R^2} \wedge d\bar{z} = i(dz \wedge \frac{\partial}{\partial z})\left(\frac{z/2R}{1+z\bar{z}/4R^2} \wedge d\bar{z}\right) =$
$i\left(\frac{1/2R}{1+z\bar{z}/4R^2} - \frac{\bar{z}z/4R^2}{(1+\bar{z}z/4R^2)^2}\right) dz \wedge d\bar{z} = \frac{i}{2R}\left(1+\frac{z\bar{z}}{4R^2}\right)^{-2} dz \wedge d\bar{z}$,
which proves the claim.

Now consider the specific Kähler polarization $P \equiv (TS^2)^{(0,1)}$. Locally, taking $z \equiv x+iy$ and $\bar{z} \equiv x-iy$ as complex coordinates on the patches of S^2, it is spanned by the vector $\frac{\partial}{\partial \bar{z}}$. Hence a point in $(TU_+)^{(0,1)}$ can be denoted by $(z, a\frac{\partial}{\partial \bar{z}})$ for $a \in \mathbb{C}$. Furthermore, on the patch U_+, the polarized sections with respect to P are holomorphic complex-valued scalar functions, as already shown in subsection 6.8. In other words, scalar functions f which satisfy $\frac{\partial}{\partial \bar{z}}f = 0$. Since the f are holomorphic they can be written as $f(z) = \sum_{i=0}^{\infty} a_i z^i$ with $a_i \in \mathbb{C}$. From (68) it follows that a section on U_+ evaluated in a point $m \in U_+$ can be written as $\psi(m, f(m))$ with ψ the diffeomorphism $\psi : \mathrm{pr}^{-1}(\mathbb{C}) \to \mathbb{C}\times\mathbb{C}$. Since $\psi(m, f(m)) \in \mathrm{pr}^{-1}(\mathbb{C}) \cong \mathbb{C}\times\mathbb{C}$ we might as well denote the points $\psi(m, f(m))$ by $(m, f(m)$. Using that $S^2 = \mathbb{C}\coprod\mathbb{C}/\sim$ it follows that on the intersections U_{+-} of a sphere of radius $k \in \mathbb{Z}/2$ with cover U we have for $z \in \mathbb{C}$,
$(z, f(z)) \sim \left(\frac{1}{z}, f(\frac{1}{z})\right) = \left(\frac{1}{z}, \sum_{i=0}^{\infty} a_i z^{-i}\right) \sim \left(z, z^k \sum_{i=0}^{\infty} a_i z^{-i}\right)$.
Since $z^k \sum_{i=0}^{\infty} a_i z_i$ should be a holomorphic function, this can only be the case if $k - i \geq 0$. Hence $k \geq i$. Hence f is a polynomial of degree $\leq k$ on S^2 with radius $k \in \mathbb{Z}/2$. Notice that the space of polynomials of degree $\leq k$ on S^2 (with radius $k \in \mathbb{Z}/2$) is indeed an irreducible subspace of the space of all smooth complex-valued scalar functions on S^2 (with radius $k \in \mathbb{Z}/2$. As a consequence, the Hilbert space on S^2 is the space of all square-integrable polynomials of degree $\leq k$ on S^2 (with radius $k \in \mathbb{Z}/2$).

7.3 Quantizing observables on the spheres of half-integer radius

In this subsection we present an explicit description of the quantization assignment in case the symplectic manifold is a sphere of half-integer radius.

We will describe this quantization assignment both in spherical and stereographic coordinates.

7.3.1 Quantizing observables on the spheres of integer radius in the spherical coordinate system

From subsubsection 5.2.1 we know that in spherical coordinates the symplectic form on a patch of S^2 with covering $U \equiv \{U_\pm\}$ equals
$\omega = R\sin(\theta)d\theta \wedge d\phi$. Hence on spheres S^2 of half-integer radius $R \in \mathbb{Z}/2$ the curvature 2-form locally equals $\Omega = \frac{\omega}{\hbar} = \frac{R}{\hbar}\sin(\theta)d\theta \wedge d\phi$. Hence on a patch U_i of such a sphere of radius $R \in \mathbb{Z}/2$ a potential 1-form is given by $\beta = -\frac{R}{\hbar}\cos(\theta)d\phi$. Indeed,
$d\beta = -d(\frac{R}{\hbar}\cos(\theta)d\phi) = -\frac{\partial}{\partial\theta}(\frac{R}{\hbar}\cos(\theta))d\theta \wedge d\phi - \frac{\partial}{\partial\phi}(\frac{R}{\hbar}\cos(\theta))d\phi \wedge d\phi = -\frac{\partial}{\partial\theta}(\frac{R}{\hbar}\cos(\theta))d\theta \wedge d\phi = \frac{R}{\hbar}\sin(\theta)d\theta \wedge d\phi$.

Since the spheres S^2 of half-integer radius $R \in \mathbb{Z}/2$ are Kähler manifolds, they admit the natural Kähler polarization $P \equiv (TS^2)^{(0,1)}$. From subsection 6.8 it follows that the only quantizable observables are the ones preserving the polarizations. Let $f \in C^\infty(S^2, \mathbb{R}; P)$. The corresponding Hamiltonian vector field in spherical coordinates equals, using (A.1),
$X_f \equiv \frac{1}{2}\omega^{kj}\frac{\partial f}{\partial x^j}\frac{\partial}{\partial x^k} = \frac{1}{2}\omega^{\theta\phi}\frac{\partial f}{\partial\theta}\frac{\partial}{\partial\phi} + \frac{1}{2}\omega^{\phi\theta}\frac{\partial f}{\partial\phi}\frac{\partial}{\partial\theta} = \frac{1}{R^2\sin(\theta)}\left(\frac{\partial f}{\partial\theta}\frac{\partial}{\partial\phi} - \frac{\partial f}{\partial\phi}\frac{\partial}{\partial\theta}\right)$.
Hence
$\beta(X_f) = -\left(\frac{R^2}{\hbar}\cos(\theta)d\phi\right)\frac{1}{R^2\sin(\theta)}\left\{\frac{\partial f}{\partial\theta}\frac{\partial}{\partial\phi} - \frac{\partial f}{\partial\phi}\frac{\partial}{\partial\theta}\right\} =$
$-\left(\frac{R^2\cos(\theta)}{R^2\sin(\theta)}\right)\frac{\partial f}{\partial\theta} = -\cot(\theta)\frac{\partial f}{\partial\theta}$.
Now let $s \in \Gamma_X(S^2)$ and write $s = g\hat{1}$ for $g \in C^\infty(S^2, \mathbb{C})$. Then,
$\nabla_{X_f}s = \nabla_{X_f}\left(g\hat{1}\right) \equiv X_f(g)\hat{1} + g\nabla_{X_f}\left(\hat{1}\right) = \{f, g\}\hat{1} - ig\beta(X_f)\hat{1} =$
$\frac{1}{R^2\sin(\theta)}\left(\frac{\partial f}{\partial\theta}\frac{\partial g}{\partial\phi} - \frac{\partial f}{\partial\phi}\frac{\partial g}{\partial\theta}\right)\hat{1} + i\cot(\theta)\frac{\partial f}{\partial\theta}g\hat{1}$.
In the third identity (49) and (67) are used. Finally, the quantization assignment $Q(f)$ in spherical coordinates becomes, using (87),

$$\begin{aligned} Q(f) &= -i\hbar\left\langle\frac{1}{R^2\sin(\theta)}\left(\frac{\partial f}{\partial\theta}\frac{\partial g}{\partial\phi} - \frac{\partial f}{\partial\phi}\frac{\partial g}{\partial\theta}\right) + i\cot(\theta)g\left(\frac{\partial f}{\partial\theta}\right)\right\rangle \\ &\quad + f \text{ for } f \in C^\infty(S^2, \mathbb{R}; P). \end{aligned} \tag{102}$$

7.3.2 Quantizing observables on the spheres of half-integer radius in the stereographic coordinate system

Notice that in spherical coordinates the symplectic form ω has a singular point on both the patches U_+ and U_-, namely the point $\theta = \pi$ in the first case and the point $\theta = 0$ in the second case. One can get around this problem in case the stereographic coordinates are used as local coordinates on the sphere. We know from subsubsection 5.2.2 that in projective coordinates the symplectic form on $\{U_-\}$ equals $\omega = \frac{i}{2R}\left(1 + \frac{z\bar{z}}{4R^2}\right)^{-2} dz \wedge d\bar{z}$. Viewing the south pole of the sphere as the point at ∞, this symplectic form only has a singular point at ∞. One can get rid of this singularity by applying the change of coordinates $z = \frac{1}{\omega}$ for $\omega \in \mathbb{C}$. As a consequence, when $z \to \infty$ then $\omega \to 0$. Applying this change of coordinates one gets, taking $2R = 1$ for notational convenience,

$\omega = \frac{1}{(1+|z|^2)^2} dz \wedge d\bar{z} = \frac{1}{(1+1/\omega^2)^2} d(\frac{1}{\omega}) \wedge d(\frac{1}{\bar{\omega}}) =$

$\frac{|\omega|^4}{(1+|\omega|^2)^2} \frac{1}{\omega^2\bar{\omega}^2} d\omega \wedge d\bar{\omega} = \frac{1}{(1+|\omega|^2)^2} dz \wedge d\bar{\omega}.$

In this new coordinates ω has no singular point on $\{U_-\}$ anymore. Notice that on $\{U_+\}$ the symplectic form ω does have a singular point at ∞, but this can be get rid of by applying the same change of coordinates again. Then ω gets back the original form in which it has no singular points on $\{U_+\}$.

Let us now calculate the quantization assignment in stereographic coordinates. We choose to do this on the chart U_- (the calculation on the chart U_+ is similar). On a sphere of half-integer radius $R \in \mathbb{Z}/2$ it locally holds for the curvature 2-form that $\Omega = \frac{\omega}{\hbar} = \frac{i}{4\hbar R^2}\left(1 + \frac{z\bar{z}}{4R^2}\right)^{-2} dz \wedge d\bar{z}$. On U_- we can choose the corresponding 1-form to be $\beta = \frac{-i}{\hbar}\frac{(\bar{z}/2R)dz}{1+z\bar{z}/4R^2}$. Indeed,

$d\beta = \frac{-i}{\hbar}\frac{\partial}{\partial z}\left(\frac{\bar{z}/2R}{1+z\bar{z}/4R^2}\right) dz \wedge dz - \frac{i}{\hbar}\frac{\partial}{\partial \bar{z}}\left(\frac{\bar{z}/2R}{1+z\bar{z}/4R^2}\right) d\bar{z} \wedge dz =$

$\frac{-i}{\hbar}\frac{\partial}{\partial \bar{z}}\left(\frac{\bar{z}/2R}{1+z\bar{z}/4R^2}\right) d\bar{z} \wedge dz = \frac{i}{\hbar}\left\langle \frac{1/2R}{1+z\bar{z}/4R^2} - \frac{z\bar{z}/4R^2}{(1+z\bar{z}/4R^2)^2}\right\rangle dz \wedge d\bar{z} =$

$\frac{i}{2R\hbar}\left(1 + \frac{z\bar{z}}{4R^2}\right)^{-2} dz \wedge d\bar{z}.$

Let $f \in C^\infty(S^2, \mathbb{R}; P)$. The corresponding Hamiltonian vector fields in projective coordinates are equal too, again using (A.1),

$X_f = \frac{1}{2}\omega^{z\bar{z}}\frac{\partial f}{\partial \bar{z}}\frac{\partial}{\partial z} + \frac{1}{2}\omega^{\bar{z}z}\frac{\partial f}{\partial z}\frac{\partial}{\partial \bar{z}} = R\left(1 + \frac{z\bar{z}}{4R^2}\right)^2\left(\frac{\partial f}{\partial \bar{z}}\frac{\partial}{\partial z} - \frac{\partial f}{\partial z}\frac{\partial}{\partial \bar{z}}\right).$

Consequently,

$\beta(X_f) = \left(\frac{-i}{\hbar}\frac{(\bar{z}/2R)dz}{1+z\bar{z}/4R^2}\right)(R)\left(1 + \frac{z\bar{z}}{4R^2}\right)^2\left(\frac{\partial f}{\partial \bar{z}}\frac{\partial}{\partial z} - \frac{\partial f}{\partial z}\frac{\partial}{\partial \bar{z}}\right) =$

$(\frac{-i}{\hbar}\frac{\bar{z}/2R}{1+z\bar{z}/4R^2})(R)\left(1 + \frac{z\bar{z}}{4R^2}\right)^2\frac{\partial f}{\partial \bar{z}} = \frac{-i\bar{z}}{2\hbar}\left(1 + z\bar{z}/4R^2\right)\frac{\partial f}{\partial \bar{z}}.$

Moreover,

$\nabla_{X_f} s = R\left(1 + \frac{z\bar{z}}{4R^2}\right)^2\left(\frac{\partial f}{\partial \bar{z}}\frac{\partial g}{\partial z} - \frac{\partial f}{\partial z}\frac{\partial g}{\partial \bar{z}}\right)\hat{1} - \frac{\bar{z}}{2\hbar}\left(1 + z\bar{z}/4R^2\right)\frac{\partial f}{\partial \bar{z}} g\hat{1}.$

Finally, the quantization assingment in stereographic coordinates becomes

$$
\begin{aligned}
Q(f) &= -i\hbar \left\langle R\left(1+\frac{z\bar{z}}{4R^2}\right)^2 \left(\frac{\partial f}{\partial \bar{z}}\frac{\partial}{\partial z} - \frac{\partial f}{\partial z}\frac{\partial}{\partial \bar{z}}\right) - \frac{\bar{z}}{2\hbar}\left(1+z\bar{z}/4R^2\right)\frac{\partial f}{\partial \bar{z}}\right\rangle \\
&\quad + f \ \text{ for } \ f \in C^\infty(S^2,\mathbb{R};P).
\end{aligned}
\tag{103}
$$

7.4 The irreducible unitary representations of SU(2)

In this subsection we first determine the irreducible unitary reprensentations of SU(2) by means of the so-called **infinitesimal method**, as described in ([26]). After this, we conclude by giving a heuristic argument how to determine the irreducible unitary representations of SU(2) by means of the orbit method.

7.4.1 Finding the irreducible representations of SU(2) by means of the infinitesimal method

To recall,

$$\mathrm{SU}(2) \equiv \{x \in \mathrm{Mat}(2,\mathbb{C}) | x^\dagger x = \mathrm{I}, \det(x) = 1\}.$$

and

$$\mathfrak{su}(2) = \{X \in \mathrm{Mat}(2,\mathbb{C}) | X = -X^\dagger, \mathrm{tr}X = 0\}.$$

Now look at the complexification of $\mathfrak{su}(2)$, that is $\mathfrak{su}(2)\otimes\mathbb{C}$. Since in $\mathfrak{su}(2)\otimes\mathbb{C}$ the scalar multiplication is extended to the complex numbers, let us look at what this implies for a particular matrix $A \in \mathfrak{su}(2)$. We have, for $z \equiv x+iy \in \mathbb{C}$, $x, y \in \mathbb{R}$, that
$(zA)^\dagger = ((x+iy)A)^\dagger = xA^\dagger - iyA^\dagger = -xA + iyA$.
Hence for a purely imaginary scalar multiplication this implies
$(iA)^\dagger = iA$.
So if we define $B \equiv iA \in \mathfrak{su}(2)\otimes\mathbb{C}$ it follows that $\mathfrak{su}(2)\otimes\mathbb{C}$ is the extension of $\mathfrak{su}(2)$ with Hermitian matrices. Since every matrix $C \in \mathrm{Mat}(2,\mathbb{C})$ can be written as the sum of a Hermitian and anti-Hermitian matrix via
$C = \frac{1}{2}(C + C^\dagger) + \frac{1}{2}(C - C^\dagger)$
with $\left(\frac{1}{2}(C + C^\dagger)\right)^\dagger = \frac{1}{2}(C + C^\dagger)$
$\left(\frac{1}{2}(C - C^\dagger)\right)^\dagger = -\frac{1}{2}(C - C^\dagger)$,
it follows that $\mathfrak{su}(2)\otimes\mathbb{C} \cong \mathfrak{sl}(2,\mathbb{C})$ with

$$\mathfrak{sl}(2) \equiv \{C \in \mathrm{Mat}(2,\mathbb{C}) | \mathrm{tr}(C) = 0\}, \tag{104}$$

the Lie algebra of general 2 by 2 complex valued traceless matrices. Since this Lie algebra has complex dimension 3 it spanned by the set of matrices:

$$X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad E = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

By ordinary matrix multiplication it can be checked that

$$[X,Y] = E, \quad [X,E] = -2X, \quad [Y,E] = 2Y. \tag{105}$$

The Lie group $SU(2)$ acts on $\mathbb{C}^2$ by ordinary matrix multiplication. The associated representation π on the space $P(\mathbb{C})$ of polynomial functions $p : \mathbb{C} \to \mathbb{C}$ is given by the formula

$$\pi(g)p(z) \equiv p(g^{-1}z) \;\; g \in SU(2), z \in \mathbb{C}. \tag{106}$$

Indeed π defines a representation, since $\pi(g)$ is clearly a linear endomorphism and, for $g_1, g_2 \in SU(2)$, we have
$\pi(g_1g_2)p(z) = p((g_1g_2)^{-1}z) = p(g_2^{-1}g_1^{-1}z) = \pi(g_1)p(g_2^{-1}z) = \pi(g_1)\pi(g_2)p(z)$.
The subspace $P_n \equiv P_n(\mathbb{C}^2)$ of homogeneous polynomials of degree n is an invariant subspace for π. The restriction of π to P_n, denoted by π_n, is an irreducible representation of $SU(2)$. We denote the associated representation of $\mathfrak{su}(2)$ in $P_n(\mathbb{C}^2)$ by π_{n*}. Then the representation of $\mathfrak{sl}(2,\mathbb{C})$ in $P_n(\mathbb{C}^2)$ equals $(\pi_{n*})^{\mathbb{C}}$, the complexification of π_{n*}. The basis of $P_n(\mathbb{C}^2)$ is given by

$$p_j(z) = z_1^j z_2^{n-j}, \;\; (z \in \mathbb{C}^2, j \in \{0,1,...,n\}). \tag{107}$$

Let $p \in P_n(\mathbb{C}^2)$, $\xi \in \mathfrak{su}(2)$ and $t \in \mathbb{R}$. Then, using (106) it follows that,
$[\pi_{n*}(\xi)p](z) = \frac{d}{dt}p(\exp^{-t\xi} z)\big|_{t=0}$,
and hence by the chain rule,
$[\pi_{n*}(\xi)p](z) = \frac{\partial p}{\partial z_1}(z)(-\xi z)_1 + \frac{\partial p}{\partial z_2}(z)(-\xi z)_2..$
The expression on the right-hand side is complex linear in ξ, meaning that $\pi_{n*}(i\xi) = i\pi_*(\xi)$. This implies that the map $\pi_{n*} : \mathfrak{su}(2) \to P_n(\mathbb{C}^2)$ can be extended to its complexification $\pi_{n*}^{\mathbb{C}}$. Hence for $\xi \in \mathfrak{sl}(2,\mathbb{C})$ and $p \in P_n(\mathbb{C}^2)$,

$$(\pi_{n*}^{\mathbb{C}})(\xi)p = -\left[(\xi z)_1 \frac{\partial}{\partial z_1}] + (\xi z)_2 \frac{\partial}{\partial z_2}\right] p. \tag{108}$$

Now consider the base elements X, Y and E of $\mathfrak{sl}(2,\mathbb{C})$. Substituting them for ξ in (108) gives

$$Xp = -z_2 \frac{\partial}{\partial z_1} p, \;\; Yp = -z_1 \frac{\partial}{\partial z_2} p, \;\; Ep = \left[-z_1 \frac{\partial}{\partial z_1} + z_2 \frac{\partial}{\partial z_2}\right] p. \tag{109}$$

It follows that the base elements X, Y and E can be associated with the operators

$$X = -z_2\frac{\partial}{\partial z_1}, \;\; Y = -z_1\frac{\partial}{\partial z_2}, \;\; E = -z_1\frac{\partial}{\partial z_1} + z_2\frac{\partial}{\partial z_2}. \tag{110}$$

One can indeed check that these operators satisfy the relations (105) and hence form a Lie algebra representation of $\mathfrak{su}(2) \otimes \mathbb{C} \cong \mathfrak{sl}(2,\mathbb{C})$ in $P_n(\mathbb{C}^2)$. The idea of the infinitesimal method is to reduce the consideration of Lie groups to consideration of the associated (complexified) Lie algebras. From the fact that the operators in (105) form a Lie algebra representation of $\mathfrak{su}(2) \otimes \mathbb{C}$ it indeed follows that every irreducible continuous representation of $SU(2)$ is equivalent to π_n for some $n \in \mathbb{N}$. For the full rigourous proof see for example ([3]) or ([26]).

To conclude this subsection, let us consider a specific case of π_n. From the discussion above one already concludes that $\forall n \in \mathbb{N}$ there is a unique representation of dimension $n+1$ (there are $n+1$ polynomials $p_j(z)$ for $j \in \{0, 1, ..., n\}$). Take $n = 3$. Then the representation is of the form depicted in figure 5.

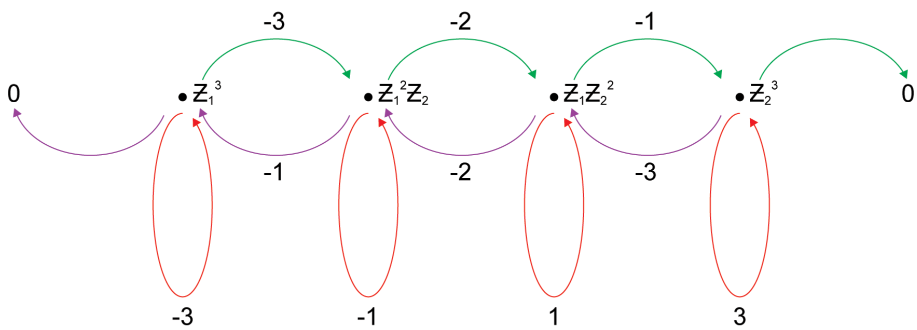

Figure 5: Schematic view of the representation π_3 of SU(2) in $P_3(\mathbb{C}^2)$. The points represent the 4 basis elements of $P_3(\mathbb{C}^2)$. The arrows represent the action of the operators X (green), Y (violet) and E (red) on the basis elements. The numbers above the green (violet) arrows denote the coefficients in front of the basis elements on the right (left) of the arrow when the corresponding operator has acted on the basis element on the left (right) of the arrow. The numbers above the red arrows denote the coefficients in front of the basis elements when the corresponding operator has acted on them.

Here z_1^3, $z_1^2 z_2$, $z_1 z_2^2$ and z_2^3 denote the basis elements of $P_3(\mathbb{C}^2)$. The arrows represent the action of the operators X, Y and E on the basis elements. The numbers above the arrows denote the coefficients in front of the basis elements after the action has taken place. As a check, we apply the operator

X to the basis elements of $P_3(\mathbb{C}^2)$. We indeed calculate,
$-z_2\frac{\partial}{\partial z_1}(z_1^3) = -3z_1^2 z_2,$
$- z_2\frac{\partial}{\partial z_1}(z_1^2 z_2) = -2z_1 z_2^2,$
$- z_2\frac{\partial}{\partial z_1}(z_1 z_2^2) = -z_2^3,$
$- z_2\frac{\partial}{\partial z_1}(z_2^3) = 0.$

7.4.2 Finding the irreducible representations of SU(2) by means of the orbit method

Consider the coadjoint orbits of $SU(2)$ satisfying the Integrality Condition, i.e. the spheres of half-integer radius. As we have derived in subsubsection 5.2.2, the stereographic coordinates z and $\bar{z}$ on U_- are given in terms of the coordinates θ and ϕ by

$$\begin{cases} z = 2R\cot(\theta/2)e^{i\phi}, \\ \bar{z} = 2R\cot(\theta/2)e^{-i\phi}, \\ 0 \leq \phi < 2\pi, \quad 0 \leq \theta < \pi. \end{cases}$$

Now consider the rescaled coordinates,

$$z_- = \frac{z}{2R} = \cot(\theta/2)e^{i\phi}, \qquad \overline{z_-} = \frac{\bar{z}}{2R} = \cot(\theta/2)e^{-i\phi}, \tag{111}$$

which span the complex plane $\mathbb{C}$ too. In terms of z_- and $\overline{z_-}$ a parametrization of U_- is given by

$$\begin{cases} x_1 = \frac{R(z_- + \overline{z_-})}{1+|z_-|^2}, \\ x_2 = \frac{-iR(z_- - \overline{z_-})}{1+|z_-|^2}, \\ x_3 = \frac{R(|z_-|^2-1)}{1+|z_-|^2}. \end{cases},$$

where $\{x_1, x_2, x_3\}$ are the standard coordinates in $\mathbb{R}^3$. Indeed, using (111),

$x_1 = \frac{R(z_- + \overline{z_-})}{1+|z_-|^2} = R\frac{\cot(\theta/2)(e^{i\phi}+e^{-i\phi})}{1+\cot^2(\theta/2)} = R\frac{2\cot(\theta/2)\cos(\phi)}{1+\cot^2(\theta/2)} = R\sin(\theta)\cos(\phi),$
$x_2 = \frac{-iR(z_- - \overline{z_-})}{1+|z_-|^2} = -iR\frac{\cot(\theta/2)(e^{i\phi}-e^{-i\phi})}{1+\cot^2(\theta/2)} = -iR\frac{2\cot(\theta/2)\sin(\phi)}{1+\cot^2(\theta/2)} = R\sin(\theta)\sin(\phi),$
$x_3 = \frac{R(|z_-|^2-1)}{1+|z_-|^2} = \frac{R(\cot^2(\theta/2)-1)}{1+\cot^2(\theta/2)} = R\cos(\theta),$

which is the parametrization of U_- in spherical coordinates (see subsubsection 5.2.1). Instead of taking x_1, x_2 and x_3 as the coordinates spanning $U_- \cong \mathbb{C}$, one can also consider the linear combinatinions $x_- = x_1 - ix_2$, $x_+ = x_1 + ix_2$ and x_3 as the coordinates spanning $U_- \cong \mathbb{C}$. In terms of z_- and $\overline{z_-}$ these equal,

$x_+ = x_1 + ix_2 = \left(\frac{R(z_- + \overline{z_-})}{1+|z_-|^2}\right) + i\left(\frac{-iR(z_- - \overline{z_-})}{1+|z_-|^2}\right) = \frac{2Rz_-}{1+|z_-|^2},$
$x_- = x_1 - ix_2 = \left(\frac{R(z_- + \overline{z_-})}{1+|z_-|^2}\right) - i\left(\frac{-iR(z_- - \overline{z_-})}{1+|z_-|^2}\right) = \frac{2R\overline{z_-}}{1+|z_-|^2},$
$x_3 = \frac{R(|z_-|^2-1)}{1+|z_-|^2}.$

The quantization assignment (103) can be rewritten in terms of z_- and $\overline{z_-}$, giving,

$$\begin{aligned} Q(f) &= -i\hbar\left\langle \frac{1}{4R}(1+|z_-|^2)^2\left(\frac{\partial f}{\partial \overline{z_-}}\frac{\partial}{\partial z_-} - \frac{\partial f}{\partial z_-}\frac{\partial}{\partial \overline{z_-}}\right) - \frac{\overline{z_-}}{2}(1+|z_-|^2)\frac{\partial f}{\partial \overline{z_-}}\right\rangle \\ &\quad + f \text{ for } f \in C^\infty(S^2, \mathbb{R}; P). \end{aligned} \tag{112}$$

Applying the assignment (112) to x_+, x_- and x_3 gives the operators

$Q(x_+) = (z_-)^2\frac{\partial}{\partial z_-} - 2Rz_-,$
$Q(x_-) = -\frac{\partial}{\partial z_-},$
$Q(x_3) = z_-\frac{\partial}{\partial z_-} - R.$

Note that here is used that $Q(f)$ acts on holomorphic sections, hence the $\frac{\partial}{\partial \overline{z_-}}$-terms can be dropped in the calculation. The determined operators $Q(x_+), Q(x_-), Q(x_3)$ satisfy the same characteristic equations

$$[r_1, r_2] = 2r_3, \;\; [r_1, r_3] = -2r_2, \;\; [r_2, r_3] = 2r_1, \tag{113}$$

as the Lie algebra generators r_1, r_2, r_3 of $\mathfrak{su}(2)$ (see subsubsection 4.2.2) . Indeed, for $s \in \mathcal{H}_P$,

$[Q(x_+), Q(x_-)]s = \left[(z_-)^2\frac{\partial}{\partial z_-} - 2Rz_-, -\frac{\partial}{\partial z_-}\right]s =$
$-\left((z_-)^2\frac{\partial}{\partial z_-} - 2Rz_-\right)\frac{\partial s}{\partial z_-} + \frac{\partial}{\partial z}\left((z_-)^2\frac{\partial}{\partial z_-} - 2Rz_-\right)s =$
$-z_-^2\frac{\partial^2 s}{\partial z_-^2} + 2Rz\frac{\partial s}{\partial z_-} + 2z_-\frac{\partial s}{\partial z_-} + z_-^2\frac{\partial^2 s}{\partial z_-^2} - 2Rs - 2Rz_-\frac{\partial s}{\partial z_-} =$
$2\left(z_-\frac{\partial}{\partial z_-} - R\right)s = 2Q(x_3)s \Leftrightarrow [Q(x_+), Q(x_-)] = 2Q(x_3).$

Similarly, one can check that $[Q(x_+), Q(x_3)] = -2Q(x_-)$ and $[Q(x_-), Q(x_3)] = 2Q(x_+)$. In subsection 7.2 we showed that for spheres of radius $k \in \frac{\mathbb{Z}}{2}$ the Hilbert space consists of polynomials in $z \in \mathbb{C}$ of degree $\leq k$, denoted by $P(\mathbb{C})$. The homogeneous polynomials in $z \in \mathbb{C}$ of degree $n \leq k$, denoted by $P_n(\mathbb{C})$, form an invariant subspace of $P(\mathbb{C})$. Consequently, $Q(x_+), Q(x_-)$ and $Q(x_3)$ form a representation of $\mathfrak{su}(2)$ in $P_n(\mathbb{C})$ for $n \leq R$. Finally, lifting the information from the Lie algebra to the Lie group (see [26]), this gives all the irreducible unitary representations of $SU(2)$.

We end this section by giving an alternative method to determine that the Hilbert space for spheres of half-integer radius $k \in \frac{\mathbb{Z}}{2}$ consists of polynomials of degree $\leq k$ compared the method described in subsection 7.2.

On U_- it holds that

$$z_- = \frac{(x_1 + ix_2)}{R - x_3}. \tag{114}$$

Indeed,

$\frac{x_1+ix_2}{1-x_3} = \left\langle \left(\frac{R(z_-+\overline{z_-})}{1+|z_-|^2}\right) + i\left(\frac{-iR(z_--\overline{z_-})}{1+|z_-|^2}\right)\right\rangle \Big/ \left\langle R - \frac{R(|z_-|^2-1)}{1+|z_-|^2}\right\rangle = \left\langle \frac{2R}{1+|z_-|^2}\right\rangle \Big/ \left\langle \frac{2R}{1+|z_-|^2}\right\rangle = z_-.$

Similarly, on U_+ it holds that

$$z_+ = \frac{(x_1 - ix_2)}{R + x_3}. \tag{115}$$

On the overlap of the patches $U_\pm$ we have $z_-z_+ = 1$, which follows from

$z_-z_+ = \left(\frac{(x_1+ix_2)}{R-x_3}\right)\left(\frac{(x_1-ix_2)}{R+x_3}\right) = \frac{x_1^2+x_2^2}{R^2-x_3^2} = \frac{x_1^2+x_2^2}{x_1^2+x_2^2} = 1.$

As the patches $U_\pm$ are topologically trivial, any line bundle is trivial when restricted to one of these patches. Thus, all we have to do is to give a prescription for gluing these trivial bundles together over, say, the equator. If one does this with the transition function $(z_-)^{-k} = (z_+)^k$ one obtains the line bundle L_k over a sphere of radius $k \in \frac{\mathbb{Z}}{2}$. To determine the global holomorphic sections of L_k in the trivialization $U_\pm$, we have to check which local holomorphic functions on $U_\pm$ can be patched together via the transition functions. Since holomorphic functions $f(z)$ can be written as a series $f(z) = \sum_{i=0}^{\infty} a_i z^i$ with $a_i \in \mathbb{C}$, it follows that a basis for holomorphic functions on U_+ is given by the monomials $(z_+)^l$ and on U_- by $(z_-)^m$ for l and m non-negative integers. We thus have to find the solution to the equation

$$(z_+)^l = (z_+)^k(z_-)^m \Leftrightarrow (z_+)^l = (z_+)^k(z_+)^{-m}, \tag{116}$$

where we used $z_+z_- = 1$. This equation has non-negative integer solutions for $l \leq k$ and $m \leq k$, implying that the Hilbert space consists of polynomials of degree $\leq k$.

8 Conclusions and outlook

-Life is infinitely stranger than anything which the mind of man could invent. We would not dare to conceive the things which are really merely commonplaces of existence.-
Sir Arthur Conan Doyle ([28])

In the first subsection of this section we will try to gain insight in the overall structure of this thesis and recall the important results derived previously. Finally, in the last subsection, we will consider were future research efforts should be directed.

8.1 Concluding remarks

This thesis consisted of three main parts, being

(1) Coadjoint orbits

(2) Geometric Quantization

(3) The irreducible unitary representations of $SU(2)$.

The parts (1) and (2) are the basic ingredients used in the Orbit Method, the method to determine all irreducible unitary representations of a Lie group. Part (3) is an application of the Orbit Method to the particular Lie group $SU(2)$.

(1):
In part (1) we introduced the notion of coadjoint orbit, the most important new mathematical object that has been brought into consideration in connection with the Orbit Method. As an application, we showed in two alternative ways that the coadjoint orbits of $SU(2)$ are spheres. Subsequently, we proved the beautiful non-trivial theorem that the coadjoint orbits of a matrix group posses symplectic structure (and mentioned that this result is still valid for general Lie groups). In particular, it followed that the coadjoint orbits of a matrix group are always even-dimensional. Then we derived the explicit shape of the Kirillov 2-form for the coadjoint orbits of SU(2) (the spheres) in both spherical and stereographic coordinates. The symplectic structure of the coadjoint orbits provided an important link between coadjoint orbits and geometric quantization, since quantization was as assignment applied to smooth real functions on symplectic manifolds (classical observables).

(2):
This brought us to part (2) of the thesis. Here we showed that Geometric Quantization, formulating a relationship between classical- and quantum mechanics in a geometric language, is a rigorous quantization scheme applicable to general curved manifolds. In particular, it is showed that the Hilbert space over a curved manifold (M, ω) does not consist of square integrable scalar functions on (M, ω), but instead of polarized, square integrable sections of a Hermitian line bundle-with-connection over (M, ω) with curvature $\Omega = \frac{\omega}{\hbar}$. Important roles in the construction of the Hilbert space are played by the Integrality Condition and the notion of Polarization. We proved the Integrality Condition to be both a sufficient and necessary condition for a Hermitian line bundle-with-connection over (M, ω) with curvature $\Omega = \frac{\omega}{\hbar}$ to exist. As an important specific type of polarization (a way of reducing the 'too big' prequantum Hilbert space) we considered the Kähler polarization, naturally defined on kähler manifolds. Finally, as an important explicit realization of geometric quantization, we showed the assignment (87) acting on sections in the Hilbert space $\mathcal{H}_P$, to satisfy all axioms (Q1)-(Q5), imposed on a quantization assignment.

(3):
In the third and last part we applied our geometric quantization construction to the spheres, the coadjoint orbits of $SU(2)$. We find that in that case the integrality condition translates into the spheres being of half-integer radius in order for smooth real functions defined on it to be quantizable. Furthermore, we showed that the Hilbert space on these spheres of half-integer radius consists of polynomials of degree $n \leq k$, for a sphere of radius $k \in \frac{\mathbb{Z}}{2}$. Subsequently, we determined the irreducible unitary representations of $SU(2)$ by the Infinitesimal Method. Finally, as an application of the Orbit Method we brought coadjoint orbits and geometric quantization together by determing all irreducible unitary representations of $SU(2)$ by means of the Orbit Method.

8.2 Outlook

Geometric Quantization is a quantization construction applicable to symplectic manifolds. Looking at the classical phase space as a symplectic manifold presupposes to take point particles as the fundamental building blocks of nature. In field theory the fundamental building blocks of nature are considered to be fields instead of point particles and therefore it is interesting to extend Geometric Quantization to fields, a first step towards making Quantum Field Theory mathematically rigorous. An introduction to such an approach can be found in, for example, ([7]). In string theory the fundamental building blocks of nature are one-dimensional extensions of point particles, being strings. Finding a rigorous quantization scheme corresponding to such a theory is another very interesting prospect.

As already indicated at the beginning of the previous section the representation theory of compact manifolds, such as $SU(2)$, was already well-known to mathematicians long before the Orbit Method was invented. In that sense, $SU(2)$ is not the most important application of the Orbit Method to consider. It is indeed shown that the Orbit Method is applicable to non-compact Lie groups as well, in case one utilizes less precise defined mathematics. At the moment one makes the notions 'mathematically precise' again, the results the theory predicts can be shown to be incorrect. As a consequence, a lot of research can still be done on the Orbit Method applied to non-compact Lie groups. Important examples of non-compact Lie groups to consider are $SL(2,\mathbb{R})$ and the famous Virasoro group (named after the physicist Miguel Angel Virasoro), which has important applications in conformal field theory and string theory. As a good start to investigate these applications of the Orbit Method, the reader is referred to ([1]) and ([27]).

A An identity at the heart of quantization

-Let us try to introduce a quantum Poisson Bracket which shall be the analogue of the classical one.-
P.A.M. Dirac ([12, p.86])

Let (M, ω) be a symplectic manifold. An essential identity in the quantization procedure is

$$[X_f, X_g] = X_{\{f,g\}}, \tag{41}$$

where $f, g \in C^\infty(M, \mathbb{R})$ and $X_f, X_g, X_{\{f,g\}}$ are Hamiltonian vector fields on M, that is, $i(X_f)\omega = 2\omega(X_f, \cdot) = -df$, $i(X_g)\omega = 2\omega(X_g, \cdot) = -dg$, $i(X_{\{f,g\}})\omega = 2\omega(X_{\{f,g\}}, \cdot) = -d\{f, g\}$. This identity will be proved in this appendix.

First of all, we will derive the form of the Hamiltonian vector fields in local coordinates. Take Y to be a vectorfield on M. Because X_f and Y are vector fields, in local coordinates they can be written as: $X_f = \xi^i \frac{\partial}{\partial x^i}$, $Y = \eta^i \frac{\partial}{\partial x^i}$. Consequently, in order to find the form of X_f in local coordinates we need to determine the coefficients ξ^i.

We have $2\omega(X_f, \cdot) = -df \Leftrightarrow 2\omega(X_f, Y) = -df(Y)$. Let us start with writing out the left-hand side of the last equality:

$$2\omega(X_f, Y) = 2\omega_{ij} dx^i \wedge dx^j \left(\xi^k \frac{\partial}{\partial x^k}\right)\left(\eta^l \frac{\partial}{\partial x^l}\right) = \omega_{ij}(\xi^i \eta^j - \xi^j \eta^i) = 2\omega_{ij}\xi^i\eta^j,$$

where the Einstein summation convention is used and the anti-symmetry of the wedge product and ω_{ij} are used in deriving the second and last equality, respectively. Writing out the right-hand side gives

$$-df(Y) = -\frac{\partial f}{\partial x^i} dx^i \left(\eta^j \frac{\partial}{\partial x^j}\right) = -\frac{\partial f}{\partial x^i}\eta^i.$$

So $2\omega(X_f, Y) = -df(Y) \Leftrightarrow \left(\frac{\partial f}{\partial x^j} + 2\omega_{ij}\xi^i\right)\eta^j = 0$. Because the η^j are arbitrary we thus have $\frac{\partial f}{\partial x^j} + 2\omega_{ij}\xi^i = 0 \Leftrightarrow \frac{\partial f}{\partial x^j} = -2\omega_{ij}\xi^i = 2\omega_{ji}\xi^i$. The fact that ω is a symplectic form gives that ω_{ji} is invertible, that is, $\omega^{kj}\omega_{ji} = \delta^k_i$. It follows that $\frac{\partial f}{\partial x^j} = 2\omega_{ji}\xi^i \Leftrightarrow \omega^{kj}\frac{\partial f}{\partial x^j} = 2\omega^{kj}\omega_{ji}\xi^i = 2\xi^k$. Hence, in local coordinates,

$$X_f = \frac{1}{2}\xi^k \frac{\partial}{\partial x^k} = \frac{1}{2}\omega^{kj}\frac{\partial f}{\partial x^j}\frac{\partial}{\partial x^k}. \tag{A.1}$$

The identity (41) relates the Lie algebras of Hamiltonina vector fields and functions on M. We already proved that $(C^\infty(M,\mathbb{R}),\{.,.\})$ has the structure of a Lie algebra. We can use identities (41) and (A.1) to show that Hamiltonian vector fields together with the Lie bracket for vector fields also form a Lie algebra. First of all, notice that the space of Hamiltonian vector fields is a linear space, where smooth real-valued functions play the role of 'scalars'. Furthermore, the Lie bracket for vector fields satisfies:

- bilinearity: for $a \in \mathbb{R}, h \in C^\infty(M,\mathbb{R})$, $[aX_f + X_g, X_h] = [X_{af+g}, X_h] = X_{\{af+g,h\}} = aX_{\{f,h\}} + X_{\{g,h\}} = a[X_f, X_h] + [X_g, X_h]$. Linearity in the second 'slot' of the commutator bracket can be derived similarly.
- anti-symmetry: $[X_f, X_g] = X_{\{f,g\}} = X_{-\{g,f\}} = -X_{\{g,f\}} = -[X_g, X_f]$.
- Jacobi identity: for $h \in C^\infty(M,\mathbb{R})$, $[X_f,[X_g,X_h]] + [X_g,[X_h,X_f]] + [X_h,[X_f,X_g]] = X_{\{f,\{g,h\}\}} + X_{\{g,\{h,f\}\}} + X_{\{h,\{f,g\}\}} = X_{\{f,\{g,h\}\}+\{g,\{h,f\}\}+\{h,\{f,g\}\}} = X_0 = 0$.

Moreover, regarding the map (Lie algebra homomorphism) $f \to X_f$ as an assignment of differential operators to functions, the identity (41) is also an illustration of the quantization paradigm (Dirac's quantum condition),

$$Poisson\ \ Brackets \to Commutators. \tag{A.2}$$

To prove identity (41) we will write out both sides of this equality in local coordinates and check that the corresponding expressions match. The left hand side equals

$$[X_f, X_g] = \tfrac{1}{4}\omega^{ij}\partial_i f \partial_j(\omega^{kl}\partial_k g \partial_l) - \tfrac{1}{4}\omega^{ij}\partial_i g \partial_j(\omega^{kl}\partial_k f \partial_l) = \\ \tfrac{1}{4}\omega^{ij}\partial_i f \partial_j(\omega^{kl}\partial_k g)\partial_l + \tfrac{1}{4}\omega^{ij}\partial_i f \omega^{kl}\partial_k g \partial_j \partial_l - \tfrac{1}{4}\omega^{ij}\partial_i g \partial_j(\omega^{kl}\partial_k f)\partial_l - \\ \tfrac{1}{4}\omega^{ij}\partial_i g \omega^{kl}\partial_k f \partial_j \partial_l.$$

By switching the indices i and k in the second term of the last expression (over both indices is summed) one easily sees that the second and fourth term in this expression cancel. The remaining expression can be written out further:

$$\tfrac{1}{4}\omega^{ij}\partial_i f \partial_j(\omega^{kl}\partial_k g)\partial_l - \tfrac{1}{4}\omega^{ij}\partial_i g \partial_j(\omega^{kl}\partial_k f)\partial_l = \\ \tfrac{1}{4}\omega^{ij}\partial_i f \omega^{kl}\partial_j\partial_k g \partial_l - \tfrac{1}{4}\omega^{ij}\partial_i g \omega^{kl}\partial_j \partial_k f \partial_l + \tfrac{1}{4}\omega^{ij}\partial_i f \partial_j \omega^{kl}\partial_k g \partial_l - \tfrac{1}{4}\omega^{ij}\partial_i g \partial_j \omega^{kl}\partial_k f \partial_l.$$

Writing out the right-hand side of identity (41) gives, using the definition of the Poisson bracket in local coordinates, that

$X_{\{f,g\}} = -\frac{1}{2}\omega^{ij}\partial_i\{f,g\}\partial_j = \frac{1}{4}\omega^{ij}\partial_i(\omega^{kl}\partial_k f\partial_l g)\partial_j =$
$\frac{1}{4}\omega^{ij}\partial_i\omega^{kl}\partial_k f\partial_l g\partial_j + \frac{1}{4}\omega^{ij}\omega^{kl}\partial_i\partial_k f\partial_l g\partial_j + \frac{1}{4}\omega^{ij}\omega^{kl}\partial_k f\partial_i\partial_l g\partial_j.$

If we apply the index interchange $j \to l$, $l \to i$, $i \to k$, $k \to j$ the second term of the last expression becomes $\frac{1}{4}\omega^{kl}\omega^{ji}\partial_k\partial_j f\partial_i g\partial_l = -\frac{1}{4}\omega^{kl}\omega^{ij}\partial_k\partial_j f\partial_i g\partial_l$, and if we apply the index interchange $j \to l$, $k \to i$, $i \to k$, $l \to j$ the third term of the last expression becomes $\frac{1}{4}\omega^{kl}\omega^{ij}\partial_k\partial_j g\partial_i f\partial_l$. Hence we see that the terms containing no derivatives of ω^{kl} match on the left- and right-hand side of identity (41). What is left to check to establish identity (41), is that

$$-\omega^{ij}\partial_i\omega^{kl}\partial_k f\partial_l g\partial_j + \omega^{ij}\partial_i f\partial_j\omega^{kl}\partial_k g\partial_l - \omega^{ij}\partial_i g\partial_j\omega^{kl}\partial_k f\partial_l = 0.$$

If we apply the index interchange $i \to k$, $k \to l$, $l \to j$, $j \to r$ to the second term of this expression, it becomes $\omega^{kr}\partial_r\omega^{lj}\partial_k f\partial_l g\partial_j$. If we apply the index interchange $i \to l$, $l \to j$, $j \to r$ to the third term of this expression, it becomes $-\omega^{lr}\partial_r\omega^{kj}\partial_k f\partial_l g\partial_j$. As a whole, the last equality then implies

$$-\omega^{ij}\partial_i\omega^{kl} + \omega^{kr}\partial_r\omega^{lj} - \omega^{lr}\partial_r\omega^{kj} = 0.$$

Using the anti-symmetrie of ω^{kr}, ω^{lr} and ω^{kj} this is equivalent to

$$\omega^{rj}\partial_r\omega^{kl} + \omega^{rk}\partial_r\omega^{lj} + \omega^{rl}\partial_r\omega^{jk} = 0. \tag{A.3}$$

The identity (A.3) is known as the Jacobi identity for Poisson forms and it will be proved in the remainder of this appendix.

Since (M,ω) is a symplectic manifold, ω is closed, that is $d\omega = 0$. In local coordinates this can be rewritten as

$$d(\omega_{ij}dx^i \wedge dx^j) = \partial_k\omega_{ij}dx^k \wedge dx^i \wedge dx^j = 0 \Leftrightarrow$$
$$\frac{1}{3}(\partial_k\omega_{ij}dx^k \wedge dx^i \wedge dx^j + \partial_i\omega_{jk}dx^i \wedge dx^j \wedge dx^k + \partial_j\omega_{ki}dx^j \wedge dx^k \wedge dx^i) = 0 \Leftrightarrow$$
$$\frac{1}{3}(\partial_k\omega_{ij}dx^i \wedge dx^j \wedge dx^k + \partial_i\omega_{jk}dx^i \wedge dx^j \wedge dx^k + \partial_j\omega_{ki}dx^i \wedge dx^j \wedge dx^k) = 0 \Leftrightarrow$$
$$\partial_k\omega_{ij} + \partial_i\omega_{jk} + \partial_j\omega_{ki} = 0.$$

The last identity is known as the Jacobi identity for symplectic forms . We will derive the Jacobi identity for symplectic forms from the Jacobi identity for Poisson forms, using that ω is a symplectic form, hence invertible. Let us start by multiplying the Jacobi identity for Poisson forms by ω_{sl}, using $\omega_{sl}\omega^{lk} = \delta_s^k$,

$$\omega^{kl}\partial_k\omega^{mn} + \omega^{kn}\partial_k\omega^{lm} + \omega^{km}\partial_k\omega^{nl} = 0 \Leftrightarrow$$
$$\omega_{sl}\omega^{kl}\partial_k\omega^{mn} + \omega_{sl}\omega^{kn}\partial_k\omega^{lm} + \omega_{sl}\omega^{km}\partial_k\omega^{nl} = 0 \Leftrightarrow$$

$$-\partial_s\omega^{mn} + \omega_{sl}\omega^{kn}\partial_k\omega^{lm} + \omega_{sl}\omega^{km}\partial_k\omega^{nl} = 0$$

Using differentiation by parts and the fact that the boundary terms, $\partial_k(\omega_{sl}\omega^{lm}) = \partial_k\delta_s^m$, vanish, we get

$$-\partial_s\omega^{mn} - \omega^{kn}\partial_k\omega_{sl}\omega^{lm} - \omega^{km}\partial_k\omega_{sl}\omega^{nl} = 0$$

We can write the first term of this expression as $\omega^{mr}\partial_s\omega_{rl}\omega^{ln}$. This follows from $\partial_s(\omega^{mn}\omega_{nl}) = 0 \Leftrightarrow \partial_s\omega^{mn}\omega_{nl} + \omega^{mn}\partial_s\omega_{nl} = 0 \Leftrightarrow \partial_s\omega^{mn} + \omega^{mr}\partial_s\omega_{rl}\omega^{ln} = 0 \Leftrightarrow -\partial_s\omega^{mn} = \omega^{mr}\partial_s\omega_{rl}\omega^{ln}$. Multiplying the resulting expression by ω_{pm} and ω_{nq} respectively, gives

$$\omega^{mr}\partial_s\omega_{rl}\omega^{ln} - \omega^{kn}\partial_k\omega_{sl}\omega^{lm} - \omega^{km}\partial_k\omega_{sl}\omega^{nl} = 0 \Leftrightarrow$$
$$\partial_s\omega_{pl}\omega^{ln} - \omega^{nk}\partial_k\omega_{sp} + \partial_p\omega_{sl}\omega^{nl} = 0 \Leftrightarrow$$
$$\partial_s\omega_{pq} + \partial_q\omega_{sp} + \partial_p\omega_{qs} = 0,$$

which is indeed the Jacobi identity for symplectic forms! Identity (41) is proved.

References

[1] A.A. Kirillov, *Lectures on the Orbit Method*, American Mathematical Society, volume 64, 2000.

[2] A.A. Kirillov, *Elements of the Theory of Representations*, Springer-Verlag Berlin Heidelberg New York, 1976.

[3] E.P. van den Ban, *Lie groups*, Lecture notes, Spring 2005.

[4] *Topics in Representation Theory: The Moment Map and the Orbit Method*, `http://www.math.columbia.edu/~woit/notes23.pdf`.

[5] *The Orbit-Stabilizer Theorem*, `http://www.math.wisc.edu/~virk/notes/pre08/pdf/orbit-stabilizer_thm.pdf`.

[6] S.J.G. Vandoren, *Lectures on Riemannian Geometry, Part II: Complex Manifolds*, Utrecht, 2008.

[7] N.M.J. Woodhouse, *Geometric Quantization*, Oxford University Press New York, 1980.

[8] N.M.J. Woodhouse, D.J.Simms, *Lectures on Geometric Quantization*, Springer-Verlag Berlin Heidelberg New York, 1976.

[9] M. Blau, *Symplectic Geometry and Geometric Quantization*, NIKHEF-H, Amsterdam.

[10] E.J.N. Looijenga, *Smooth manifolds*, Notes to a course by Eduard Looijenga, fall 2008.

[11] *The 10 weirdest physics facts, from relativity to quantum physics*, `http://bit.ly/3C5k2J`.

[12] P.A.M. Dirac, *The principles of Quantum Mechanics (Fourth Edition)*, Clarendon Press, Oxford, 1958.

[13] *Quantum Theory: Paul Dirac*, `http://www.spaceandmotion.com/quantum-theory-paul-dirac-quotes.htm`.

[14] P.A.M. Dirac, *Proc. Roy. Soc. London ser. A*, **109**, 642-53, 1926.

[15] *BrainyQuote*, `http://bit.ly/ga5UXv`.

[16] G.E. Bredon, *Topology and geometry*, Springer-Verlag New York, 1993.

[17] I. Newton, *Principia mathematica*, Prometheus Books, New York, 1995.

[18] J.M. Souriau, *Quantification Géométrique*, Commun. Math. Phys. 1, 1966, 374; *Structure des Systèmes Dynamiques*, Dunod, Paris, 1970.

[19] B. Kostant, *Quantization and Unitary Representations*, in *Lecture Notes in Mathematics*, Vol. 170, Springer, Berlin, 1970.

[20] R. Abraham and J.E. Marsden, *Foundations of Mechanics*, Second Edition, Benjamin, Reading MA, 1978.

[21] S. Kobayashi and K. Nomizu, *Foundations of differential geometry*, Vol. 1, Wiley, New York, 1963.

[22] A. Weil, *Variétés Kählériennes*, Hermann, Paris, 1958.

[23] T. Frankel, *The geometry of physics*, Cambridge University Press, Second Edition, 2004.

[24] J. Sniatycki, *Geometric Quantization and Quantum Mechanics*, Springer, New York, 1980.

[25] *BrainyQuote*, `http://bit.ly/hkIBvq`.

[26] D.P. Želobenko, *Compact Lie groups and their representations*, American Mathematical Society, Providence, Rhode Island, 1973.

[27] E. Witten, *Coadjoint Orbits of the Virasoro Group*, Joseph Henry Laboratories, Princeton University, Princeton, NJ 08544, USA.

[28] *Englishforums*, `http://bit.ly/hQAZTl`.

[29] J.J. Sakurai, *Modern Quantum Mechanics*, Addison-Wesley Publishing Company, Los Angeles, 1994.

[30] *Lecture 3: Tangent Vectors and the Tangent Space*, `http://www.zweigmedia.com/diff_geom/Sec3.html`.

[31] *Topics in Representation Theory: The Killing Form, Reflections and Classification of Root Systems*, http://www.math.columbia.edu/~woit/notes9.pdf.

[32] *COMPACT LIE ALGEBRAS*, http://www.pas.rochester.edu/~rajeev/phy406/Symmetries9.pdf

Index