University of Crete
Department of Mathematics

Notes on Symplectic Geometry

Konstantin Athanassopoulos

Iraklion, 2015

These notes were written primarily for the author's own personal use. Some parts had been presented in a series of lectures in the Geometry Seminar of the Department of Mathematics of the University of Crete during the academic year 2006-07. The purpose of the lectures was to present the field of Symplectic Geometry/Topology to interested faculty members and graduate students, as it was relatively unknown to greek universities at that time, and to some extend still is.

Contents

Chapter 1

Introduction

1.1 Newtonian mechanics

In newtonian mechanics the state of a mechanical system is described by a finite number of real parameters. The set of all possible positions, of a material point for example, is a finite dimensional smooth manifold M, called the *configuration space*. A motion of the system is a smooth curve $\gamma : I \to M$, where $I \subset \mathbb{R}$ is an open interval. The velocity field of γ is smooth curve $\dot{\gamma} : I \to TM$. The (total space of the) tangent bundle TM of M is called the *phase space*.

According to Newton, the total force is a vector field F that acts on the points of the configuration space. Locally, a motion is a solution of the second order differential equation $F = m\ddot{x}$, where m is the mass. Equivalently, $\dot{\gamma}$ is locally a solution of the first order differential equation

$$\begin{pmatrix} \dot{x} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} v \\ \frac{1}{m} F(x, v, t) \end{pmatrix}.$$

Consider a system of N particles in $\mathbb{R}^3$ subject to some forces. If x_i denotes the position of the i-th particle then the configuration space is $(\mathbb{R}^3)^N$ and Newton's law of motion is

$$m_i \frac{d^2 x_i}{dt^2} = F_i(x_1, ..., x_N, \dot{x}_1, ..., \dot{x}_N, t), \quad 1 \leq i \leq N,$$

where m_i is the mass and F_i is the force on the i-th particle. Relabeling the variables setting q^{3i}, q^{3i+1} and q^{3i+2} the coordinates of x_i in this order, the configuration space becomes $\mathbb{R}^n$, $n = 3N$, and the equations of motion take the form

$$m_j \frac{d^2 q^j}{dt^2} = F_j(q^1, ..., q^n, \dot{q}^1, ..., \dot{q}^n, t), \quad 1 \leq j \leq n$$

Suppose that the forces do not depend on time and are conservative. This means that there is a smooth function $V : \mathbb{R}^n \to \mathbb{R}$ such that

$$F_j(q^1, ..., q^n, \dot{q}^1, ..., \dot{q}^n) = -\frac{\partial V}{\partial q^j}, \quad 1 \leq j \leq n.$$

For instantce, this is the case if N particles interact by gravitational attraction. Rewritting Newton's law as a system of first oder ordinary differential equations

$$\frac{dq^j}{dt} = v^j, \quad m_j \frac{d^2 v^j}{dt^2} = -\frac{\partial V}{\partial q^j}, \quad 1 \leq j \leq n,$$

or changing coordinates to $p_j = m_j v^j$ we have

$$\frac{dq^j}{dt} = \frac{1}{m_j} p_j, \quad \frac{d^2 p_j}{dt^2} = -\frac{\partial V}{\partial q^j}, \quad 1 \leq j \leq n,$$

The solutions of the above system of ordinary differential equations are the integral curves of the smooth vector field

$$X = \sum_{j=1}^{n} \frac{1}{m_j} p_j \frac{\partial}{\partial q^j} - \sum_{j=1}^{n} \frac{\partial V}{\partial q^j} \frac{\partial}{\partial p_j}.$$

Note that the smooth function

$$H(q^1, ..., q^n, p_1, ..., p_n) = \sum_{j=1}^{n} \frac{1}{2m_j} p_j^2 + V(q^1, ..., q^n)$$

is constant along solutions, because

$$dH = \sum_{j=1}^{n} \frac{\partial V}{\partial q^j} dq^j + \sum_{j=1}^{n} \frac{1}{m_j} p_j dp_j$$

and so $dH(X) = 0$. Actually, H completely determines X in the following sense. Let

$$\omega = \sum_{j=1}^{n} dq^j \wedge dp_j.$$

Then,

$$i_X \omega = \sum_{j=1}^{n} dq^j(X) dp_j - \sum_{j=1}^{n} dp_j(X) dq^j = \sum_{j=1}^{n} \frac{1}{m_j} p_j dp_j + \sum_{j=1}^{n} \frac{\partial V}{\partial q^j} dq^j = dH.$$

The smooth 2-form ω is closed and non-degenerate. The latter means that given any smooth 1-form η the equation $i_Y \omega = \eta$ has a unique solution Y. Indeed, for any smooth vector field Y we have

$$Y = \sum_{j=1}^{n} \omega(Y, \frac{\partial}{\partial p_j}) \frac{\partial}{\partial q^j} - \sum_{j=1}^{n} \omega(Y, \frac{\partial}{\partial q^j}) \frac{\partial}{\partial p_j}$$

and so $i_Y \omega = 0$ if and only if $Y = 0$.

Returning to Newtonian mechanics, we give the following definition.

Definition 1.1. An (autonomous) newtonian mechamical system is a triple (M, g, X), where M is a smooth manifold, g is a Riemannian metric on M and X is a smooth vector field on TM such that $\pi_* X = id$. A motion of (M, g, X) is a smooth curve $\gamma : I \to M$ such that $\dot{\gamma} : I \to TM$ is an integral curve of X. The smooth function $T : TM \to \mathbb{R}$ defined by $T(v) = \dfrac{1}{2} g(v, v)$ is called the kinetic energy.

Examples 1.2. (a) Let $M = \mathbb{R}$, so that we may identify TM with $\mathbb{R}^2$ and π with the projection onto the first coordinate. If g is the euclidean riemannian metric on $\mathbb{R}$ and

$$X = v\frac{\partial}{\partial x} + \frac{1}{m}(-k^2 x - \rho v)\frac{\partial}{\partial v}, \qquad k > 0, \quad \rho \geq 0,$$

then obviously $\pi_* X = id$ and a motion is a solution of the second order differential equation

$$m\ddot{x} = -k^2 x - \rho\dot{x}.$$

This mechanical system describes the oscillator.

(b) The geodesic vector field G of a Riemannian n-manifold (M, g) defines a newtonian mechanical system. Locally it has the expression

$$G = \sum_{k=1}^{n} v^k \frac{\partial}{\partial x^k} - \sum_{i,j,k=1}^{n} \Gamma_{ij}^k v^i v^j \frac{\partial}{\partial v^k}$$

where Γ_{ij}^k are the Christofell symbols.

(c) The motion of a particle of unit mass on the unit circle S^1 under the influence of a vertical downward unit force is governed by Newton's law which states

$$\ddot{x} = -\sin x$$

or equivalently

$$\begin{pmatrix} \dot{x} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} v \\ -\sin x \end{pmatrix}, \qquad x \bmod 2\pi.$$

Here the phase space is the tangent bundle $TS^1 \cong S^1 \times \mathbb{R}$ and the corresponding vector field can be lifted on the universal covering space $\mathbb{R} \times \mathbb{R}$ to the smooth vector field

$$X(x, v) = v\frac{\partial}{\partial x} - \sin x \cdot \frac{\partial}{\partial v},$$

which is invariant under horizontal translations by integer multiples of 2π. The force acting is conservative with potential $V(x) = -\cos x$. The mechanical energy

$$H(x, v) = \frac{1}{2}v^2 - \cos x$$

is a constant of motion.

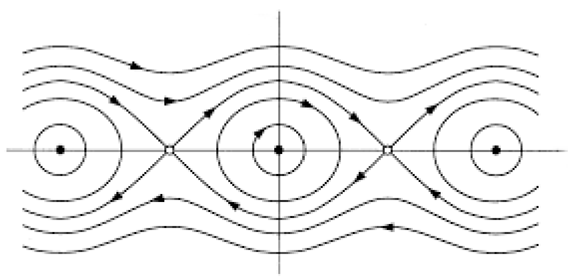
We observe that H is a Morse function. Indeed, the critical points of H are the points $(n\pi, 0)$, $n \in \mathbb{Z}$, where $H(2k\pi, 0) = -1$ and $H((2k+1)\pi, 0) = 1$ for every $k \in \mathbb{Z}$. Moreover, the Hessian at the critical points is

$$D^2 H(n\pi, 0) = \begin{pmatrix} \cos n\pi & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} (-1)^n & 0 \\ 0 & 1 \end{pmatrix}.$$

Thus, $(n\pi, 0)$ is a non-degenerate critical point of H of Morse index 0 if n is even and of Morse index 1 if n is odd. By the Morse Lemma, for every $k \in \mathbb{Z}$ there exist an open neighbourhood $U_k \subset ((2k-1)\pi, (2k+1)\pi) \times \mathbb{R}$ of $(2k\pi, 0)$ and $\epsilon > 0$ such that $H^{-1}(c) \cap U_k$ is a smooth simple close curve for $-1 < \epsilon < -1 + \epsilon$. Since there are no critical values in the interval $(-1, 1)$, for every $-1 < c < 1$ the level set $H^{-1}(c) \cap U_k$ is a smooth simple closed. The level set $H^{-1}(c)$ is a countable union of smooth simple closed curves for $|c| < 1$. If $c \geq 1$, then

$$H^{-1}(c) = \{(x, \sqrt{2(c + \cos x)}) : x \in \mathbb{R}\} \cup \{(x, -\sqrt{2(c + \cos x)}) : x \in \mathbb{R}\}.$$

The two graphs are disjoint if $c > 1$. If $c = 1$, they intersect transversally at the points $((2k+1)\pi, 0)$, $k \in \mathbb{Z}$, by the Morse Lemma. The phase portrait of X is depicted in the following picture.



Often a mechanical system has potential energy. This is a smooth function $V : M \to \mathbb{R}$. Let grad$V$ be the gradient of V with respect to the Riemannian metric. If for every $v \in TM$ we set

$$\overline{\mathrm{grad}V} = \frac{d}{dt}\bigg|_{t=0} (v + t\mathrm{grad}V(\pi(v)),$$

then $\overline{\mathrm{grad}V} \in \mathcal{X}(TM)$ and $\pi_*\overline{\mathrm{grad}V} = 0$, since $\pi(v + t\mathrm{grad}V(\pi(v))) = \pi(v)$, for every $t \in \mathbb{R}$. Locally, if $g = (g_{ij})$ and $(g_{ij})^{-1} = (g^{ij})$, then

$$\mathrm{grad}V = \sum_{i,j=1}^{n} g^{ij} \frac{\partial V}{\partial x^i} \frac{\partial}{\partial x^j} \qquad \text{and} \qquad \overline{\mathrm{grad}V} = \sum_{i,j=1}^{n} g^{ij} \frac{\partial V}{\partial x^i} \frac{\partial}{\partial v^j}.$$

Definition 1.3. A newtonian mechanical system with potential energy is a triple (M, g, V), where (M, g) is a Riemannian manifold and $V : M \to \mathbb{R}$ is a smooth function called the potential energy.

The corresponding vector field on TM is $Y = G - \overline{\mathrm{grad}V}$, where G is the geodesic vector field. The smooth function $E = T + V \circ \pi : TM \to \mathbb{R}$ is called the mechanical energy.

Proposition 1.4. (Conservation of energy) *In a newtonian mechanical system with potential energy (M, g, V) the mechanical energy is a constant of motion.*

Proof. We want to show that $Y(E) = 0$. We compute locally, where we have

$$E = \frac{1}{2}\sum_{i,j=1}^{n} g_{ij}v^i v^j + V \qquad \text{and}$$

$$Y = \sum_{k=1}^{n} v^k \frac{\partial}{\partial x^k} - \sum_{k=1}^{n}\Big(\sum_{i,j=1}^{n}\Gamma_{ij}^k v^i v^j + \sum_{i=1}^{n}\frac{\partial V}{\partial x^i}g^{ik}\Big)\frac{\partial}{\partial v^k}.$$

Recall that

$$\Gamma_{ij}^k = \frac{1}{2}\sum_{l=1}^{n} g^{kl}\big(\frac{\partial g_{il}}{\partial x^j} + \frac{\partial g_{jl}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^l}\big).$$

We can now compute

$$Y(E) = \sum_{k=1}^{n} v^k\frac{\partial V}{\partial x^k} + \frac{1}{2}\sum_{i,j,k=1}^{n}\frac{\partial g_{ij}}{\partial x^k}v^i v^j v^k - \sum_{k=1}^{n}\Big(\sum_{i,j=1}^{n}\Gamma_{ij}^k v^i v^j + \sum_{i=1}^{n}\frac{\partial V}{\partial x^i}g^{ik}\Big)\frac{\partial V}{\partial v^k}$$

$$-\sum_{k=1}^{n}\Big(\sum_{i,j=1}^{n}\Gamma_{ij}^k v^i v^j + \sum_{i=1}^{n}\frac{\partial V}{\partial x^i}g^{ik}\Big)\Big(\sum_{i=1}^{n} g_{ik}v^i\Big)$$

$$= \sum_{k=1}^{n} v^k\frac{\partial V}{\partial x^k} - \Big(\sum_{i=1}^{n}\frac{\partial V}{\partial x^i}g^{ik}\Big)\Big(\sum_{i=1}^{n} g_{ik}v^i\Big) + \frac{1}{2}\sum_{i,j,k=1}^{n}\frac{\partial g_{ij}}{\partial x^k}v^i v^j v^k$$

$$-\sum_{k=1}^{n}\Big(\sum_{r=1}^{n} g_{rk}v^r\Big)\Big(\sum_{i,j=1}^{n}\frac{1}{2}\sum_{l=1}^{n} g^{kl}\big(\frac{\partial g_{il}}{\partial x^j} + \frac{\partial g_{jl}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^l}\big)v^i v^j\Big)$$

$$= \frac{1}{2}\sum_{i,j,k=1}^{n}\frac{\partial g_{ij}}{\partial x^k}v^i v^j v^k - \frac{1}{2}\sum_{i,j,l=1}^{n}\big(\frac{\partial g_{il}}{\partial x^j} + \frac{\partial g_{jl}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^l}\big)v^i v^j v^l = 0. \qquad \square$$

A smooth curve $\gamma : I \to M$ is a motion of a newtonian mechanical system on M with potential energy V if and only if γ satisfies the second order differential equation

$$\nabla_{\dot\gamma}\dot\gamma = -\text{grad}V$$

where ∇ is the Levi-Civita connection on M. In case V has an upper bound, then a motion is a geodesic with respect to a new Riemannian metric on M, possibly reparametrized. So, suppose that there exists some $e > 0$ such that $V(x) < e$ for every $x \in M$. On M we consider the new Riemannian metric $g^* = (e - V)g$. Let γ be a montion with mechanical energy e, that is

$$\frac{1}{2}g(\dot\gamma(t), \dot\gamma(t)) + V(\gamma(t)) = e$$

for every $t \in I$. Since $e > V(\gamma(t))$, we have $\dot{\gamma}(t) \neq 0$ for every $t \in I$. The function $s : I \to \mathbb{R}$ with

$$s(t) = \sqrt{2} \int_{t_0}^{t} (e - V(\gamma(\tau))d\tau,$$

where $t_0 \in I$, is smooth and strictly increasing. Let $\gamma^* = \gamma \circ s^{-1}$.

Theorem 1.5. (Jacobi-Maupertuis) *If γ is a motion of the mechanical system (M, g, V) and $V(x) < e$ for every $x \in M$, then its reparametrization γ^* is a geodesic with respect to the Riemannian metric $g^* = (e - V)g$.*

Proof. It suffices to carry out the computation locally. The Christofell symbols of the metric g^* are given by the formula

$$\Delta_{ij}^k = \Gamma_{ij}^k + \frac{1}{2(e-V)} \left(-\frac{\partial V}{\partial x^i} \delta_{jk} - \frac{\partial V}{\partial x^j} \delta_{ik} + \sum_{l=1}^{n} \frac{\partial V}{\partial x^l} g^{lk} g_{ij} \right).$$

If in the local coordinates we have $\gamma = (x^1, x^2, ..., x^n)$, then

$$\frac{dx^k}{ds} = \frac{dx^k}{dt} \cdot \frac{dt}{ds} = \frac{1}{\sqrt{2}(e-V)} \cdot \frac{dx^k}{dt}$$

and so

$$\frac{d^2x^k}{ds^2} = \frac{1}{2(e-V)^2} \cdot \frac{d^2x^k}{d^2t} + \frac{1}{2(e-V)^3} \cdot \frac{dx^k}{dt} \sum_{l=1}^{n} \frac{\partial V}{\partial x^l} \frac{dx^l}{dt}.$$

Since

$$\frac{d^2x^k}{d^2t} = -\sum_{i,j=1}^{n} \Gamma_{ij}^k \frac{dx^i}{dt} \frac{dx^j}{dt} - \sum_{l=1}^{n} \frac{\partial V}{\partial x^l} g^{kl}$$

and

$$\sum_{i,j=1}^{n} g_{ij} \frac{dx^i}{dt} \frac{dx^j}{dt} = 2(e-V)$$

substituting we get

$$\frac{d^2x^k}{ds^2} + \sum_{i,j=1}^{n} \Delta_{ij}^k \frac{dx^i}{ds} \frac{dx^j}{ds} =$$

$$-\frac{1}{2(e-V)^2} \sum_{i,j=1}^{n} \Gamma_{ij}^k \frac{dx^i}{dt} \frac{dx^j}{dt} - \frac{1}{2(e-V)^2} \sum_{l=1}^{n} \frac{\partial V}{\partial x^l} g^{kl} + \frac{1}{2(e-V)^3} \frac{dx^k}{dt} \sum_{l=1}^{n} \frac{\partial V}{\partial x^l} \frac{dx^l}{dt}$$

$$+\frac{1}{2(e-V)^2} \sum_{i,j=1}^{n} \Gamma_{ij}^k \frac{dx^i}{dt} \frac{dx^j}{dt} - \frac{1}{4(e-V)^3} \frac{dx^k}{dt} \sum_{i=1}^{n} \frac{\partial V}{\partial x^i} \frac{dx^i}{dt} - \frac{1}{4(e-V)^3} \frac{dx^k}{dt} \sum_{j=1}^{n} \frac{\partial V}{\partial x^j} \frac{dx^j}{dt}$$

$$+\frac{1}{4(e-V)^3} \left(\sum_{l=1}^{n} \frac{\partial V}{\partial x^l} g^{kl} \right) \left(\sum_{i,j=1}^{n} g_{ij} \frac{dx^i}{dt} \frac{dx^j}{dt} \right) =$$

$$-\frac{1}{2(e-V)^2} \sum_{l=1}^{n} \frac{\partial V}{\partial x^l} g^{kl} + \frac{1}{4(e-V)^3} \left(\sum_{l=1}^{n} \frac{\partial V}{\partial x^l} g^{kl} \right) \left(\sum_{i,j=1}^{n} g_{ij} \frac{dx^i}{dt} \frac{dx^j}{dt} \right) = 0. \qquad \square$$

1.2 Lagrangian mechanics

Let (M, g, V) be a newtonian mechanical system with potential energy V and let $L : TM \to \mathbb{R}$ be the smooth function $L = T - V \circ \pi$, where T is the kinetic energy and $\pi : TM \to M$ is the tangent bundle projection.

Theorem 2.1. (d'Alembert-Lagrange) *A smooth curve $\gamma : I \to M$ is a motion of the mechanical system (M, g, V) if and only if*

$$\frac{d}{dt}\Big(\frac{\partial L}{\partial v^i}(\dot\gamma(t))\Big) = \frac{\partial L}{\partial x^i}(\dot\gamma(t))$$

for every $t \in I$ and $i = 1, 2..., n$, where n is the dimension of M.

Proof. Suppose that in local coordinates we have $\gamma = (x^1, x^2, ..., x^n)$. Recall that γ is a motion of (M, g, V) if and only if

$$\ddot x^k = -\sum_{i,j=1}^{n} \Gamma^k_{ij}\dot x^i \dot x^j - \sum_{l=1}^{n} \frac{\partial V}{\partial x^l} g^{lk}.$$

Since

$$L(\dot\gamma) = \frac{1}{2}\sum_{i,j=1}^{n} g_{ij}\dot x^i \dot x^j - V(\gamma),$$

for every $i = 1, 2, ..., n$ we have

$$\frac{d}{dt}\Big(\frac{\partial L}{\partial v^i}(\dot\gamma(t))\Big) - \frac{\partial L}{\partial x^i}(\dot\gamma(t)) = \frac{d}{dt}\left(\sum_{j=1}^{n} g_{ij}\dot x^j\right) - \frac{1}{2}\sum_{m,l=1}^{n}\frac{\partial g_{ml}}{\partial x^i}\dot x^m \dot x^l + \frac{\partial V}{\partial x^i}(\gamma(t)) =$$

$$\sum_{j=1}^{n}\sum_{l=1}^{n}\frac{\partial g_{ij}}{\partial x^l}\dot x^l \dot x^j + \sum_{j=1}^{n} g_{ij}\ddot x^j - \frac{1}{2}\sum_{m,l=1}^{n}\frac{\partial g_{ml}}{\partial x^i}\dot x^m \dot x^l + \frac{\partial V}{\partial x^i}(\gamma(t)) =$$

$$\sum_{m,l=1}^{n}\left(\frac{\partial g_{im}}{\partial x^l} - \frac{1}{2}\frac{\partial g_{ml}}{\partial x^i}\right)\dot x^m \dot x^l + \sum_{j=1}^{n} g_{ij}\ddot x^j + \frac{\partial V}{\partial x^i}(\gamma(t)).$$

Taking the image of the vector with these coordinates by $(g_{ij})^{-1} = (g^{ij})$, we see that the equations in the statement of the theorem are equivalent to

$$0 = \sum_{i=1}^{n} g^{ik}\left(\sum_{m,l=1}^{n}\left(\frac{\partial g_{im}}{\partial x^l} - \frac{1}{2}\frac{\partial g_{ml}}{\partial x^i}\right)\dot x^m \dot x^l\right) + \sum_{j=1}^{n} g^{ik} g_{ij}\ddot x^j + \sum_{i=1}^{n}\frac{\partial V}{\partial x^i} g^{ik} =$$

$$\ddot x^k + \sum_{m,l=1}^{n}\sum_{i=1}^{n} g^{ik}\left(\frac{\partial g_{im}}{\partial x^l} - \frac{1}{2}\frac{\partial g_{ml}}{\partial x^i}\right)\dot x^m \dot x^l + \sum_{i=1}^{n}\frac{\partial V}{\partial x^i} g^{ik} =$$

$$\ddot x^k + \sum_{m,l=1}^{n}\Gamma^k_{ml}\dot x^m \dot x^l + \sum_{i=1}^{n}\frac{\partial V}{\partial x^i} g^{ik}. \qquad \square$$

Generalizing we give the following definition.

Definition 2.2. An autonomous Lagrangian system is a couple (M, L), where M is a smooth manifold and $L : TM \to \mathbb{R}$ is a smooth function, called the Lagrangian. A Lagrangian motion is a smooth curve $\gamma : I \to M$ which locally satisfies the system of differential equations

$$\frac{d}{dt}\Big(\frac{\partial L}{\partial v^i}(\dot\gamma(t))\Big) = \frac{\partial L}{\partial x^i}(\dot\gamma(t))$$

for $t \in I$ and $i = 1,\ 2...,\ n$, where n is the dimension of M. These equations are called the Euler-Lagrange equations.

The variational interpretation of the Euler-Lagrange equations is given by the Least Action Principle due to Hamilton.

Theorem 2.3. (Least Action Principle) *Let (M, L) be a Lagrangian system. A smooth curve $\gamma : [a, b] \to M$ is a Lagrangian motion if and only if for every smooth variation $\Gamma : (-\epsilon, \epsilon) \times [a, b] \to M$ of γ with fixed endpoints, so that $\Gamma(0, t) = \gamma(t)$ for $a \le t \le b$, we have*

$$\frac{\partial}{\partial s}\Big|_{s=0}\int_a^b L\Big(\frac{\partial \Gamma}{\partial t}(s,t)\Big)dt = 0.$$

Proof. It suffices to carry out the computations locally. We have

$$\frac{\partial}{\partial s}\Big|_{s=0}\int_a^b L\Big(\frac{\partial \Gamma}{\partial t}(s,t)\Big)dt = \int_a^b \frac{\partial}{\partial s}\Big|_{s=0} L\Big(\frac{\partial \Gamma}{\partial t}(s,t)\Big)dt =$$

$$\int_a^b \left[\sum_{i=1}^n \frac{\partial L}{\partial x^i}(\dot\gamma(t))\frac{\partial \Gamma}{\partial s}(0,t) + \sum_{i=1}^n \frac{\partial L}{\partial v^i}(\dot\gamma(t))\frac{\partial}{\partial s}\Big|_{s=0}\Big(\frac{\partial \Gamma}{\partial t}(s,t)\Big)\right]dt =$$

$$\int_a^b \left[\sum_{i=1}^n \frac{\partial L}{\partial x^i}(\dot\gamma(t))\frac{\partial \Gamma}{\partial s}(0,t) + \sum_{i=1}^n \frac{\partial L}{\partial v^i}(\dot\gamma(t))\frac{\partial}{\partial t}\Big(\frac{\partial \Gamma}{\partial s}(0,t)\Big)\right]dt =$$

$$\int_a^b \left[\sum_{i=1}^n \frac{\partial L}{\partial x^i}(\dot\gamma(t))\frac{\partial \Gamma}{\partial s}(0,t) + \frac{d}{dt}\left(\sum_{i=1}^n \frac{\partial L}{\partial v^i}(\dot\gamma(t))\frac{\partial \Gamma}{\partial s}(0,t)\right) - \sum_{i=1}^n \frac{d}{dt}\left(\frac{\partial L}{\partial v^i}(\dot\gamma(t))\right)\frac{\partial \Gamma}{\partial s}(0,t)\right]dt =$$

$$\int_a^b \sum_{i=1}^n \left[\frac{\partial L}{\partial x^i}(\dot\gamma(t)) - \frac{d}{dt}\left(\frac{\partial L}{\partial v^i}(\dot\gamma(t))\right)\right]\frac{\partial \Gamma}{\partial s}(0,t)dt,$$

because $\frac{\partial \Gamma}{\partial s}(0,a) = \frac{\partial \Gamma}{\partial s}(0,b) = 0$ since the variation is with fixed endpoints. This means that

$$\frac{\partial}{\partial s}\Big|_{s=0}\int_a^b L\Big(\frac{\partial \Gamma}{\partial t}(s,t)\Big)dt = 0$$

if and only if

$$\frac{d}{dt}\Big(\frac{\partial L}{\partial v^i}(\dot\gamma(t))\Big) = \frac{\partial L}{\partial x^i}(\dot\gamma(t))$$

for $i = 1, 2..., n$, because $\frac{\partial \Gamma}{\partial s}(0, t)$ can take any value. □

Example 2.4. Consider a particle of charge e and mass m in $\mathbb{R}^3$ moving under the influence of an electromagnetic field with electrical and magnetic components E and B, respectively. The fields E and B satisfy Maxwell's equations

$$\text{curl}E + \frac{1}{c}\frac{\partial B}{\partial t} = 0, \quad \text{div}B = 0,$$

$$\text{curl}B - \frac{1}{c}\frac{\partial E}{\partial t} = 4\pi J, \quad \text{div}E = 4\pi\rho$$

where c is the speed of light, ρ is the charge density and J is the charge current density. By the Poincaré lemma, there exists a vector potential $A = (A_1, A_2, A_3)$ such that $B = \text{curl}A$. So

$$\text{curl}(E + \frac{1}{c}\frac{\partial A}{\partial t}) = 0$$

and there exists a scalar potential $V : \mathbb{R}^3 \to \mathbb{R}$ such that

$$E = -\text{grad}V - \frac{1}{c}\frac{\partial A}{\partial t},$$

the gradient being euclidean.

Suppose the electromagnetic field does not depend on time and let $L : \mathbb{R}^3 \to \mathbb{R}$ be the Lagrangian

$$L(x, v) = \frac{1}{2}m\langle v, v\rangle + e(\frac{1}{c}\langle A(x), v\rangle - V(x)).$$

We shall describe only the first for $i = 1$ of the corresponding Euler-Lagrange equations, the other two being analogous. The right hand side is

$$\frac{\partial L}{\partial x^1} = -e\frac{\partial V}{\partial x^1} + \frac{e}{c}\langle\frac{\partial A}{\partial x^1}, v\rangle.$$

The left hand side is

$$\frac{d}{dt}(\frac{\partial L}{\partial v^1}) = m\dot{v}^1 + \frac{e}{c}[v^1\frac{\partial A_1}{\partial x^1} + v^2\frac{\partial A_1}{\partial x^2} + v^3\frac{\partial A_1}{\partial x^3}].$$

It follows that the first of the Euler-Lagrange equations takes the form

$$m\dot{v}^1 = -e\frac{\partial V}{\partial x^1} + \frac{e}{c}[v^2(\frac{\partial A_2}{\partial x^1} - \frac{\partial A_1}{\partial x^2}) - v^3(\frac{\partial A_1}{\partial x^3} - \frac{\partial A_3}{\partial x^1})].$$

The right hand side is the first coordinate of the vector $e(E + \frac{1}{c}v \times B)$. Since the other two equations are analogous and have the same form by cyclically permuting indices, we conclude that the Euler-Lagrange equations give Lorentz's equation of motion

$$m\frac{d^2x}{dt^2} = e(E + \frac{1}{c}\frac{dx}{dt} \times B).$$

We globalize the above situation as follows. Let (M, g) be a pseudo-Riemannian n-manifold, A be a smooth 1-form on M and $V : M \to \mathbb{R}$ a smooth function. The Lagrangian

$$L(x, v) = \frac{1}{2} m g(v, v) + A_x(v) - V(x)$$

generalizes the motion of a charged particle of mass m under the influence of an electromagnetic field. Let $(U, x^1, x^2, ..., x^n)$ be a local system of coordinates on M and $(\pi^{-1}(U), x^1, x^2, ..., x^n, v^1, v^2, ..., v^n)$ be the corresponding local system of coordinates on TM. In local coordinates L is given by the formula

$$L(x^1, x^2, ..., x^n, v^1, v^2, ..., v^n) = \frac{1}{2} m \sum_{i,j=1}^{n} g_{ij} v^i v^j + \sum_{i=1}^{n} A_i v^i,$$

where (g_{ij}) is the matrix of the pseudo-Riemannian metric g and $A = \sum_{i=1}^{n} A_i dx^i$ on U. A smooth curve $\gamma : I \to M$ is a Lagrange motion if and only if it satisfies the Euler-Lagrange equations. In our case the right hand side of the Euler-Lagrange equations is

$$\frac{\partial L}{\partial x^k}(\dot{\gamma}(t)) = \frac{1}{2} m \sum_{i,j=1}^{n} \frac{\partial g_{ij}}{\partial x^k} \frac{dx^i}{dt} \frac{dx^j}{dt} + \sum_{i=1}^{n} \frac{\partial A_i}{\partial x^k} \frac{dx^i}{dt} - \frac{\partial V}{\partial x^k},$$

and the left hand side

$$\frac{d}{dt}\big(\frac{\partial L}{\partial v^k}(\dot{\gamma}(t))\big) = m \sum_{i,j=1}^{n} \frac{\partial g_{ik}}{\partial x^j} \frac{dx^j}{dt} \frac{dx^i}{dt} + m \sum_{i=1}^{n} g_{ik} \frac{d^2 x^i}{dt^2} + \sum_{i=1}^{n} \frac{\partial A_k}{\partial x^i} \frac{dx^i}{dt}.$$

So the Euler-Lagrange equations are equivalent to

$$\sum_{i=1}^{n} (\frac{\partial A_i}{\partial x^k} - \frac{\partial A_k}{\partial x^i}) \frac{dx^i}{dt} - \frac{\partial V}{\partial x^k} = m \sum_{i=1}^{n} g_{ik} \frac{d^2 x^i}{dt^2} + m \sum_{i,j=1}^{n} (\frac{\partial g_{ik}}{\partial x^j} - \frac{1}{2} \frac{\partial g_{ij}}{\partial x^k}) \frac{dx^i}{dt} \frac{dx^j}{dt}.$$

On the other hand

$$dA(\dot{\gamma}(t), \frac{\partial}{\partial x^k}) = \sum_{i,j=1}^{n} \frac{\partial A_i}{\partial x^j} dx^j \wedge dx^i (\dot{\gamma}(t), \frac{\partial}{\partial x^k}) = \sum_{i=1}^{n} (\frac{\partial A_k}{\partial x^i} - \frac{\partial A_i}{\partial x^k}) \frac{dx^i}{dt}.$$

Recall that the covariant derivative formula along γ gives

$$\nabla_{\dot{\gamma}} \dot{\gamma} = \sum_{l=1}^{n} \frac{d^2 x^l}{dt^2} \frac{\partial}{\partial x^l} + \sum_{i,j,l=1}^{n} \Gamma_{ij}^{l} \frac{dx^i}{dt} \frac{dx^j}{dt} \frac{\partial}{\partial x^l}$$

and so

$$g(m \nabla_{\dot{\gamma}} \dot{\gamma}, \frac{\partial}{\partial x^k}) = m \sum_{l=1}^{n} g_{lk} \frac{d^2 x^l}{dt^2} + m \sum_{i,j,l=1}^{n} g_{lk} \Gamma_{ij}^{l} \frac{dx^i}{dt} \frac{dx^j}{dt}.$$

Since

$$\sum_{l=1}^{n} g_{lk} \Gamma_{ij}^{l} = \frac{1}{2} \left(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{jk}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k} \right),$$

we get

$$\sum_{i,j,l=1}^{n} g_{lk}\Gamma_{ij}^{l}\frac{dx^i}{dt}\frac{dx^j}{dt} = \sum_{i,j=1}^{n}(\frac{\partial g_{ik}}{\partial x^j} - \frac{1}{2}\frac{\partial g_{ij}}{\partial x^k})\frac{dx^i}{dt}\frac{dx^j}{dt}.$$

We conclude now that the Euler-Lagrange equations have the form

$$g(m\nabla_{\dot\gamma}\dot\gamma, \frac{\partial}{\partial x^k}) = -dA(\dot\gamma, \frac{\partial}{\partial x^k}) - g(\mathrm{grad}V, \frac{\partial}{\partial x^k}), \qquad k = 1, 2, ..., n$$

or independently of local coordinates

$$m\nabla_{\dot\gamma}\dot\gamma = -\mathrm{grad}(i_{\dot\gamma}(dA)) - \mathrm{grad}V,$$

where the gradient is taken with respect to the pseudo-Riemannian metric g.

As in newtonian mechanical systems with potential energy, one can define the notion of mechanical energy for Lagrangian systems also. In order to do this, we shall need to define first the Legendre transformation. So let $L : TM \to \mathbb{R}$ be a Lagrangian and $p \in M$, $v \in T_pM$. The derivative

$$(L|_{T_pM})_{*v} : T_v(T_pM) \cong T_pM \to \mathbb{R}$$

can be considered as an element of the dual tangent space T_p^*M.

Definition 2.5. The Legendre transformation of a Lagrangian system (M, L) is the map $\mathcal{L} : TM \to T^*M$ defined by $\mathcal{L}(p, v) = (L|_{T_pM})_{*v}$. In other words, for every $w \in T_pM$ we have

$$\mathcal{L}(p, v)(w) = \frac{d}{dt}\Big|_{t=0} L(p, v + tw).$$

It is worth to note that $\mathcal{L}$ is not in general a vector bundle morphism, as it may not be linear on fibers. For instance, if $M = \mathbb{R}$ and $L(x, v) = e^v$ (this Lagrangian has no physical meaning), then $\mathcal{L} : T\mathbb{R} \to T^*\mathbb{R} \cong \mathbb{R}^2$ is given by $\mathcal{L}(x, v) = (x, e^v)$, which is not linear in the variable v.

Example 2.6. If $L = \frac{1}{2}g - V$ is the Lagrangian of a newtonian mechanical system with potential energy (M, g, V), then for every $p \in M$ and v, $w \in T_pM$ we have $\mathcal{L}(p, v)(w) = g(v, w)$. Thus, in this particular case the Legendre transformation $\mathcal{L} : TM \to T^*M$ is the natural isomorphism defined by the Riemannian metric.

Definition 2.7. The energy of a Lagrangian system (M, L) is the smooth function $E : TM \to \mathbb{R}$ defined by $E(p, v) = \mathcal{L}(p, v)(v) - L(p, v)$.

If $(x^1, x^2, ..., x^n)$ is a system of local coordinates on M with corresponding local coordinates $(x^1, x^2, ..., x^n, v^1, v^2, ..., v^n)$ on TM, then

$$E(x^1, x^2, ..., x^n, v^1, v^2, ..., v^n) = \sum_{i=1}^{n}\frac{\partial L}{\partial v^i}v^i - L(x^1, x^2, ..., x^n, v^1, v^2, ..., v^n).$$

In the case of a newtonian mechanical system with potential energy (M, g, V) the above definition gives

$$E(p,v) = \mathcal{L}(p,v)(v) - L(p,v) = g(v,v) - \frac{1}{2}g(v,v) + V(p) = \frac{1}{2}g(v,v) + V(p),$$

which coincides with the previous definition.

Example 2.8. We shall compute the Legendre transformation and the energy of the Lagrangian system of example 2.4 using the same notation. Considering local coordinates $(x^1, x^2, ..., x^n, v^1, v^2, ..., v^n)$ on TM, we have

$$(L|_{T_pM})(v^1, v^2, ..., v^n) = \frac{1}{2}m\sum_{i,j=1}^{n} g_{ij}v^iv^j + \sum_{i=1}^{n} A_iv^i - V(p).$$

Differentiating we get

$$(L|_{T_pM})_{*v} = m\sum_{i,j=1}^{n} g_{ij}v^idv^j - \sum_{i=1}^{n} A_idv^i.$$

We conclude now that

$$\mathcal{L}(p,v)(w) = (L|_{T_pM})_{*v}(w) = mg(v,w) + A_p(w).$$

The energy here is

$$E(p,v) = \mathcal{L}(p,v)(v) - L(p,v) = \frac{1}{2}mg(v,v) + V(p),$$

and so does not depend on the 1-form A, which represents the magnetic field. This reflects the fact that the magnetic field does not produce work.

Theorem 2.9. (Conservation of energy) *In a Lagrangian system the energy is a constant of motion.*

Proof. Considering local coordinates on M, let $\gamma = (x^1, x^2, ..., x^n)$ be a Lagrangian motion. Then

$$E(\dot\gamma(t)) = \sum_{i=1}^{n} \frac{\partial L}{\partial v^i}\frac{dx^i}{dt} - L(\dot\gamma(t))$$

and differentiating

$$\frac{d}{dt}(E(\dot\gamma(t))) = \sum_{i,j=1}^{n}\left(\frac{\partial^2 L}{\partial v^i\partial x^j}\frac{dx^i}{dt}\frac{dx^j}{dt} + \frac{\partial^2 L}{\partial v^i\partial v^j}\frac{dx^i}{dt}\frac{d^2x^j}{dt^2}\right) + \sum_{i=1}^{n}\frac{\partial L}{\partial v^i}\frac{d^2x^i}{dt^2}$$

$$-\sum_{i=1}^{n}\frac{\partial L}{\partial x^i}\frac{dx^i}{dt} - \sum_{i=1}^{n}\frac{\partial L}{\partial v^i}\frac{d^2x^i}{dt^2} =$$

$$\sum_{i,j=1}^{n}\left(\frac{\partial^2 L}{\partial v^i\partial x^j}\frac{dx^i}{dt}\frac{dx^j}{dt} + \frac{\partial^2 L}{\partial v^i\partial v^j}\frac{dx^i}{dt}\frac{d^2x^j}{dt^2}\right) - \sum_{i=1}^{n}\frac{\partial L}{\partial x^i}\frac{dx^i}{dt}.$$

But from the Euler-Lagrange equations we have

$$\frac{\partial L}{\partial x^i} = \frac{d}{dt}(\frac{\partial L}{\partial v^i}) = \sum_{j=1}^{n}\left(\frac{\partial^2 L}{\partial v^i \partial x^j}\frac{dx^j}{dt} + \frac{\partial^2 L}{\partial v^i \partial v^j}\frac{d^2x^j}{dt^2}\right)$$

and so substituting we get $\frac{d}{dt}(E(\dot\gamma(t))) = 0$. □

Apart from the energy, one can have constants of motion from symmetries of the Lagrangian.

Theorem 2.10. (Noether) *Let (M, L) be a Lagrangian system and X a complete smooth vector field on M with flow $(\phi_t)_{t\in\mathbb{R}}$. If $L((\phi_t)_{*p}(v)) = L(v)$ for every $v \in T_pM$, $p \in M$ and $t \in \mathbb{R}$, then the smooth function $f_X : TM \to \mathbb{R}$ defined by*

$$f_X(v) = \lim_{s\to 0}\frac{L(v + sX(\pi(v))) - L(v)}{s}$$

is a constant of motion.

Proof. Considering local coordinates, let $\phi_t = (\phi_t^1, \phi_t^2, ..., \phi_t^n)$. Since L is $(\phi_t)_*$ invariant, if $\gamma = (x^1, x^2, ..., x^n)$ is a Lagrangian motion, differentiating the equation $L((\phi_s)_{*\gamma(t)}(\dot\gamma(t))) = L(\dot\gamma(t))$ with respect to s, we have

$$\sum_{i=1}^{n}\frac{\partial L}{\partial x^i}\left(\frac{\partial \phi_s^i}{\partial s}\right)_{s=0} + \sum_{i,j=1}^{n}\frac{\partial L}{\partial v^i}\left(\frac{\partial^2 \phi_s^i}{\partial x^j \partial s}\right)_{s=0}\frac{dx^j}{dt} = 0.$$

Since $f_X(v)$ is the directional derivative of $L|_{T_{\pi(v)}M}$ in the direction of $X(\pi(v))$ and

$$X = \sum_{i=1}^{n}\left(\frac{\partial \phi_t^i}{\partial t}\right)_{t=0}\frac{\partial}{\partial x^i},$$

we have

$$f_X(\dot\gamma(t)) = \sum_{i=1}^{n}\frac{\partial L}{\partial v^i}\left(\frac{\partial \phi_s^i}{\partial s}\right)_{s=0}.$$

Using now the Euler-Lagrange equations we compute

$$\frac{d}{dt}(f_X(\dot\gamma(t))) = \sum_{i=1}^{n}\frac{d}{dt}\left(\frac{\partial L}{\partial v^i}\right)\left(\frac{\partial \phi_s^i}{\partial s}\right)_{s=0} + \sum_{i=1}^{n}\frac{\partial L}{\partial v^i}\frac{d}{dt}\left(\frac{\partial \phi_s^i}{\partial s}\right)_{s=0} =$$

$$\sum_{i=1}^{n}\frac{\partial L}{\partial x^i}\left(\frac{\partial \phi_s^i}{\partial s}\right)_{s=0} + \sum_{i,j=1}^{n}\frac{\partial L}{\partial v^i}\left(\frac{\partial^2 \phi_s^i}{\partial x^j \partial s}\right)_{s=0}\frac{dx^j}{dt} = 0. \qquad \square$$

Examples 2.11. (a) Let (M, g, V) be a newtonian mechanical system with potential energy and X be a complete vector field, which is a symmetry of the system. Then

$f_X(v) = g(v, X(\pi(v)))$. The restriction to a fiber of the tangent bundle of f_X is linear in this case.

(b) Let X be a complete vector field which we assume to be a symmetry of the Lagrangian system of example 2.3. For instance, this is the case if the flow of X preserves the pseudo-Riemannian metric on M and the 1-form A. Then the Lagrangian is X-invariant and the first integral provided form Noether's theorem is $f_X(v) = mg(v, X) - A(X)$.

Let (M, L) be a Lagrangian system. Let $(x^1, ..., x^n, v^1, ..., v^n)$ be a system of local coordinates in TM coming from local coordinates $(x^1, ..., x^n)$ on M and let $(q^1, ..., q^n, p_1, ..., p_n)$ be the corresponding local coordinates on T^*M, so that $x^j = q^j$, $1 \leq j \leq n$. The local representation of the Legendre transformation is

$$\mathcal{L}(x^1, ..., x^n, \sum_{j=1}^{n} v^j \frac{\partial}{\partial x^j}) = (x^1, ..., x^n, \sum_{j=1}^{n} \frac{\partial L}{\partial v^j} dx^j).$$

The local forms

$$\sum_{j=1}^{n} \frac{\partial L}{\partial v^j} dx^j$$

over all charts on TM fit together and give a global smooth 1-form θ_L on TM. This may be verified directly. Alternatively, we note that

$$\mathcal{L}^*(\sum_{i=1}^{n} p_i dq^i) = \sum_{i=1}^{n} (p_i \circ \mathcal{L}) d(x^i \circ \mathcal{L}) = \sum_{i=1}^{n} \frac{\partial L}{\partial v^i} dx^i.$$

The local 1-forms $\sum_{i=1}^{n} p_i dq^i$ on T^*M fit together to a global smooth 1-form θ on T^*M. Actually, θ is precisely the 1-form defined by

$$\theta_a(v) = a(\pi_{*a}(v))$$

for $v \in T_a(T^*M)$ and $a \in T^*M$, where $\pi : T^*M \to M$ is the cotangent bundle projection. Indeed,

$$\theta|_{\text{locally}} = \sum_{i=1}^{n} \theta(\frac{\partial}{\partial q^i}) dq^i + \sum_{i=1}^{n} \theta(\frac{\partial}{\partial p_i}) dp_i.$$

If now $a = (q^1, q^2, ..., q^n, p_1, p_2, ..., p_n)$, then $\pi_{*a}(\frac{\partial}{\partial p_i}) = 0$, and therefore $\theta(\frac{\partial}{\partial p_i}) = 0$. Moreover,

$$\theta(\frac{\partial}{\partial q^i}) = a(\pi_{*a}(\frac{\partial}{\partial q^i})) = p_i.$$

It follows that

$$\theta|_{\text{locally}} = \sum_{i=1}^{n} p_i dq^i.$$

The smooth 1-form θ is called the *Liouville canonical 1-form* on T^*M.

Remark 2.12. The 2-form $d\theta_L$ in local coordinates $(x^1, ..., x^n, v^1, ..., v^n)$ on TM is given by the formula

$$d\theta_L|_{\text{locally}} = \sum_{i,j=1}^{n} \frac{\partial^2 L}{\partial x^j \partial v^i} dx^j \wedge dx^i + \sum_{i,j=1}^{n} \frac{\partial^2 L}{\partial v^j \partial v^i} dv^j \wedge dx^i.$$

It follows that $d\theta_L$ is non-degenerate if and only if the vertical Hessian matrix

$$\left(\frac{\partial^2 L}{\partial v^j \partial v^i} \right)_{1 \le i,j \le n}$$

is everywhere invertible. A Lagrangian system is called non-degenerate if the vertical Hessian of the Lagrangian is everywhere invertible.

1.3 The equations of Hamilton

A Lagrangian system (M, L) is called hyperregular if the Legendre transformation $\mathcal{L} : TM \to T^*M$ is a diffeomorphism. For example a newtonian mechanical system with potential energy and the system of example 2.4 are hyperregular.

Definition 3.1. In a hyperregular Lagrangian system as above, the smooth function $H = E \circ \mathcal{L}^{-1} : T^*M \to \mathbb{R}$, where E is the energy, is called the Hamiltonian function of the system.

Example 3.2. Let (M, g, V) is a newtonian mechanical system with potential energy. The Legendre transformation gives

$$q^i = x^i, \quad p_i = \frac{\partial L}{\partial v^i} = \sum_{j=1}^{n} g_{ij} v^j.$$

The inverse Legendre transformation is given by

$$x^i = q^i, \quad v^i = \sum_{j=1}^{n} g^{ij} p_j.$$

So we have

$$E = \frac{1}{2} \sum_{i,j=1}^{n} g_{ij} v^i v^j + V(x^1, x^2, ..., x^n),$$

$$L = \frac{1}{2} \sum_{i,j=1}^{n} g_{ij} v^i v^j - V(x^1, x^2, ..., x^n)$$

and therefore

$$H = \frac{1}{2} \sum_{i,j=1}^{n} g^{ij} p_i p_j + V(q^1, q^2, ..., q^n).$$

Theorem 3.3. (Hamilton) *Let (M, L) be a hyperregular Lagrangian system on the n-dimensional manifold M. A smooth curve $\gamma : I \to M$ is a Lagrangian motion if and only if the smooth curve $\mathcal{L}\circ\dot{\gamma} : I \to T^*M$ locally solves the system of differential equations*

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q^i}, \quad i = 1, 2, ..., n.$$

Proof. In the local coordinates $(x^1, x^2, ..., x^n)$ of a chart on M the Legendre transformation is given by the formulas

$$q^i = x^i, \quad p_i = \frac{\partial L}{\partial v^i}, \quad i = 1, 2, ..., n,$$

where $(x^1, x^2, ..., x^n, v^1, v^2, ..., v^n)$ are the local coordinates of the corresponding chart on TM. Inversing,

$$x^i = q^i, \quad v^i = y^i(q^1, q^2, ..., q^n, p_1, p_2, ..., p_n), \quad i = 1, 2, ..., n$$

for some smooth functions y^1, y^2,...,y^n. From the definitions of the energy E and the Hamiltonian H, we have

$$H = E \circ \mathcal{L}^{-1} = \sum_{j=1}^{n} p_j y^j - L(q^1, q^2, ..., q^n, y^1, y^2, ..., y^n)$$

and differentiating the chain rule gives

$$\frac{\partial H}{\partial p_i} = y^i + \sum_{j=1}^{n} p_j \frac{\partial y^j}{\partial p_i} - \sum_{j=1}^{n} \frac{\partial L}{\partial v^j}\frac{\partial y^j}{\partial p_i} = y^i$$

$$\frac{\partial H}{\partial q^i} = \sum_{j=1}^{n} p_j \frac{\partial y^j}{\partial q^i} - \frac{\partial L}{\partial x^i} - \sum_{j=1}^{n} \frac{\partial L}{\partial v^j}\frac{\partial y^j}{\partial q^i} = -\frac{\partial L}{\partial x^i}.$$

If $\gamma(t) = (x^1(t), x^2(t), ..., x^n(t))$ is a smooth curve in local coordinates on M, then

$$\mathcal{L}(\dot{\gamma}(t)) = (x^1(t), x^2(t), ..., x^n(t), \frac{\partial L}{\partial v^1}(\dot{\gamma}(t)), \frac{\partial L}{\partial v^2}(\dot{\gamma}(t)), ..., \frac{\partial L}{\partial v^n}(\dot{\gamma}(t))).$$

Now γ is a Lagrangian motion if and only if

$$\dot{x}^i = v^i \quad \text{and} \quad \frac{d}{dt}\left(\frac{\partial L}{\partial v^i}(\dot{\gamma})\right) = \frac{\partial L}{\partial x^i}(\dot{\gamma})$$

or equivalently

$$\dot{q}^i = \dot{x}^i = v^i = y^i = \frac{\partial H}{\partial p_i} \quad \text{and}$$

$$\dot{p}_i = \frac{d}{dt}\left(\frac{\partial L}{\partial v^i}(\dot{\gamma})\right) = \frac{\partial L}{\partial x^i}(\dot{\gamma}) = -\frac{\partial H}{\partial q^i}. \qquad \square$$

The equations provided by Theorem 3.3 on T^*M are *Hamilton's equations.* The cotangent bundle T^*M is called the phase space of the Lagrangian system (M, L).

Corollary 3.4. *The Hamiltonian is constant on solutions of Hamilton's equations.*

Proof. Indeed, if $\gamma(t) = (q^1(t), ..., q^n(t), p_1(t), ..., p_n(t))$ is the local form of a solution of Hamilton's equations then

$$dH(\gamma(t))(\dot{\gamma}(t)) = \sum_{i=1}^{n} \frac{\partial H}{\partial q^i}\dot{q}^i(t)+\sum_{i=1}^{n} \frac{\partial H}{\partial p_i}\dot{p}_i(t) = \sum_{i=1}^{n} \frac{\partial H}{\partial q^i}\frac{\partial H}{\partial p_i}+\sum_{i=1}^{n} \frac{\partial H}{\partial p_i}(-\frac{\partial H}{\partial q^i}) = 0. \quad \square$$

The equations of Hamilton have a global formulation on T^*M in the sense that a solution is the integral curve of a smooth vector field defined globally on T^*M. Recall that the Liouville canonical 1-form θ on T^*M has a local expression

$$\theta|_{\text{locally}} = \sum_{i=1}^{n} p_i dq^i.$$

Let $\omega = -d\theta$, so that

$$\omega|_{\text{locally}} = \sum_{i=1}^{n} dq^i \wedge dp_i.$$

Since for every smooth vector field Y on T^*M we have

$$Y = \sum_{i=1}^{n} \omega(Y, \frac{\partial}{\partial p_i})\frac{\partial}{\partial q^i} - \sum_{i=1}^{n} \omega(Y, \frac{\partial}{\partial q^i})\frac{\partial}{\partial p_i}$$

it follows that ω is a non-degenerate, closed 2-form on T^*M. Thus, given a smooth function $H : T^*M \to \mathbb{R}$, there exists a unique smooth vector field X on T^*M such that $i_X\omega = dH$, called the Hamiltonian vector field. Locally this global equation takes the form

$$\sum_{i=1}^{n} dq^i(X)dp_i - \sum_{i=1}^{n} dp_i(X)dq^i = \sum_{i=1}^{n} \frac{\partial H}{\partial q^i}dq^i + \sum_{i=1}^{n} \frac{\partial H}{\partial p_i}dp_i$$

and therefore $dp_i(X) = -\dfrac{\partial H}{\partial q^i}$ and $dq^i(X) = \dfrac{\partial H}{\partial p_i}$. These are precisely Hamilton's equations.

Chapter 2

Basic symplectic geometry

2.1 Symplectic linear algebra

A *synplectic form* on a (real) vector space V of finite dimension is a non-degenerate, skew-symmetric, bilinear form $\omega : V \times V \to \mathbb{R}$. This means that the map $\tilde{\omega} : V \to V^*$ defined by $\tilde{\omega}(v)(w) = \omega(v, w)$, for v, $w \in V$, is a linear isomorphism. The pair (V, ω) is then called a *symplectic vector space.*

Lemma 1.1. (Cartan) *Let V be a vector space of dimension n and ω be a skew-symmetric, bilinear form on V. If $\omega \neq 0$, then the rank of $\tilde{\omega}$ is even. If $\dim \tilde{\omega}(V) = 2k$, there exists a basis l^1, l^2,...,l^{2k} of $\tilde{\omega}(V)$ such that*

$$\omega = \sum_{j=1}^{k} l^{2j-1} \wedge l^{2j}.$$

Proof. Let $\{v_1, v_2, ..., v_n\}$ be a basis of V and $\{v_1^*, v_2^*, ..., v_n^*\}$ be the corresponding dual basis of V^*. If $a_{ij} = \omega(v_i, v_j)$, $i < j$, then

$$\omega = \sum_{i<j} a_{ij} v_i^* \wedge v_j^*.$$

Since $\omega \neq 0$, there are some $1 \leq i < j \leq n$ such that $a_{ij} \neq 0$. We may assume that $a_{12} \neq 0$, changing the numbering if necessary. Let

$$l^1 = \frac{1}{a_{12}} \tilde{\omega}(v_1) = v_2^* + \frac{1}{a_{12}} \sum_{j=3}^{n} a_{1j} v_j^*,$$

$$l^2 = \tilde{\omega}(v_2) = -a_{12} v_1^* + \sum_{j=3}^{n} a_{2j} v_j^*.$$

The set $\{l^1, l^2, v_3^*, ..., v_n^*\}$ is now a new basis of V^*. If $\omega_1 = \omega - l^1 \wedge l^2$, then

$$\tilde{\omega}_1(v_1) = a_{12} l^1 - l^1(v_1) l^2 + l^2(v_1) l^1 = a_{12} l^1 - 0 - a_{12} l^1 = 0,$$

$$\tilde{\omega}_1(v_2) = l^2 - l^1(v_2)l^2 + l^2(v_2)l^1 = l^2 - l^2 + 0 = 0.$$

Thus, ω_1 is an element of the subalgebra of the exterior algebra of V generated by v_3^*,,...,v_n^*. If $\omega_1 = 0$, then $\omega = l^1 \wedge l^2$. If $\omega_1 \neq 0$, we repeat the above taking ω_1 in the place of ω. So, inductively, we arrive at the conclusion, since V has finite dimension. □

Corollary 1.2. *If ω is a skew-symmetric, bilinear form of rank $2k$ of a vector space, then k is the maximal positive integer such that $\omega \wedge ... \wedge \omega \neq 0$ (k times).*

Proof. Indeed from Cartan's lemma, the $(k+1)$-fold wedge product of ω with itself is equal to 0 and the k-fold is $\omega \wedge ... \wedge \omega = k! \cdot l^1 \wedge ... \wedge l^{2k} \neq 0$. □

By Cartan's lemma, if (V, ω) is a symplectic vector space of finite dimension, there exists some $n \in \mathbb{N}$ such that $\dim V = 2n$ and there exists a basis $\{a_1, ..., a_n, b_1, ..., b_n\}$ of V^* such that

$$\omega = a_1 \wedge b_1 + a_2 \wedge b_2 + \cdots + a_n \wedge b_n.$$

This basis is dual to a basis $\{v_1, ..., v_n, u_1, ..., u_n\}$ of V which is called a symplectic basis and is characterized by the properties $\omega(v_i, v_j) = \omega(u_i, u_j) = 0$ and $\omega(v_i, u_j) = \delta_{ij}$ for $1 \leq i, j \leq n$.

If $W \leq V$, we set $W^\perp = \{v \in V : \omega(w, v) = 0 \quad \text{for every} \quad w \in W\}$. Obviously, $W^\perp \leq V$ and $\tilde{\omega}(W^\perp) = \{a \in V^* : a|_W = 0\}$, since $\tilde{\omega}$ is an isomorphism.

Lemma 1.3. *Let (V, ω) be a symplectic vector space of dimension $2n$ and W_1, W_2, W be subspaces of V. Then the following hold:*

(a) $\dim W + \dim W^\perp = \dim V = 2n$.
(b) $W^{\perp\perp} = W$.
(c) $W_1 \leq W_2$ *if and only if* $W_2^\perp \leq W_1^\perp$.
(d) $W_1^\perp \cap W_2^\perp = (W_1 + W_2)^\perp$.
(e) $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$.

Proof. (a) Since $\tilde{\omega}(W^\perp)$ coincides with the anihilator of W, its dimension is $\dim V - \dim W$ and $W^\perp$ has the same dimension, because $\tilde{\omega}$ is an isomorphism.

(b) Evidently, $W \leq W^{\perp\perp}$ and since $\dim W = \dim W^{\perp\perp}$, by (a), we have $W = W^{\perp\perp}$.

(c) If $W_1 \leq W_2$, then obviously $W_2^\perp \leq W_1^\perp$. Conversely, if $W_2^\perp \leq W_1^\perp$, then from (b) we have $W_1 = W_1^{\perp\perp} \leq W_2^{\perp\perp} = W_2$.

(d) From (c) we have $(W_1 + W_2)^\perp \leq W_1^\perp \cap W_2^\perp$ and if $v \in W_1^\perp \cap W_2^\perp$ and $w_1 \in W_1$, $w_2 \in W_2$, then

$$\omega(v, w_1 + w_2) = \omega(v, w_1) + \omega(v, w_2) = 0 + 0 = 0.$$

Thus, $v \in (W_1 + W_2)^\perp$, which shows that $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$.

(e) From (b) and (d) we have $(W_1 \cap W_2)^\perp = (W_1^\perp + W_2^\perp)^{\perp\perp} = W_1^\perp + W_2^\perp$. □

Example 1.4. Let W be a vector space of dimension n. On $W \times W^*$ consider the skew-symmetric, bilinear form ω defined by

$$\omega((w,a),(w',a')) = a'(w) - a(w').$$

If now $\tilde{\omega}(w,a) = 0$, then $0 = \omega((w,a),(w',0)) = -a(w')$ for every $w' \in W$. Thus $a = 0$. Similarly, $0 = \omega((w,a),(0,a')) = a'(w)$ for every $a' \in W^*$. Hence $w = 0$. This shows that $(W \times W^*, \omega)$ is a symplectic vector space.

Let (V, ω) be a symplectic vector space of dimension $2n$. A subspace $W \leq V$ is called
(i) isotropic, if $W \leq W^\perp$, and then $\dim W \leq n$,
(ii) coisotropic, if $W^\perp \leq W$, and then $\dim W \geq n$,
(iii) Lagrangian, if $W = W^\perp$, and then $\dim W = n$, and
(iv) symplectic, if $W \cap W^\perp = \{0\}$, and then $\dim W$ is even.

For instance, in Example 1.4, the subspaces $W \times \{0\}$ and $\{0\} \times W^*$ are Lagrangian. Obviously, any 1-dimensional subspace is isotropic.

Proposition 1.5. *Every isotropic subspace of a symplectic vector space (V, ω) is contained in a Lagrangian subspace.*

Proof. Let $W \leq V$ be an isotropic subspace, which is not Lagrangian itself. There exists $v \in W^\perp \setminus W$. Then $\langle v \rangle$ is an isotropic subspace of V and therefore $\langle v \rangle \leq \langle v \rangle^\perp \cap W^\perp$. By Lemma 1.3(c), $W \leq \langle v \rangle^\perp$, and thus $W \leq \langle v \rangle^\perp \cap W^\perp$, since W is isotropic. From Lemma 1.3(d) we have now

$$\langle v \rangle + W \leq \langle v \rangle^\perp \cap W^\perp = (\langle v \rangle + W)^\perp,$$

which means that $\langle v \rangle + W$ is isotropic. Since $\dim V$ is finite, repeating this process we arrive after a finite number of steps at a Lagrangian subspace. □

Corollary 1.6. *Every symplectic vector space contains a Lagrangian subspace.* □

The Lagrangian subspaces of a symplectic vector space can be characterized as follows.

Proposition 1.7. *Let W be a linear subspace of a symplectic vector space (V, ω) of dimension $2n$. The following assertions are equivalent.*
(i) W is Lagrangian.
(ii) W is isotropic and $\dim W = n$.
(iii) W is isotropic and has an isotropic complement in V.

Proof. Obviously, assertions (i) and (ii) are equivalent. To prove that they imply (iii), we construct and isotropic complement of the Lagrangian subspace W. Let $v_1 \notin W$ and $U_1 = \langle v_1 \rangle$. Then U_1 is isotropic and $W \cap U_1 = \{0\}$. Therefore,

$$W + U_1^\perp = W^\perp + U_1^\perp = V.$$

In the second step we choose $v_2 \in U_1^\perp$ such that $v_2 \notin W + U_1$ and put $U_2 = U_1 \oplus \langle v_1 \rangle$. Then U_2 is isotropic, because

$$U_2 = \langle v_1, v_2 \rangle \leq U_1^\perp \cap \langle v_2 \rangle^\perp = U_2^\perp$$

and $W \cap U_2 = \{0\}$, so that $W + U_2^\perp = V$. Inductively, if an isotropic subspace U_{k-1} of dimension $k-1$ has been defined such that $W \cap U_{k-1} = \{0\}$, we put $U_k = U_{k-1} \oplus \langle v_k \rangle$, where $v_k \in U_{k-1}^\perp$ is such that $v_k \notin W + U_{k-1}$. Obviously, $W \cap U_k = \{0\}$, so that $W + U_k^\perp = V$ and

$$U_k = U_{k-1} \oplus \langle v_k \rangle \leq U_{k-1}^\perp \cap \langle v_k \rangle^\perp = U_k^\perp.$$

Since at each step the dimension increases by 1, we have $\dim U_n = n$ and U_n is a Lagrangian complement of W in V.

For the converse it suffices to prove that (iii) implies $W^\perp \leq W$. Let $v \in W^\perp$. If U is an isotropic complement of W in V, there exist unique $w \in W$ and $u \in U$ such that $v = w + u$. Since U is isotropic, we have

$$u = v - w \in W^\perp \cap U^\perp = (W + U)^\perp = V^\perp = \{0\}.$$

Consequently, $v = w \in W$. □

A linear map $f : (V, \omega) \to (V', \omega')$ between symplectic vector spaces is called *symplectic* if $f^*\omega' = \omega$. Evidently, every symplectic linear map is injective.

Theorem 1.8. *For every positive integer n there exists exactly one symplectic vector space of dimension $2n$, up to symplectic linear isomorphism.*

Proof. Let (V, ω) be a symplectic vector space of dimension $2n$. It suffices to construct a symplectic linear isomorphism from (V, ω) to the standard Example 1.4. By Corollary 1.6, there exists a Lagrangian subspace W of V. By Proposition 1.7, there exists a Lagrangian subspace U of V such that $V = W \oplus U$. Let $F : U \to W^*$ be the linear map defined by $F(u)(w) = \omega(w, u)$. Since W is Lagrangian, F is a linear isomorphism. Therefore, $f = id \oplus F : V \to W \oplus W^*$ is a linear isomorphism. Moreover,

$$F(u_2)(w_1) - F(u_1)(w_2) = \omega(w_1, u_2) - \omega(w_2, u_1) = \omega(w_1 + u_1, w_2 + u_2)$$

for every w_1, $w_2 \in W$ and u_1, $u_2 \in U$, because W and U are Lagrangian. It follows that f is symplectic. □

Example 1.9. Let h denote the usual hermitian product on $\mathbb{C}^n$. As a real vector space $\mathbb{C}^n \cong \mathbb{R}^{2n}$ carries the symplectic structure

$$\omega(z, w) = -\mathrm{Im} h(z, w).$$

A real subspace W is isotropic if and only if $h(w_1, w_2) \in \mathbb{R}$ for every w_1, $w_2 \in W$. Let $J : \mathbb{C}^n \to \mathbb{C}^n$ be multiplication by i. Then $h(z, u) = h(J(z), J(u))$ for every z, $u \in \mathbb{C}^n$ and so W is isotropic if and omly if $J(W)$ is. Obviously, $W \cap J(W) = \{0\}$, since W is a real subspace. It follows that W is Lagrangian if and only if $h(w_1, w_2) \in \mathbb{R}$ for every w_1, $w_2 \in W$ and $\mathbb{C}^n = W \oplus J(W)$. Thus, $\mathbb{C}^n$ is the complexification of the real space W and the real and imaginary subspaces W and $J(W)$, respectively, are Lagrangian.

Let (V, ω) be a symplectic vector space. A complex structure on V is a linear automorphism $J : V \to V$ such that $J^2 = -id$. It is said to be compatible with the symplectic structure if it is symplectic and $\omega(v, J(v)) > 0$ for all non-zero $v \in V$.

Theorem 1.10. *On every symplectic vector space (V, ω) there exists a compatible complex structure J and a positive definite inner product g given by the formula $g(u, v) = \omega(u, J(v))$ for every u, $v \in V$.*

Proof. Let $\langle , \rangle$ be any positive definite inner product on V. All adjoints below are taken with respect to this inner product. Since ω is non-degenerate and skew-symmetric, there exists a unique skew-symmetric linear automorphism $A : V \to V$ such that $\omega(u, v) = \langle A(u), v\rangle$ for every u, $v \in V$. So, $A^t = -A$ and $-A^2 = A^t A$ is a positive definite self-adjoint linear automorphism of V, which has a unique square root. This means that there exists a unique positive definite self-adjoint linear automorphism $B : V \to V$ such that $B^2 = -A^2$. Let $J = AB^{-1}$. Then

$$J^t = (B^{-1})^t A^t = -B^{-1} A = BA^{-1} = J^{-1},$$

which means that J is orthogonal. Moreover,

$$B^2 = -A^2 = AA^t = JBB^t J^t = JB^2 J^{-1} = (JBJ^{-1})^2.$$

From the uniqueness of the square root of $-A^2$ we get $B = JBJ^{-1}$ and therefore $JB = BJ$ and $J = AB^{-1} = B^{-1}A$. We conclude that A and B commute and

$$J^2 = A(B^2)^{-1} A = A(-A^2)^{-1} A = -id.$$

Also,

$$\begin{aligned}\omega(J(u), J(v)) &= \langle A^2 B^{-1}(u), AB^{-1}(v)\rangle = \langle -B(u), B^{-1}A(v)\rangle \\ &= \langle -u, BB^{-1}A(v)\rangle = \langle -u, A(v)\rangle = \langle A(u), v\rangle = \omega(u, v)\end{aligned}$$

for every u, $v \in V$. Finally, the formula $g(u, v) = \omega(u, J(v))$ defines a positive definite inner product, because $g(u, v) = \langle AJ^{-1}(u), v\rangle = \langle B(u), v\rangle$ for every u, $v \in V$. □

Having a compatible complex structure J on a symplectic vector space (V, ω), the latter becomes a complex vector space by setting $i \cdot v = J(v)$ for every $v \in V$. Moreover, $h = g - i\omega$ is a positive definite hermitian product on V. The triple (V, ω, J) is called a Kähler vector space.

Returning to the proof of Theorem 1.10, we note that the compatible complex structure J is a function of the initially chosen positive definite inner product. Every compatible complex structure arises in this way, because $\omega(u, v) = g(J(u), v)$. If we start with two positive definite inner products s_1 and s_2 on V which lead to compatible complex structures J_1 and J_2, respectively, then the linear path $s_t =$

$(1-t)s_1 + ts_2$, $0 \leq t \leq 1$, gives a path from J_1 to J_2. This shows that the space $\mathcal{J}(V,\omega)$ of the compatible complex structures is path connected. We shall prove later that it is actually contractible.

2.2 The symplectic linear group

Recall that for every $n \in \mathbb{N}$ there exists only one symplectic vector space of dimension $2n$ (up to symplectic linear isomorphism). So we need only consider $\mathbb{R}^n \times (\mathbb{R}^n)^*$ with the canonical symplectic structure ω of Example 1.4. If $\langle , \rangle$ denotes the euclidean inner product on $\mathbb{R}^{2n}$, then $\mathbb{R}^n \times (\mathbb{R}^n)^*$ can be identified with $\mathbb{R}^{2n} \cong \mathbb{R}^n \oplus \mathbb{R}^n$, where the symplectic form is given by the formula

$$\omega((x,y),(x',y')) = \langle x, y'\rangle - \langle x', y\rangle.$$

The complex structure $J : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is the orthogonal transformation $J(x,y) = (-y,x)$ for x, $y \in \mathbb{R}^n$. Then, $J^2 = -id$ and

$$\omega((x,y),(x',y')) = \langle J(x,y),(x',y')\rangle.$$

A linear map $f : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ with matrix A (with respect to the canonical basis) is symplectic if and only if

$$\langle Jv, w\rangle = \omega(v,w) = \omega(Av,Aw) = \langle JAv, Aw\rangle = \langle A^tJAv, w\rangle,$$

for every v, $w \in \mathbb{R}^{2n}$. So, f is symplectic if and only if $A^tJA = J$. The set of symplectic linear maps

$$\mathrm{Sp}(n,\mathbb{R}) = \{A \in \mathbb{R}^{2n\times 2n} : A^tJA = J\}$$

is a Lie group, as a closed subgroup of $GL(2n,\mathbb{R})$, and is called the symplectic group.

To see that $\mathrm{Sp}(n,\mathbb{R})$ is a Lie group in an elementary way which gives directly its Lie algebra, let $F : GL(2n,\mathbb{R}) \to \mathfrak{so}(2n,\mathbb{R})$ be the smooth map

$$F(A) = A^tJA.$$

Then, $\mathrm{Sp}(n,\mathbb{R}) = F^{-1}(J)$ and it suffices to show that J is a regular value of F. The derivative of F at A is $F_{*A}(H) = H^tJA + A^tJH$, $H \in \mathbb{R}^{2n\times 2n}$. Let $A \in \mathrm{Sp}(n,\mathbb{R})$ and $B \in \mathfrak{so}(2n,\mathbb{R})$. If $H = -\frac{1}{2}AJB$, since $A^tJ = JA^{-1}$, then

$$H^tJA + A^tJH = -\frac{1}{2}(JB)^tJ + J(-\frac{1}{2}JB) = -\frac{1}{2}B^tJ^tJ - \frac{1}{2}J^2B = B.$$

This shows that F_{*A} is a linear epimorphism. Hence $\mathrm{Sp}(n,\mathbb{R})$ is a Lie group with Lie algebra

$$\mathfrak{sp}(n,\mathbb{R}) = \{H \in \mathbb{R}^{2n\times 2n} : H^tJ + JH = 0\}.$$

equipped with the usual Lie bracket of matrix groups and has dimension $2n^2 + n$.

Let $h : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ be the usual hermitian product defined by

$$h(u,v) = \langle u, v\rangle - i\omega(u,v).$$

If now $A \in GL(2n, \mathbb{R})$, then identifying $\mathbb{R}^{2n}$ with $\mathbb{C}^n$ we have
$A \in U(n)$ if and only if $h(Au, Av) = h(u, v)$ for every $u,\ v \in \mathbb{R}^{2n}$
if and only if $\langle Au, Av\rangle = \langle u, v\rangle$ and $\omega(Au, Av) = \omega(u, v)$ for every $u,\ v \in \mathbb{R}^{2n}$
if and only if $A \in O(2n, \mathbb{R}) \cap \mathrm{Sp}(n, \mathbb{R})$
if and only if $JA = AJ$ and $\langle Au, Av\rangle = \langle u, v\rangle$ for every $u,\ v \in \mathbb{R}^{2n}$
if and only if $A \in O(2n, \mathbb{R}) \cap GL(n, \mathbb{C})$
if and only if $\langle AJu, Av\rangle = \langle JAu, Av\rangle = \langle Ju, v\rangle$ for every $u,\ v \in \mathbb{R}^{2n}$
if and only if $A \in \mathrm{Sp}(n, \mathbb{R}) \cap GL(n, \mathbb{C})$.
In other words,

$$O(2n, \mathbb{R}) \cap \mathrm{Sp}(n, \mathbb{R}) = O(2n, \mathbb{R}) \cap GL(n, \mathbb{C}) = \mathrm{Sp}(n, \mathbb{R}) \cap GL(n, \mathbb{C}) = U(n).$$

Note that every element $A \in \mathrm{Sp}(n, \mathbb{R})$ preserves the volume and $\det A = 1$. If λ is an eigenvalue of A with multiplicity $k \in \mathbb{N}$, then since

$$\det(A - \lambda I_{2n}) = \det(A^t - \lambda I_{2n}) = \det(J^{-1}(A^t - \lambda I_{2n})J) =$$

$$\det(A^{-1} - \lambda I_{2n}) = \det A^{-1} \cdot \det(I_{2n} - \lambda A) = \lambda^{2n} \cdot \det(A - \frac{1}{\lambda} I_{2n})$$

and $\lambda \neq 0$, because A is invertible, we conclude that $\frac{1}{\lambda}$ is also an eigenvalue of A of multiplicity k. Hence tha total multiplicity of all eigenvalues not equal to 1 or -1 is even. It follows that if -1 is an eigenvalue of A, it occurs with even multiplicity and the same holds for 1, if it is an eigenvalue of A. Finally, if λ, μ are two real eigenvalues of A and $\lambda\mu \neq 1$, then the corresponding eigenspaces are ω-orthogonal, because if $Ax = \lambda x$ and $Ay = \mu y$, then

$$\langle J(x), y\rangle = \omega(x, y) = \omega(Ax, Ay) = \langle JA(x), A(y)\rangle = \lambda\mu\langle J(x), y\rangle$$

and therefore we must have $\omega(x, y) = 0$.

There is a symplectic version of the well known polar decomposition for the general linear groups. If $A \in \mathrm{Sp}(n, \mathbb{R})$, then

$$(A^t)^t J A^t = AJA^t = AJJA^{-1}J^{-1} = -J^{-1} = J$$

and

$$(A^t A)^t J A^t A = A^t A J A^t A = A^t J A = J$$

and therefore A^t and $R = A^t A$ are also symplectic. Also R is symmetric and $RJR = J$ or $JR = R^{-1}J$ and $RJ = JR^{-1}$. Moreover, R is positive definite (with respect to the euclidean inner product). So there exists a unique positive definite, symmetric matrix S such that $S^2 = R$, and $S \in \mathrm{Sp}(n, \mathbb{R})$, since

$$(-JS^{-1}J)^2 = -JS^{-2}J = -(JS^2J)^{-1} = -(JRJ)^{-1} = -(-R^{-1})^{-1} = R$$

and so $-JS^{-1}J = S$, by uniqueness, or equivalently $J = SJS$, which means that $S \in \mathrm{Sp}(n, \mathbb{R})$. This argument shows that the square root of any positive definite, symmetric, symplectic matrix is also symplectic.

Now if $U = AS^{-1}$, then

$$\langle Ux, Uy\rangle = \langle S^{-1}x, RS^{-1}y\rangle = \langle S^{-1}x, Sy\rangle = \langle x, y\rangle$$

for every x, $y \in \mathbb{R}^{2n}$, since S is symmetric. Therefore, $U \in O(2n, \mathbb{R}) \cap \mathrm{Sp}(n, \mathbb{R}) = U(n)$ and we get the polar decomposition $A = US$. Note that given such a decomposition, we have $A^t = SU^{-1}$ and so $A^tA = SU^{-1}US = S^2$, which means that S must be necessarily the unique square root of A^tA. This implies the uniqueness of the polar decomposition.

Lemma 2.1. *If R is positive definite, symmetric, symplectic matrix, then $R^a \in Sp(n, \mathbb{R})$ for every real number $a > 0$.*

Proof. From the spectral theorem, $\mathbb{R}^{2n}$ is decomposed into a direct sum (actually orthogonal with respect to the euclidean inner product) of the eigenspaces $V(\lambda)$ of R, where λ is a (necessarily real) eigenvalue of R. Since R is positive definite, $\lambda > 0$ and $V(\lambda)$ is the eigenspace of R^a corresponding to the eigenvalue λ^a. As we saw above, if λ, μ are two eigenvalues of R and $\lambda\mu \neq 1$, then $V(\lambda)$ and $V(\mu)$ are ω-orthogonal. In particular, if $\lambda \neq 1$ then $V(\lambda)$ is isotropic. In case $\lambda\mu = 1$, we have

$$\omega(R^a(x), R^a(y)) = (\lambda\mu)^a\omega(x, y) = \omega(x, y)$$

for every $x \in V(\lambda)$ and $y \in V(\mu)$. Since every vector of $\mathbb{R}^{2n}$ is a sum of eigenvectors of R, the conclusion follows. □

Corollary 2.2. *The unitary group $U(n)$ is a strong deformation retract of $Sp(n, \mathbb{R})$. In particular, $Sp(n, \mathbb{R})$ is connected and the homogeneous space $Sp(n, \mathbb{R})/U(n)$ is contractible.*

Proof. Recalling the polar decomposition of a symplectic matrix A, we see that the map $r : \mathrm{Sp}(n, \mathbb{R}) \to U(n)$ defined by $r(A) = A(A^tA)^{-1/2}$ is a retraction. Also, $H : \mathrm{Sp}(n, \mathbb{R}) \times [0, 1] \to \mathrm{Sp}(n, \mathbb{R})$ defined by

$$H_s(A) = A(A^tA)^{-s/2}, \quad 0 \leq s \leq 1,$$

is a homotopy $H : id \simeq j \circ r \text{ rel } U(n)$, where $j : U(n) \hookrightarrow \mathrm{Sp}(n, \mathbb{R})$ denotes the inclusion. This homotopy rel $U(n)$ descends to a homotopy on $\mathrm{Sp}(n, \mathbb{R})/U(n)$ and so the latter is contractible. □

Proposition 2.3. *The unitary group $U(n)$ is a maximal compact subgroup of $Sp(n, \mathbb{R})$.*

Proof. Let G be a compact subgroup of $\mathrm{Sp}(n, \mathbb{R})$ and μ_G denote the Haar measure of G. Let

$$R = \int_G (A^tA) d\mu_G(A).$$

Since μ_G is right and left invariant, because G is compact, we have $B^tRB = R$ for all $B \in G$. As above, $S = R^{1/2}$ is symplectic and so $SBS^{-1} \in \mathrm{Sp}(n, \mathbb{R})$. Also, $SBS^{-1} \in O(2n, \mathbb{R})$, because

$$\langle SBS^{-1}(x), SBS^{-1}(y)\rangle = \langle BS^{-1}(x), RBS^{-1}(y)\rangle = \langle S^{-1}(x), B^tRBS^{-1}(y)\rangle$$

$$= \langle S^{-1}(x), RS^{-1}(y)\rangle = \langle S^{-1}(x), S(y)\rangle = \langle x, y\rangle$$

for every x, $y \in \mathbb{R}^{2n}$. Hence $SGS^{-1} \leq U(n)$, which proves the assertion. □

We recall that the group $SU(n)$ is simply connected. This can be proved by induction as follows. It is trivial for $n = 1$, since $SU(1)$ is the trivial group. Assume that $n > 1$ and $SU(n-1)$ is simply connected. The map $p : SU(n) \to S^{2n-1}$ which sends $A \in SU(n)$ to its first column is a fibration with fiber $SU(n-1)$. From the homotopy exact sequence of p we get an exact sequence

$$\{1\} = \pi_1(SU(n-1)) \to \pi_1(SU(n)) \to \pi_1(S^{2n-1}) = \{1\}$$

and hence $\pi_1(SU(n)) = \{1\}$.

Proposition 2.4. *The determinant map* $\det : U(n) \to S^1$ *induces an isomorphism on fundamental groups.*

Proof. The determinant map is a smooth submersion and therefore a fibration, since $U(n)$ is compact. Its fiber is $SU(n)$ and from the homotopy exact sequence we get the exact sequence

$$\{1\} = \pi_1(SU(n)) \to \pi_1(U(n)) \xrightarrow{(\det)_\sharp} \pi_1(S^1) \to \pi_0(SU(n)) = \{1\}. \quad \square$$

Corollary 2.5. $\pi_1(\mathrm{Sp}(n, \mathbb{R})) \cong \mathbb{Z}$. □

The symplectic group is related to the space of compatible complex structures of a symplectic vector space. The standard complex structure J belongs to $\mathcal{J}(\mathbb{R}^{2n}, \omega)$ and the corresponding positive definite inner product is the euclidean.

Let $I \in \mathcal{J}(\mathbb{R}^{2n}, \omega)$. If L a Lagrangian subspace, then $I(L)$ is a Lagrangian complement of L. Let $\{v_1, ..., v_n\}$ be an orthonormal basis of L with respect to the positive definite inner product which corresponds to I. Then $\{v_1, ..., v_n, I(v_1), ..., I(v_n)\}$ is an orthonormal (with respect to the same inner product) symplectic basis. Hence there exists $A \in GL(2n, \mathbb{R})$ such that $I = AJA^{-1}$.

The Lie group $\mathrm{Sp}(n, \mathbb{R})$ acts continuously on $\mathcal{J}(\mathbb{R}^{2n}, \omega)$ by conjugation. We shall use this action in order to prove essentialy that the latter is homeomorphic to the space of positive definite, symmetric, symplectic matrices.

Theorem 2.6. *The space* $\mathcal{J}(\mathbb{R}^{2n}, \omega)$ *is homeomorphic to the homogeneous space* $Sp(n, \mathbb{R})/U(n)$ *and is therefore contractible.*

Proof. Since the isotropy group of J is $\mathrm{Sp}(n, \mathbb{R}) \cap GL(n, \mathbb{C}) = U(n)$, it suffices to prove that the action of $\mathrm{Sp}(n, \mathbb{R})$ on $\mathcal{J}(\mathbb{R}^{2n}, \omega)$ by conjugation is transitive. Let $I \in \mathcal{J}(\mathbb{R}^{2n}, \omega)$ and let $A \in O(2n, \mathbb{R})$ be such that $I = AJA^{-1}$. Let $g(v, w) = \omega(v, I(w))$ be the corresponding positive definite inner product and $\tilde{g}$ the positive definite inner product which corresponds to J with respect to $A^*\omega$. Then

$$\tilde{g}(v, w) = (A^*\omega)(v, J(w)) = g(A(v), A(w))$$

for every v, $w \in \mathbb{R}^{2n}$. There exists $B \in GL(2n, \mathbb{R})$ such that

$$\tilde{g}(v, w) = \langle B(v), B(w) \rangle$$

for every v, $w \in \mathbb{R}^{2n}$. We can choose such B so that it commutes with J. Indeed, if $\{v_1, ..., v_n, J(v_1), ..., J(v_n)\}$ is a $\tilde{g}$-orthonormal and $A^*\omega$-symplectic basis and $\{u_1, ..., u_n, J(u_1), ..., J(u_n)\}$ is a $\langle , \rangle$-orthonormal and ω-symplectic basis, we can define B setting $B(v_j) = u_j$ and $B(J(v_j)) = J(u_j)$, $1 \leq j \leq n$, and then B has the desired properties. Now we have

$$\omega(BA^{-1}(v), BA^{-1}(w)) = \langle JBA^{-1}(v), BA^{-1}(w) \rangle = \langle BJA^{-1}(v), BA^{-1}(w) \rangle$$

$$= \langle BA^{-1}I(v), BA^{-1}(w) \rangle = \tilde{g}(A^{-1}I(v), A^{-1}(w)) = g(I(v), w) = \omega(v, w)$$

for every v, $w \in \mathbb{R}^{2n}$ and therefore $BA^{-1} \in \mathrm{Sp}(n, \mathbb{R})$. Since $(BA^{-1})I(BA^{-1})^{-1} = J$, this proves that $\mathrm{Sp}(n, \mathbb{R})$ acts transitively on $\mathcal{J}(\mathbb{R}^{2n}, \omega)$. $\square$

As a last issue in this section we shall discuss the set of Lagrangian subspaces. Let $L(n)$ denote the set of all Lagrangian linear subspaces of $\mathbb{R}^{2n}$. A n-dimensional subspace $W \leq \mathbb{R}^{2n}$ is Lagrangian if and only if ω is zero on W or equivalently $J(W)$ is orthogonal to W with respect to the euclidean inner product $\langle , \rangle$

Example 2.7. Let $A \in \mathbb{R}^{n \times n}$ and let $W = \{(x, A(x)) : x \in \mathbb{R}^n\}$ be the graph of A. Since W has dimension n, it is a Lagrangian subspace of $\mathbb{R}^{2n}$ if and only if it is isotropic or equivalently

$$0 = \langle (x, A(x)), J(y, A(y)) \rangle = \langle (x, A(x)), (-A(y), y) \rangle = -\langle x, A(y) \rangle + \langle A(x), y \rangle$$

for every x, $y \in \mathbb{R}^n$. Hence W is Lagrangian if and only if A is symmetric.

Let $W \leq \mathbb{R}^{2n}$ be a Lagrangian subspace. Let $B : W \to \mathbb{R}^n = \mathbb{R}^n \times \{0\}$ be an orthogonal isomorphism (with respect to the euclidean inner product) and let

$$A : W \oplus J(W) = \mathbb{R}^{2n} \to \mathbb{R}^{2n} = \mathbb{R}^n \oplus J(\mathbb{R}^n)$$

be defined by $A(v + J(u)) = B(v) + J(B(u))$. Then, $B = A|_W$ and $AJ = JA$. Also,

$$\omega(A(v + J(u)), A(v' + J(u'))) = \langle JB(v) - B(u), B(v') + JB(u') \rangle$$

$$= 0 + \langle v, u' \rangle - \langle u, v' \rangle - 0 = \omega(v + u, v' + u')$$

for every v, v', u, $u' \in W$, since J and B are orthogonal transformations. Thus, $A \in O(2n, \mathbb{R}) \cap GL(n, \mathbb{C}) = U(n)$ and $U(n)$ acts transitively on $L(n)$. The isotropy group of the Lagrangian subspace $\mathbb{R}^n \times \{0\}$ is the subgroup of $U(n)$ consisting of real matrices, that is $O(n, \mathbb{R})$. Hence $L(n) = U(n)/O(n, \mathbb{R})$ and so it has the structure of a homogeneous smooth manifold of dimension

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

For every $U \in U(n)$ and $A \in O(n, \mathbb{R})$ we have $\det(UA) = \det U \cdot \det A = \pm \det U$. So, we have a well defined smooth function $(\det)^2 : L(n) \to S^1$. Obviously, $(\det)^2$

is a submersion, and therefore a fibration, since $U(n)$ is compact. Let now U_1, $U_2 \in U(n)$ be such that $(\det U_1)^2 = (\det U_2)^2$. There exists $A \in O(n,\mathbb{R})$ such that $\det U_1 = \det(U_2 A)$ and therefore $U_1(U_2A)^{-1} \in SU(n)$ and

$$(U_1(U_2A)^{-1})U_2A \cdot O(n,\mathbb{R}) = U_1 \cdot O(n,\mathbb{R}).$$

Consequently, the group $SU(n)$ acts transitively on each fiber of $(\det)^2$ with isotropy group $SO(n,\mathbb{R})$. This shows that each fiber of $(\det)^2$ is diffeomorphic to the homogeneous space $SU(n)/SO(n,\mathbb{R})$, which is simply connected, since $SU(n)$ is simply connected and $SO(n,\mathbb{R})$ is connected, by the homotopy exact sequence of the fibration

$$SO(n,\mathbb{R}) \hookrightarrow SU(n) \to SU(n)/SO(n,\mathbb{R}).$$

From the homotopy exact sequence of the fibration

$$SU(n)/SO(n,\mathbb{R}) \hookrightarrow L(n) \stackrel{(\det)^2}{\longrightarrow} S^1$$

we have the exact sequence

$$\{1\} = \pi_1(SU(n)/SO(n,\mathbb{R})) \to \pi_1(L(n)) \to \mathbb{Z} \to \pi_0(SU(n)/SO(n,\mathbb{R})) = \{1\}.$$

It follows that $(\det)^2$ induces an isomorphism $\pi_1(L(n)) \cong \mathbb{Z}$, and therefore also an isomorphism on singular cohomology $((\det)^2)^* : \mathbb{Z} \cong H^1(L(n);\mathbb{Z})$. The cohomology class $((\det)^2)^*(1)$ is called the Maslov class.

2.3 Symplectic manifolds

A *symplectic manifold* is a pair (M,ω), where M is a smooth manifold and ω is a closed 2-form on M such that (T_pM,ω_p) is a symplectic vector space for every $p \in M$. Necessarily then M is even dimensional and if $\dim M = 2n$, then $\frac{1}{n!}\omega^n$ is a volume $2n$-form on M. So M is orientable and ω determines in this way an orientation. However, not every orientable, even-dumensional, smooth manifold admits a symplectic structure. If (M,ω) is a compact, symplectic manifold of dimension $2n$, then ω defines a real cohomology class $a = [\omega] \in H^2(M;\mathbb{R})$ and the cohomology class $a^n = a \cup \cdots \cup a \in H^{2n}(M;\mathbb{R})$ is represented by $\omega^n = \omega \wedge \cdots \wedge \omega$. So, $a^n \neq 0$ and the symplectic form ω cannot be exact. It follows that if M is an orientable, compact, smooth manifold such that $H^2(M;\mathbb{R}) = \{0\}$, then M admits no symplectic structure. For example, the n-sphere S^n cannot be symplectic for $n > 2$, as well as the 4-manifold $S^1 \times S^3$.

A smooth map $f : (M,\omega) \to (M',\omega')$ between symplectic manifolds is called *symplectic* if $f^*\omega' = \omega$. If f is also a diffeomorphism, it is called *symplectomorphism*. In this way symplectic manifolds form a category. The product of two symplectic manifolds (M_1,ω_1) and (M_2,ω_2) is the symplectic manifold

$$(M_1 \times M_2, \pi_1^*\omega_1 + \pi_2^*\omega),$$

where $\pi_j : M_1 \times M_2 \to M_j$, $j = 1,2$, are the projections.

Example 3.1. For every positive integer n, the space $\mathbb{R}^{2n}$ is a symplectic manifold, by considering on each tangent space $T_p\mathbb{R}^{2n} \cong \mathbb{R}^{2n}$ the canonical symplectic vector space structure. If dx^1, dx^2,..., dx^n, dy^1, dy^2,..., dy^n are the canonical basic differential 1-forms on $\mathbb{R}^{2n}$, then the canonical symplectic manifold structure is defined by the 2-form

$$\sum_{i=1}^{n} dx^i \wedge dy^i.$$

Example 3.2. Another simple example is the 2-shpere with its standard area 2-form ω given by the formula $\omega_x(u,v) = \langle x, u \times v\rangle$ for u, $v \in T_xS^2$ and $x \in S^2$, where $\times$ denotes the exterior product in $\mathbb{R}^3$. With this area 2-form the total area of S^2 is 4π. More generally, let $M \subset \mathbb{R}^3$ be an oriented surface. The Gauss map $N : M \to S^2$ associates to every $x \in M$ the outward unit normal vector $N(x) \perp T_xM$. Then, as in the case of S^2, the formula $\omega_x(u,v) = \langle N(x), u \times v\rangle$ for u, $v \in T_xM$ defines a symplectic 2-form on M.

Example 3.3. The basic example of a symplectic manifold is the cotangent bundle T^*M of any smooth n-manifold M with the symplectic 2-form $\omega = -d\theta$, where θ is the Liouville canonical 1-form on T^*M. Recall the locally ω is given by the formula

$$\omega|_{\text{locally}} = \sum_{i=1}^{n} dq^i \wedge dp_i$$

and compare with Example 3.1.

Another important example of a symplectic manifold is the complex projective space, which is an example of a Kähler manifold.

Example 3.4. For $n \geq 1$ let $\mathbb{C}P^n$ denote the complex projective space of complex dimension n and $\pi : \mathbb{C}^{n+1} \setminus \{0\} \to \mathbb{C}P^n$ be the quotient map. Recall that there is a canonical atlas $\{(V_j, \phi_j) : 0 \leq j \leq n\}$, where $V_j = \{[z_0, ..., z_n] \in \mathbb{C}P^n : z_j \neq 0\}$ and

$$\phi_j[z_0, ..., z_n] = (\frac{z_0}{z_j}, ..., \frac{z_{j-1}}{z_j}, \frac{z_{j+1}}{z_j}, ..., \frac{z_n}{z_j}).$$

The quotient map π is a submersion. To see this note first that $\phi_0 \circ \pi : \pi^{-1}(V_0) \to \mathbb{C}^n$ is given by the formula

$$(\phi_0 \circ \pi)(z_0, ..., z_n) = (\frac{z_1}{z_0}, ..., \frac{z_n}{z_0}).$$

Let $z = (z_0, ..., z_n) \in \pi^{-1}(V_0)$ and $v = (v_0, ..., v_n) \in T_z\mathbb{C}^{n+1} \cong \mathbb{C}^{n+1}$ be non-zero. Then $v = \dot{\gamma}(0)$, where $\gamma(t) = z + tv$, and

$$(\phi_0 \circ \pi \circ \gamma)(t) = \left(\frac{z_1 + tv_1}{z_0 + tv_0}, ..., \frac{z_n + tv_n}{z_0 + tv_0}\right)$$

so that

$$(\phi_0 \circ \pi \circ \gamma)'(0) = \left(\frac{v_1}{z_0} - \frac{z_1 v_0}{z_0^2}, ..., \frac{v_n}{z_0} - \frac{z_n v_0}{z_0^2}\right).$$

This implies that $v \in \text{Ker}\ \pi_{*z}$ if and only if $[v_0, ..., v_n] = [z_0, ..., z_n]$. In other words $\text{Ker}\ \pi_{*z} = \{\lambda z : \lambda \in \mathbb{C}\}$. Obviously, for every $(\zeta_0, ..., \zeta_n) \in \mathbb{C}^n$ there exists $v = (v_0, ..., v_n) \in \mathbb{C}^{n+1}$ such that

$$\zeta_j = \frac{v_j}{z_0} - \frac{z_j v_0}{z_0^2}.$$

Since similar things hold for any other chart (V_j, ϕ_j) instead of (V_0, ϕ_0), this shows that π is a submersion.

The inclusion $S^{2n+1} \hookrightarrow \mathbb{C}^{n+1} \setminus \{0\}$ is an embedding and so its derivative at every point of S^{2n+1} is a linear monomorphism. For every $z \in S^{2n+1}$ we have

$$\text{Ker}(\pi|_{S^{2n+1}})_{*z} = \text{Ker}\pi_{*z} \cap T_z S^{2n+1} = \{\lambda z : \lambda \in \mathbb{C} \text{ and } \text{Re}\lambda = 0\}$$

which is a real line. On the other hand, $\pi^{-1}(\pi(z)) \cap S^{2n+1}$ is the trace of the smooth curve $\sigma : \mathbb{R} \to S^{2n+1}$ with $\sigma(t) = e^{it} z$ for which $\sigma(0) = z$ and $\dot{\sigma}(0) = iz$. Therefore $\text{Ker}(\pi|_{S^{2n+1}})_{*z}$ is generated by $\dot{\sigma}(0)$.

Let h be the usual hermitian product on $\mathbb{C}^{n+1}$. If

$$W_z = \{\eta \in T_z \mathbb{C}^{n+1} : h(\eta, z) = 0\},$$

then $\pi_{*z}|_{W_z} : W_z \to T_{[z]}\mathbb{C}P^n$ is a linear isomorphism for every $z \in \mathbb{C}^{n+1} \setminus \{0\}$. Indeed, for every $v \in T_z \mathbb{C}^{n+1}$ there are unique $\lambda \in \mathbb{C}$ and $\eta \in W_z$ such that $v = \lambda z + \eta$. Obviously,

$$\lambda = \frac{h(v, z)}{h(z, z)}, \qquad \eta = v - \frac{h(v, z)}{h(z, z)} \cdot z.$$

The restricted hermitian product on W_z can be transfered isomorphically by π_{*z} on $T_{[z]}\mathbb{C}P^n$. If now

$$g_{[z]}(v, w) = \text{Re}\ h((\pi_{*z}|_{W_z})^{-1}(v), (\pi_{*z}|_{W_z})^{-1}(w))$$

for v, $w \in T_{[z]}\mathbb{C}P^n$, then g is Riemannian metric on $\mathbb{C}P^n$ called the *Fubini-Study metric*. If $z \in S^{2n+1}$, then $W_z = \{v \in T_z S^{2n+1} : \langle v, \dot{\sigma}(0) \rangle = 0\}$.

Each element $A \in U(n+1)$ induces a diffeomorphism $\tilde{A} : \mathbb{C}P^n \to \mathbb{C}P^n$. Moreover, $A(W_z) = W_{A(z)}$ for every $z \in \mathbb{C}^{n+1} \setminus \{0\}$ and therefore $\tilde{A}$ is an isometry of the Fubini-Study metric. In this way, $U(n+1)$ acts on $\mathbb{C}P^n$ by isometries. The action is transitive and so $\mathbb{C}P^n$ is a homogeneous Riemannian manifold with respect to the Fubibi-Study metric. Indeed, $U(n+1)$ acts transitively on S^{2n+1}, because if $z \in S^{2n+1}$, there exist $E_1, \ldots E_n \in \mathbb{C}^{n+1}$ such that $\{E_1, \ldots E_n, z\}$ is an h-orthonormal basis of $\mathbb{C}^{n+1}$. The matrix U with columns $E_1, \ldots, E_n$, z is an element of $U(n+1)$ such that $U(e_j) = E_j$ for $1 \leq j \leq n$ and $U(e_{n+1}) = z$. This last equality shows that $U(n+1)$ acts transitively on $\mathbb{C}P^n$.

The isotropy group of $[e_{n+1}] = [0, \ldots, 0, 1]$ consists of all $A \in U(n+1)$ such that $\lambda A(e_{n+1}) = e_{n+1}$ for some $\lambda \in S^1$. This means that

$$\lambda A = \begin{pmatrix} B & 0 \\ 0 & 1 \end{pmatrix}$$

for some $B \in U(n)$. Since $\tilde{A} = \widetilde{\lambda A}$, this implies that the isotropy group of $[e_{n+1}]$ is $U(n)$, considered as a subgroup of $U(n+1)$ as above, and therefore $\mathbb{C}P^n$ is diffeomorphic to the homogeneous space $U(n+1)/U(n)$.

If $A \in U(n+1)$, then $\det A \in S^1$ and taking $a \in S^1$ such that $a^n = \det A$ we have $a^{-1}A \in SU(n+1)$ and $\tilde{A} = \widetilde{a^{-1}A}$. Hence $SU(n+1)$ acts also transitively on $\mathbb{C}P^n$ and $\mathbb{C}P^n$ is diffeomorphic to $SU(n+1)/U(n)$, if we identify $U(n)$ with the subgroup of $SU(n+1)$ consisting of matrices of the form

$$\begin{pmatrix} B & 0 \\ 0 & \frac{1}{\det B} \end{pmatrix}$$

for $B \in U(n)$. If $A \in SU(n+1)$ belongs to the isotropy group of $[e_{n+1}]$ and λA has the above form, then $\det B = \lambda^{n+1}$ and putting $B' = \dfrac{1}{\lambda}B$, we have now

$$A = \begin{pmatrix} B' & 0 \\ 0 & \frac{1}{\lambda} \end{pmatrix}$$

where $\det B' = \lambda$. Therefore $A \in U(n)$, as a subgroup of $SU(n+1)$.

The scalar multiplication with the imaginary unit i defines a linear automorphism $J : W_z \to W_z$ such that $J^2 = -id$ and $h(Jv, Jw) = h(v, w)$ for every v, $w \in W_z$. Conjugating with $\pi_{*z}|_{W_z}$, we get a linear automorphism $J_{[z]}$ of $T_{[z]}\mathbb{C}P^n$ depending smoothly on $[z]$, which is a linear isometry, such that $J_{[z]}^2 = -id$. In other words, the Fubini-Study metric is a hermitian Riemannian metric.

If we set $\omega_{[z]}(v, w) = g_{[z]}(J_{[z]}v, w)$ for v, $w \in T_{[z]}\mathbb{C}P^n$, then

$$\omega_{[z]}(w, v) = g_{[z]}(J_{[z]}w, v) = g_{[z]}(v, J_{[z]}w) = -g_{[z]}(J_{[z]}v, w) = -\omega_{[z]}(v, w).$$

So we get a differential 2-form on $\mathbb{C}P^n$, which is obviously non-degenerate since $J_{[z]}$ is a linear isomorphism. To see that ω is symplectic, it remains to show that it is closed. This will be an application of the following general criterion.

Proposition 3.5. (Mumford's criterion) *Let M be a complex manifold and G be a group of diffeomorphisms of M which preserve the complex structure J of M and a complex hermitian metric h on M. If $J_z \in \rho_z(G_z)$ for every $z \in M$, where $\rho_z : G_z \to Aut_{\mathbb{C}}(T_zM)$ is the isotropic linear representation of the isotropy group G_z, then the 2-form $\omega(.,.) = \mathrm{Re}\ h(J.,.)$ is closed.*

Proof. Since every $g \in G$ leaves J and h invariant, it leaves ω invariant and therefore $d\omega$ also. For $g \in G_z$ and v, w, $u \in T_zM$ we have

$$(d\omega)_z(\rho_z(g)u, \rho_z(g)v, \rho_z(g)w) = (d\omega)_z(u, v, w)$$

because $\rho_z(g) = g_{*z}$. Since $J_z \in \rho_z(G_z)$, there exists $g \in G_z$ such that $J_z = \rho_z(g)$ and so

$$\begin{aligned}(d\omega)_z(u, v, w) = (d\omega)_z(J_zu, J_zv, J_zw) &= (d\omega)_z(J_z^2u, J_z^2v, J_z^2w) \\ &= -(d\omega)_z(u, v, w).\end{aligned}$$

Hence $d\omega = 0$. □

In the case of $\mathbb{C}P^n$ we apply Mumford's criterion for $G = SU(n+1)$, since $\mathbb{C}P^n$ is diffeomorphic to the homogeneous space $SU(n+1)/U(n)$. The isotropy group of

$[z] \in \mathbb{C}P^n$ is $G_{[z]} \cong U(W_z)$. Indeed, recall that $W_{e_{n+1}} = \mathbb{C}^n \times \{0\}$ and as we saw above there exists $A \in U(n+1)$ such that $A(e_{n+1}) = z$. Then, $G_{[z]} = \tilde{A}G_{[e_{n+1}]}\tilde{A}^{-1}$ and $A(W_z) = W_{e_{n+1}}$. Since every element of $G_{[z]}$ is a $\mathbb{C}$-linear, the isotropic linear representation $\rho_{[z]} : SU(n+1)_{[z]} \to U(n)$ is precisely the above group isomorphism $G_{[z]} \cong U(W_z)$. It follows that $J_{[z]} = iI_n \in U(W_z)$, because it is just multiplication by i, and by Mumford's criterion the non-degenerate 2-form $\omega_{[z]}(v,w) = g_{[z]}(J_{[z]}(v), w)$ for v, $w \in T_{[z]}\mathbb{C}P^n$, $z \in \mathbb{C}P^n$ is closed and hence symplectic.

This concludes the description of the symplectic structure of complex projective spaces. The complex projective space is an example of a Kähler manifold. The class of Kähler manifolds is important in Differential Geometry, Symplectic Geometry and Mathematical Physics. They will be descussed briefly in a subsequent section.

Having in mind the symplectic structure of the complex projective space we give the following.

Definition 3.6. An *almost symplectic structure* on a smooth manifold M of dimension $2n$ is non-degenerate, smooth 2-form on M. An *almost complex structure* on M is a smooth bundle endomorphism $J : TM \to TM$ such that $J^2 = -id$.

The following proposition leads to vector bundle obstructions for a compact manifold to be symplectic.

Proposition 3.7. *A smooth manifold M of dimension $2n$ has an almost complex structure if and only if it has an almost symplectic structure.*

Proof. Let J be an almost complex structure on M. Let g_0 be any Riemannian metric on M and g be the Riemannian metric defined by

$$g(v,w) = g_0(v,w) + g_0(Jv, Jw)$$

for v, $w \in T_pM$, $p \in M$. Then,

$$g(Jv,Jw) = g_0(Jv,Jw) + g_0(J^2v, J^2w) = g_0(Jv,Jw) + g_0(-v,-w) = g(v,w).$$

The smooth 2-form ω defined by

$$\omega(v,w) = g(Jv,w)$$

is non-degenerate, because $\omega(v, Jv) > 0$ for $v \neq 0$.

The proof of the converse is similar to the proof of Theorem 1.10, where now linear maps are replaced by vector bundle morphisms. Suppose that ω is an almost symplectic structure on M and let again g be any Riemannian metric on M. There exists a smooth bundle automorphism $A : TM \to TM$ (depending on g) such that $\omega(v,w) = g(Av,w)$ for all v, $w \in T_pM$, $p \in M$. Since ω is non-degenerate and skew-symmetric, A is an automorphism and skew-symmetric (with respect to g). Therefore, $-A^2$ is positive definite and symmetric (with respect to g). So, it has a unique square root, which means that there is a unique smooth bundle automorphism $B : TM \to TM$ such that $B^2 = -A^2$. Moreover, B commutes with

A. Then, $J = AB^{-1}$ is an almost complex structure on M. □

Remark 3.8. In the proof of the converse statement in Proposition 3.7 we have used the easily proved fact that if ω_t, $t \in \mathbb{R}$, is a family of symplectic bilinear forms on $\mathbb{R}^{2n}$, and g_t, $t \in \mathbb{R}$, is a family of positive definite inner products all depending smoothly on t, then following the proof of Theorem 1.10 we end up with a smooth family of corresponding compatible complex structures J_t, $t \in \mathbb{R}$, on $\mathbb{R}^{2n}$. This guarantees the smoothness of the almost complex structure J on M.

If (M, ω) is a symplectic manifold, an almost complex structure J on M is called compatible with ω if $g_x(u, v) = -\omega(J(u), v)$, for u, $v \in T_xM$, $x \in M$, is a Riemannian metric on M preserved by J. As the proof of Proposition 3.7 shows, any symplectic manifold carries compatible almost complex structures. As we commented after the proof of Theorem 1.10, if J_0 and J_1 are two almost complex structures compatible with ω, there exists a smooth family J_t, $0 \leq t \leq 1$, of compatible almost complex structures from J_0 to J_1. Actually the arguments used to prove Theorem 2.6 can be globalized to prove that the space $\mathcal{J}(M, \omega)$ of compatible almost complex structures is contractible. This is important for uniqueness of invariants arising from a compatible almost complex structure.

Example 3.9. Let $M \subset \mathbb{R}^3$ be an oriented surface with Gauss map $N : M \to S^2$. An almost complex structure J on M can be defined by the formula

$$J_x(v) = N(x) \times v$$

for $v \in T_xM$, $x \in M$, since

$$J_x^2(v) = N(x) \times (N(x) \times v) = \langle N(x), v\rangle N(x) - \langle N(x), N(x)\rangle v = -v.$$

Recall from Example 3.2 that the symplectic 2-form ω on M is defined by

$$\omega_x(u, v) = \langle N(x), u \times v\rangle = \langle N(x) \times u, v\rangle = \langle J_x(u), v\rangle.$$

So J is compatible with ω.

Example 3.10. Not every almost complex manifold is symplectic. We shall construct an almost complex structure on the 6-sphere S^6 using the exterior product in $\mathbb{R}^7$, defined from the Cayley algebra of octonions, in the same way as we did in the previous Example 3.9. As we know S^6 carries no symplectic structure. It is still unknown whether S^6 can be made a complex manifold.

As a vector space the (non-associative) Cayley algebra of the octonions is isomorphic to $\mathbb{R}^8$. Each octonion $a = (a_r, a_0, a_1, a_2, a_3, a_4, a_5, a_6)$ can be written as

$$a = a_r \cdot 1 + \sum_{j=0}^{6} a_j e_j$$

where $\{e_0, e_1, e_2, e_3, e_4, e_5, e_6\}$ is the canonical basis of $\mathbb{R}^7$. The first component a_r is the real part and can be considered a real number, and the second is the imaginary part, which is a vector $\vec{a} \in \mathbb{R}^7$.

The multiplication is defined as

$$(a_r + \vec{a}) \cdot (b_r + \vec{b}) = a_r b_r - \langle \vec{a}, \vec{b} \rangle + a_r \vec{b} + b_r \vec{a} + \sum_{i \neq j} a_i b_j e_i \cdot e_j$$

where $\langle , \rangle$ is the euclidean inner product and $e_i \cdot e_j$ is given by the following multiplication table

	e_0	e_1	e_2	e_3	e_4	e_5	e_6
e_0	-1	e_2	$-e_1$	e_4	$-e_3$	e_6	$-e_5$
e_1	$-e_2$	-1	e_0	$-e_5$	e_6	e_3	$-e_4$
e_2	e_1	$-e_0$	-1	e_6	e_5	$-e_4$	$-e_3$
e_3	$-e_4$	e_5	$-e_6$	-1	e_0	$-e_1$	e_2
e_4	e_3	$-e_6$	$-e_5$	$-e_0$	-1	e_2	e_1
e_5	$-e_6$	$-e_3$	e_4	e_1	$-e_2$	-1	e_0
e_6	e_5	e_4	e_3	$-e_2$	$-e_1$	$-e_0$	-1

If $\vec{a}$, $\vec{b} \in \mathbb{R}^7$ are considered as purely imaginary octonions, then

$$\vec{a} \cdot \vec{b} = -\langle \vec{a}, \vec{b} \rangle + \sum_{i \neq j} a_i b_j e_i \cdot e_j.$$

Letting the imaginary part be

$$\vec{a} \times \vec{b} = \sum_{i \neq j} a_i b_j e_i \times e_j$$

where $e_i \times e_j = e_i \cdot e_j$ for $i \neq j$ and $e_i \times e_i = 0$, we get a skew-symmetric, bilinear product $\times : \mathbb{R}^7 \times \mathbb{R}^7 \to \mathbb{R}^7$ with the additional properties:

(i) $\langle \vec{a} \times \vec{b}, \vec{c} \rangle = \langle \vec{a}, \vec{b} \times \vec{c} \rangle$ and

(ii) $\vec{a} \times (\vec{b} \times \vec{c}) + (\vec{a} \times \vec{b}) \times \vec{c} = 2\langle \vec{a}, \vec{c} \rangle \vec{b} - \langle \vec{b}, \vec{c} \rangle \vec{a} - \langle \vec{b}, \vec{a} \rangle \vec{c}$

for every $\vec{a}$, $\vec{b}$, $\vec{c} \in \mathbb{R}^7$. From (ii) follows that $\vec{a} \times \vec{b}$ is orthogonal to both $\vec{a}$, $\vec{b}$ and

(iii) $\vec{a} \times (\vec{a} \times \vec{b}) = \langle \vec{a}, \vec{b} \rangle \vec{a} - \langle \vec{a}, \vec{a} \rangle \vec{b}$.

Identifying $T_x S^6$ with a linear subspace of $\mathbb{R}^6$ for every $x \in S^6$ as usual, we can define $J_x : T_x S^6 \to T_x S^6$ to be the linear automorphism given by $J_x(v) = x \times v$ and then

$$J_x^2(v) = x \times (x \times v) = -\langle x, x \rangle v = -v.$$

Since J depends smoothly on $x \in S^6$, it is an almost complex structure on S^6.

The tangent bundle of a symplectic manifold (M, ω), or more generally of an almost complex manifold, is the realification of a complex vector bundle. This situation is a particular case of a symplectic vector bundle. A symplectic vector bundle (E, ω) over a smooth manifold M is a real smooth vector bundle $p : E \to M$ with a symplectic form ω_x on each fiber $E_x = p^{-1}(x)$ which varies smoothly with $x \in M$. In other words, ω is a smooth section of the bundle $E^* \wedge E^*$, where E^* is the dual bundle. Two symplectic vector bundles (E_1, ω_1) and (E_2, ω_2) over M are isomorphic if there exists a smooth vector bundle isomorphism $f : E_1 \to E_2$ such that $f^*(\omega_2) = \omega_1$.

Let (E,ω) be a $2n$-dimensional symplectic vector bundle over M and let $U \subset M$ be an open set over which the bundle is trivial. This is equivalent to saying that there exist smooth sections $e_1,\ e_2,\dots,\ e_{2n} : U \to p^{-1}(U) \subset E$ such that the set $\{e_1(x),\dots,e_{2n}(x)\}$ is a basis of the fiber E_x for every $x \in U$. If $\{e_1^*(x),\dots,e_{2n}^*(x)\}$ is the dual basis of E_x^*, there are smooth functions $a_{ij} : U \to \mathbb{R}$, $1 \le i < j \le 2n$, such that

$$\omega|_{p^{-1}(U)} = \sum_{i<j} a_{ij} e_i^* \wedge e_j^*.$$

Let $x_0 \in U$. Since $\omega_{x_0} \neq 0$, we may assume that $a_{12}(x_0) \neq 0$ and so $a_{12}(x) \neq 0$ for x in a smaller neighbourhood of x_0. Continuing now in the same way as in the proof of Cartan's Lemma 1.1, we end up with an open neighbourhood V of x_0 such that $(p^{-1}(V), \omega|_{p^{-1}(V)})$ is isomorphic to $V \times \mathbb{R}^{2n}$ equiped with the standard symplectic form on the fiber $\mathbb{R}^{2n}$. This implies that the structure group of every symplectic vector bundle (E,ω) can be reduced from $GL(2n,\mathbb{R})$ to the symplectic group $Sp(n,\mathbb{R})$. Since $Sp(n,\mathbb{R})$ is connected and $U(n)$ is a maximal compact subgroup of $Sp(n,\mathbb{R})$, it follows from the Iwasawa decomposition and the reduction theorem for fiber bundles that the structure group of a symplectic vector bundle can be further reduced to $U(n)$. Hence every symplectic vector bundle is a complex vector bundle. As in the case of symplectic vector spaces, one can define compatible complex structures on (E,ω) and the corresponding space $\mathcal{J}(E,\omega)$. The same arguments show that $\mathcal{J}(E,\omega)$ is not empty and contractible. So every symplectic vector bundle has a complex structure which is well-defined up to homotopy.

Returning to the case of a symplectic manifold (M,ω) of dimension $2n$, the tangent bundle of M can be considered as a complex vector bundle of complex dimension n to which correspond Chern classes $c_k \in H^{2k}(M;\mathbb{Z})$, $1 \le k \le n$. The Chern classes are related to the Pontryagin classes of the tangent bundle of M through polynomial (quadratic) equations, which can serve as obstructions to the existence of a symplectic structure on M, since not every compact, orientable, smooth $2n$-manifold has cohomology classes satisfying these equations. For instance, using these equations and Hirzebruch's Signature Theorem, one can show that the connected sum $\mathbb{C}P^2 \# \mathbb{C}P^2$ cannot be a symplectic manifold.

2.4 Local description of symplectic manifolds

Even though we have defined the symplectic structure in analogy to the Riemannian structure, their local behaviour differs drastically. In this section we shall show that in the neighbourhood of any point on a symplectic $2n$-manifold (M,ω) there are suitable local coordinates $(q^1,...,q^n,p_1,...,p_n)$ such that

$$\omega|_{\text{locally}} = \sum_{i=1}^n dq^i \wedge dp_i.$$

This shows that in symplectic geometry there are no local invariants, in contrast to Riemannian geometry, where there are highly non-trivial local invariants. In other words, the study of symplectic manifolds is of global nature and one expects to use mainly topological methods.

The method of proof of the local isomorphy of all symplectic manifolds, we shall present, is based on Moser's trick.

Lemma 4.1. (Moser) *Let M and N be two smooth manifolds and $F : M \times \mathbb{R} \to N$ be a smooth map. For every $t \in \mathbb{R}$ let $X_t : M \to TN$ be the smooth vector field along $F_t = F(.,t)$ defined by*

$$X_t(p) = \frac{\partial}{\partial s}\Big|_{s=t} F(p,s) \in T_{F_t(p)}N.$$

If $(\omega_t)_{t\in\mathbb{R}}$ is a smooth family of k-forms on N, then

$$\frac{d}{dt}(F_t^*\omega_t) = F_t^*(\frac{d\omega_t}{dt} + i_{X_t}d\omega_t) + d(F_t^* i_{X_t}\omega_t).$$

If moreover F_t is a diffeomorphism for every $t \in \mathbb{R}$, then

$$\frac{d}{dt}(F_t^*\omega_t) = F_t^*(\frac{d\omega_t}{dt} + i_{X_t}d\omega_t + di_{X_t}\omega_t).$$

Note that if F_t is not a diffeomorphism then X_t is not in general a vector field on N. The meaning of the symbol $F_t^* i_{X_t}\omega_t$ will be clear in the proof.

Proof. (a) First we shall prove the formula in the special case $M = N = P \times \mathbb{R}$ and $F_t = \psi_t$, where $\psi_t(x,s) = (x, s+t)$. Then

$$\omega_t = ds \wedge a(x,s,t)dx^k + b(x,s,t)dx^{k+1},$$

where

$$a(x,s,t)dx^k = \sum_{i_1<i_2<...<i_k} a_{i_1 i_2 ... i_k}(x,s,t)dx^{i_1} \wedge dx^{i_2} \wedge ... \wedge dx^{i_k}$$

and similarly for $b(x,s,t)dx^{k+1}$. So $\psi_t^*\omega_t = ds \wedge a(x,s+t,t)dx^k + b(x,s+t,t)dx^{k+1}$ and

$$\frac{d}{dt}(\psi_t^*\omega_t) = ds \wedge \frac{\partial a}{\partial s}(x,s+t,t)dx^k + \frac{\partial b}{\partial s}(x,s+t,t)dx^{k+1}$$

$$+ds \wedge \frac{\partial a}{\partial t}(x,s+t,t)dx^k + \frac{\partial b}{\partial t}(x,s+t,t)dx^{k+1},$$

Obviously,

$$\psi_t^*(\frac{d\omega_t}{dt}) = ds \wedge \frac{\partial a}{\partial t}(x,s+t,t)dx^k + \frac{\partial b}{\partial t}(x,s+t,t)dx^{k+1}. \tag{1}$$

On the other hand $X_t = \frac{\partial}{\partial s}$. So $i_{X_t}\omega_t = a(x,s,t)dx^k$ and

$$d(i_{X_t}\omega_t) = \sum_{i_1<i_2<...<i_k} da_{i_1 i_2 ... i_k}(x,s,t) \wedge dx^{i_1} \wedge dx^{i_2} \wedge ... \wedge dx^{i_k} =$$

$$\sum_{i_1<i_2<...<i_k} \Big(\frac{\partial a_{i_1 i_2 ... i_k}}{\partial s}(x,s,t)ds \wedge dx^{i_1} \wedge dx^{i_2} \wedge ... \wedge dx^{i_k} +$$

$$\sum_{j\notin\{i_1<i_2<...<i_k\}} \frac{\partial a_{i_1 i_2 \dots i_k}}{\partial x^j} dx^j \wedge dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_k} \Bigg).$$

We shall write for brevity

$$d(i_{X_t}\omega_t) = \frac{\partial a}{\partial s}(x, s, t) ds \wedge dx^k + d_x a(x, s, t) dx^{k+1}.$$

So

$$\psi_t^*(d(i_{X_t}\omega_t)) = \frac{\partial a}{\partial s}(x, s+t, t) ds \wedge dx^k + d_x a(x, s+t, t) dx^{k+1}. \tag{2}$$

Using the symbol d_x in the same way, we have

$$d\omega_t = -ds \wedge d_x a(x, s, t) dx^k + \frac{\partial b}{\partial s}(x, s, t) ds \wedge dx^{k+1} + d_x b(x, s, t) dx^{k+2},$$

and thus

$$\psi_t^*(i_{X_t} d\omega_t) = -d_x a(x, s+t, t) dx^{k+1} + \frac{\partial b}{\partial s}(x, s+t, t) dx^{k+1}. \tag{3}$$

Summing up now (1), (2) and (3) we get

$$\psi_t^*(\frac{d\omega_t}{dt}) + \psi_t^* d(i_{X_t}\omega_t) + \psi_t^*(i_{X_t} d\omega_t) = \frac{d}{dt}\psi_t^*\omega_t.$$

(b) The general case follows from part (a) using the decomposition $F_t = F \circ \psi_t \circ j$, where $j : M \to M \times \mathbb{R}$ is the inclusion $j(p) = (p, 0)$ and ψ_t is the same as in part (a). Now we have

$$X_t(p) = \frac{\partial}{\partial s}\bigg|_{s=t} F(p, s) = F_{*(p,t)}\left(\frac{\partial}{\partial s}\right)_{(p,t)}.$$

If each F_t is a diffeomorphism, then X_t is a vector field on N and $i_{X_t}\omega_t$ is defined. If not, the term $F_t^*(i_{X_t}\omega_t)$ has the following meaning. By definition,

$$F_t^*(i_{X_t}\omega_t)_p(v_1, ..., v_{k-1}) = (\omega_t)_{F_t(p)}(X_t(p), (F_t)_{*p}(v_1), ..., (F_t)_{*p}(v_{k-1})) =$$

$$(\omega_t)_{F(p,t)}(F_{*(p,t)}\left(\frac{\partial}{\partial s}\right)_{(p,t)}, F_{*(p,t)}(v_1, 0), ..., F_{*(p,t)}(v_{k-1}, 0)) =$$

$$(F^*\omega_t)_{(p,t)}(\left(\frac{\partial}{\partial s}\right)_{(p,t)}, (v_1, 0), ..., (v_{k-1}, 0)) = (i_{\partial/\partial s}F^*\omega_t)_{(p,t)}((v_1, 0), ..., (v_{k-1}, 0)) =$$

$$j^*\psi_t^*(i_{\partial/\partial s}F^*\omega_t)_p(v_1, ..., v_{k-1})$$

for v_1,...,$v_{k-1} \in T_pM$. Therefore, $F_t^*(i_{X_t}\omega_t) = j^*\psi_t^*(i_{\partial/\partial s}F^*\omega_t)$ and similarly $F_t^*(i_{X_t}d\omega_t) = j^*\psi_t^*(i_{\partial/\partial s}d(F^*\omega_t))$. Since j^* does not depend on t, we have

$$\frac{d}{dt}(F_t^*\omega_t) = j^*\frac{d}{dt}(\psi_t^*F^*\omega_t)$$

and applying part (a) to $F^*\omega_t$ we get

$$\frac{d}{dt}(F_t^*\omega_t) = j^*\psi_t^*(\frac{d(F^*\omega_t)}{dt}) + j^*\psi_t^*(i_{\partial/\partial s}d(F^*\omega_t)) + j^*d(\psi_t^* i_{\partial/\partial s}(F^*\omega_t)) =$$

$$j^*\psi_t^*F^*(\frac{d\omega_t}{dt}) + F_t^*(i_{X_t}d\omega_t) + d(F_t^*(i_{X_t}\omega_t)) =$$

$$F_t^*(\frac{d\omega_t}{dt}) + F_t^*(i_{X_t}d\omega_t) + d(F_t^*(i_{X_t}\omega_t)). \quad \square$$

Corollary 4.2. *Let X be a smooth vector field on a smooth manifold M. If ω is a differential form on M, then $L_X\omega = i_X d\omega + di_X\omega$.*

Proof. If X is complete and $(\phi_t)_{t\in\mathbb{R}}$ is its flow, we apply Lemma 4.1 for $F_t = \phi_t$, $M = N$ and $\omega_t = \omega$ and we have

$$L_X\omega = \frac{d}{dt}\Big|_{t=0} \phi_t^*\omega = i_X d\omega + di_X\omega.$$

If X is not complete, then M has an open covering $\mathcal{U}$ such that for every $U \in \mathcal{U}$ there exists some $\epsilon > 0$ and a local flow map $\phi : (-\epsilon, \epsilon) \times U \to M$ of X. Again we apply Lemma 4.1 for $F_t = \phi_t$ on U this time to get the desired formula on every $U \in \mathcal{U}$, hence on M. $\square$

We are now in a position to prove the main theorem of this section.

Theorem 4.3. (Darboux) *Let ω_0 and ω_1 be two symplectic 2-forms on a smooth $2n$-manifold M and $p \in M$. If $\omega_0(p) = \omega_1(p)$, there exists an open neighbourhood U of p in M and a diffeomorphism $F : U \to F(U) \subset M$, where $F(U)$ is an open neighbourhood of p, such that $F(p) = p$ and $F^*\omega_1 = \omega_0$.*

Proof. Let $\omega_t = (1-t)\omega_0 + t\omega_1$, $0 \le t \le 1$. Since $\omega_t(p) = \omega_0(p) = \omega_1(p)$, there exists an open neighbourhood U_1 of p diffeomorphic to $\mathbb{R}^{2n}$ such that $\omega_t|_{U_1}$ is symplectic for every $0 \le t \le 1$. By the lemma of Poincaré, there exists a 1-form a on U_1 such that $\omega_0 - \omega_1 = da$ on U_1 and $a(p) = 0$. For every $0 \le t \le 1$ there exists a smooth vector field Y_t on U_1 such that $i_{Y_t}\omega_t = a$. Obviously, $Y_t(p) = 0$ and the above hold for every $-\epsilon < t < 1 + \epsilon$, for some $\epsilon > 0$. Now $\bar{Y} = (\frac{\partial}{\partial s}, Y_s)$ is a smooth vector field on $(-\epsilon, 1+\epsilon) \times U_1$. If ϕ_t is the flow of $\bar{Y}$, then $\phi_t(s, x) = (s + t, f_t(s, x))$, for some smooth $f_t : (-\epsilon, 1+\epsilon) \times U_1 \to M$. Therefore, $\phi_t(0, x) = (t, F_t(x))$, where $F_t : U_1 \to F_t(U_1)$ is a diffeomorphism. Since $\phi_t(0, p) = (0, p)$, that is $F_t(p) = p$, there exists an open neighbourhood U of p such that F_t is defined on U and $F_t(U) \subset U_1$ for every $0 \le t \le 1$. Obviously, $Y_t = \frac{\partial F_t}{\partial t}$ and so from Lemma 4.1 we have

$$\frac{d}{dt}(F_t^*\omega_t) = F_t^*(\frac{d\omega_t}{dt} + i_{Y_t}d\omega_t + di_{Y_t}\omega_t) = F_t^*(\omega_1 - \omega_0 + 0 + da) = 0.$$

Hence $F_t^*\omega_t = F_0^*\omega_0 = \omega_0$ for every $0 \leq t \leq 1$, since $F_0 = id$. □

Corollary 4.4. *Let (M, ω) be a symplectic $2n$-manifold and $p \in M$. There exists an open neighbourhood U of p and a diffeomorphism $F : U \to F(U) \subset \mathbb{R}^{2n}$ such that*

$$\omega|_U = F^*\big(\sum_{i=1}^{n} dx^i \wedge dy^i\big).$$

Proof. Let (W, ψ) be a chart of M with $p \in W$, $\psi(W) = \mathbb{R}^{2n}$ and $\psi(p) = 0$. Then the 2-form $\omega_1 = (\psi^{-1})^*\omega$ on $\mathbb{R}^{2n}$ is symplectic. Composing with a linear transformation if necessary, we may assume that $\omega_1(0) = \omega_0(0)$, where ω_0 is the standard symplectic 2-form on $\mathbb{R}^{2n}$. By Darboux's theorem, there exists an open neighbourhood V of 0 in $\mathbb{R}^{2n}$ and a diffeomorphism $\phi : V \to \phi(V)$ with $\phi(0) = 0$ and $\phi^*\omega_1 = \omega_0$. It suffices to set now $F = (\psi^{-1} \circ \phi)^{-1}$. □

At this point we cannot resist the temptation to use Moser's trick in order to prove the following result, also due to J. Moser.

Theorem 4.5. (Moser) *Let M be a connected, compact, oriented, smooth n-manifold and ω_0, ω_1 be two representatives of the orientation. If*

$$\int_M \omega_0 = \int_M \omega_1,$$

there exists a diffeomorphism $f : M \to M$ such that $f^*\omega_1 = \omega_0$.

Proof. For every $0 \leq t \leq 1$ the n-form $\omega_t = (1-t)\omega_0 + t\omega_0$ is a representative of the orientation, that is a positive volume element of M. Since

$$\int_M (\omega_0 - \omega_1) = 0,$$

there exists a $(n-1)$-form a on M such that $\omega_0 - \omega_1 = da$. There exists a unique smooth vector field X_t on M such that $i_{X_t}\omega_t = a$. As in the proof of Darboux's theorem, there exists a smooth isotopy $F : M \times [0,1] \to M$ with $F_0 = id$ and

$$X_t = \frac{\partial F_t}{\partial t},$$

because M is compact. Again from Lemma 4.1 we have

$$\frac{d}{dt}(F_t^*\omega_t) = F_t^*(\omega_1 - \omega_0 - 0 + da) = 0.$$

Hence $F_t^*\omega_t = \omega_0$ for every $0 \leq t \leq 1$. □

2.5 Lagrangian submanifolds

Let (M, ω) be a symplectic manifold and $j : L \to M$ be an immersion. If $j^*\omega = 0$, then L is called an isotropic immersed submanifold of M. In other words, $j_{*x}(T_xL)$ is an isotropic linear subspace of $T_{j(x)}M$ for every $x \in L$. Analogously, L is called a coisotropic (respectively symplectic) immersed submanifold if $j_{*x}(T_xL)$ is coisotropic (respectively symplectic) for every $x \in L$. The same terms are used for (embedded) submanifolds of M and for subbundles of TM restricted to submanifolds of M with the obvious definitions.

A submanifold $L \subset M$ is called Lagrangian if it is isotropic and there exists an isotropic subbundle E of $TM|_L$ such that $TM|_L = TL \oplus E$.

Proposition 5.1. *If (M, ω) is a symplectic manifold and $L \subset M$ is a submanifold of M, then L is Lagrangian if and only if L is isotropic and* $\dim L = \dfrac{1}{2} \dim M$.

Proof. The direct assertion is obvious. For the converse, let L be isotropic of dimension half the dimension of M. Then T_xL has an isotropic complement in T_xM for every $x \in L$. There exists an compatible almost complex structure J on M and so $g_x(u, v) = -\omega_x(J_x(u), v)$, u, $v \in T_xM$, $x \in M$, is a Riemannian metric on M. If $E = J(TL)$, then E is a complementaty to TL smooth isotropic subbundle of $TM|_L$. □

For example, the real projective space $\mathbb{R}P^n$ is a Lagrangian submanifold of the complex projective space $\mathbb{C}P^n$ with its standard symplectic structure.

Example 5.2. Recall that the standard symplectic 2-form on the cotangent bundle T^*M of a smooth manifold M is defined as $\omega = -d\theta$, where θ is the Liouville canonical 1-form defined by $\theta_a(v) = a(\pi_{*a}(v))$ for $v \in T_aT^*M$, $a \in T^*M$,, where $\pi : T^*M \to M$ is the cotangent bundle projection. We observe that if a is any smooth 1-form on M, then $a^*\theta = a$. Indeed, for every $v \in T_xM$, $x \in M$ we have

$$(a^*\theta)_x(v) = a_x(\pi_{*a_x}(a_{*x}(v))) = a_x((\pi \circ a)_{*x}(v)) = a_x(v).$$

It follows that $da = a^*(d\theta) = -a^*\omega$.

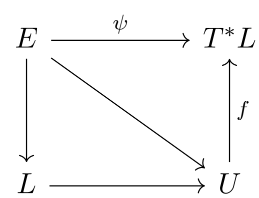
Let now $L = a(M) = \{(x, a_x) : x \in M\} \subset T^*M$ be the graph of a. It is clearly a submanifold of T^*M diffeomorphic to M of dimension half the dimension of T^*M. Since $a^*\omega = -da$, we conclude that L is a Lagrangian submanifold of T^*M if and only if a is closed. If a is closed and $j : L \to T^*M$ is the inclusion, then $d(j^*\theta) = 0$. Since a maps M diffeomorphically onto L, this implies that M is covered by open sets $U \subset M$ for which thare are smooth functions $f_U : U \to \mathbb{R}$ such that $a|_U = df_U$. Each such function f_U is called a generating function for the Lagrangian submanifold L. In the trivial case $a = 0$, we have that M is a Lagrangian submanifold of T^*M considered as the zero-section.

The main theorem of this section is a result of A. Weinstein and B. Kostant, which reduces the local studies in a neighbourhood of a Lagrangian submanifold to the case of the zero-section of its cotangent bundle. It can be considered a generalization of Darboux's theorem.

First we shall need the following.

Lemma 5.3. *Let (M,ω) be a symplectic manifold and $L \subset M$ be a Lagrangian submanifold. There are open neighbourhoods U of L in M and V of L in T^*L for which there exists a diffeomorphism $f : U \to V$ such that $f|_L = id$ and $f^*(-d\theta|_L) = \omega|_L$, where θ is the Liouville canonical 1-form on T^*L.*

Proof. Let E be an isotropic complement of TL in $TM|_L$. Let $\psi : E \to T^*L$ be the vector bundle morphism defined by $\psi(v)(u) = \omega(u,v)$. If $\psi(v) = 0$, then $v \in (TL)^\perp = TL$ and so $v \in E \cap TL = \{0\}$. This implies that ψ is an isomorphism of vector bundles over L. Since E is complementary to TL, ψ induces a diffeomorphism $f : U \to V \subset T^*L$ of an open neighbourhood U of L in M onto an open neighbourhood V of L in T^*L, by the tubular neighbourhood theorem.



Obviously, $f_*|_{TL} = id$ and $f_*|_E = \psi$. Taking into account the splittings $TM|_L = TL \oplus E$ and $T(T^*L)|_L \cong TL \oplus T^*L$ we have

$$f^*(-d\theta)((v_1,v_2),(w_1,w_2)) = -d\theta((v_1,\psi(v_2)),(w_1,\psi(w_2)))$$

$$= \psi(w_2)(v_1) - \psi(v_2)(w_1) = \omega(v_1,w_2) - \omega(w_1,v_2) = \omega((v_1,v_2),(w_1,w_2)),$$

for every (v_1,v_2), $(w_1,w_2)) \in TL \oplus E$, since L is Lagrangian. □

Theorem 5.4. (Weinstein) *Let M be a smooth manifold and $L \subset M$ be an embedded submanifold. Let ω_0, ω_1 be two symplectic 2-forms on M such that $\omega_0|_L = \omega_1|_L$, meaning that $\omega_0(v,w) = \omega_1(v,w)$ for every v, $w \in T_xM$ and $x \in L$. Then, there esists an open neighbourhood U of L in M and a diffeomorphism $f : U \to f(U) \subset M$ onto an open neighbourhood $f(U)$ of L in M such that $f|_L = id$ and $f^*\omega_1 = \omega_0$.*

Proof. By the tubular neighbourhood theorem, we may assume that there exists a smooth strong deformation retraction $\phi : M \times [0,1] \to M$, that is $\phi_0 : M \to L \subset M$ is a smooth retraction of M onto L, $\phi_1 = id$ and $\phi_t(x) = x$ for all $x \in L$ and $0 \le t \le 1$, where $\phi_t = \phi(.,t)$. For any smooth k-form σ on M we have

$$\sigma - \phi_0^*\sigma = \int_0^1 \frac{d}{dt}(\phi_t^*\sigma)dt = \int_0^1 \phi_t^*(i_{X_t}d\sigma)dt + d\int_0^1 \phi_t^*(i_{X_t}\sigma)dt,$$

by Moser's Lemma 4.1, where $X_t = \left.\frac{\partial}{\partial s}\right|_{s=t} \phi(.,s)$.

If we put now

$$I(\eta) = \int_0^1 \phi_t^*(i_{X_t}\eta)dt$$

for every smooth form η on M, we get

$$\sigma - \phi_0^*\sigma = I(d\sigma) + dI(\sigma).$$

In other words, I is a cochain homotopy of the deRham cochain complex of M into itself between the identity and the induced cochain map by ϕ_0.

Now we set $\sigma = \omega_1 - \omega_0$ and $\omega_t = \omega_0 + t\sigma = (1-t)\omega_0 + t\omega_1$. By assumption, $\sigma|_L = 0$ and so $\phi_0^*\sigma = 0$. Also, $\sigma = dI(\sigma)$, since σ is closed. Furthermore, $X_t(x) = 0$ for $x \in L$, because $\phi_t(x) = x$ for all $0 \leq t \leq 1$ and so $I(\sigma)|_L = 0$. Obviously, $\omega_t|_L = \omega_0|_L = \omega_1|_L$ and $\omega_t|_L$ is non-degenerate for all $0 \leq t \leq 1$. By compactness of $[0,1]$, there is an open neighbourhood of L on which ω_t is non-degenerate for all $0 \leq t \leq 1$. On this neighbourhood there exists a smooth 1-parameter family of smooth vector fields Z_t, $0 \leq t \leq 1$, such that

$$i_{Z_t}\omega_t = -I(\sigma)$$

and $Z_t|_L = 0$. There exists a smaller neighbourhood of L on which the flow of each Z_t is defined at least on the interval $[-2,2]$. Thus, the time-1-map f_t of the flow of Z_t is defined in this neighbourhood of L and $f_t|_L = id$ for all $0 \leq t \leq 1$. Now we have

$$f_1^*\omega_1 - f_0^*\omega_0 = \int_0^1 \frac{d}{dt}(f_t^*\sigma)dt = \int_0^1 [f_t^*\sigma + df_t^*(i_{Z_t}\omega_t)]dt$$

$$= \int_0^1 f_t^*(\sigma + di_{Z_t}\omega_t)dt = \int_0^1 f_t^*(\sigma - dI(\sigma))dt = 0.$$

Consequently, $\omega_0 = (f_1 \circ (f_0)^{-1})^*\omega_1$ and if we put $f = f_1 \circ f_0^{-1}$, then $f|_L = id$ and $f^*\omega_1 = \omega_0$. □

Corollary 5.5. (Kostant) *Let (M,ω) be a symplectic manifold and $L \subset M$ be a Lagrangian submanifold. There exists an open neighbourhood U of L in M and a diffeomorphism $h : U \to V$ of U onto an open neighbourhood V of L in T^*L such that $h|_L = id$ and $h^*(-d\theta) = \omega$, where θ is the Liouville canonical 1-form on T^*L.*

Proof. This is a combination of Lemma 5.3 and Weinstein's Theorem 5.4. □

2.6 Hamiltonian vector fields and Poisson bracket

Let (M,ω) be a symplectic $2n$-manifold. A smooth vector field X on M is called *Hamiltonian* if there exists a smooth function $H : M \to \mathbb{R}$ such that $i_X\omega = dH$. In other words,

$$\omega_p(X_p, v_p) = v_p(H)$$

for every $v_p \in T_pM$ and $p \in M$. We usually write $X = X_H$ and obviously $X_H = \tilde{\omega}^{-1}(dH)$.

If $M = T^*\mathbb{R}^n \cong \mathbb{R}^{2n}$ with the canonical symplectic 2-form

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

we have $\tilde{\omega}(\frac{\partial}{\partial q^i}) = dp_i$ and $\tilde{\omega}(\frac{\partial}{\partial p^i}) = -dq^i$. Thus,

$$X_H = \tilde{\omega}^{-1}\Big(\sum_{i=1}^{n} \frac{\partial H}{\partial q^i} dq^i + \sum_{i=1}^{n} \frac{\partial H}{\partial p_i} dp_i\Big) = \sum_{i=1}^{n} \Big(\frac{\partial H}{\partial p_i} \cdot \frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i} \cdot \frac{\partial}{\partial p_i}\Big).$$

So the integral curves of X_H are the solutions of Hamilton's differential equations

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q^i}, \qquad 1 \le i \le n.$$

According to Darboux's theorem, this is true locally, with respect to suitable local coordinates, on every symplectic $2n$-manifold.

A smooth vector field X on M is called *symplectic* or *locally Hamiltonian* if $L_X\omega = d(i_X\omega) = 0$. In this case the lemma of Poincaré implies that every point $p \in M$ has an open neighbourhood V diffeomorphic to $\mathbb{R}^{2n}$ for which there exists a smooth function $H_V : V \to \mathbb{R}$ such that $i_X\omega|_V = dH_V$. From Lemma 4.1, we have

$$\frac{d}{dt}\phi_t^*\omega = \phi_t^*(d(i_X\omega)),$$

and $\phi_0^*\omega = \omega$, where ϕ_t is the flow of X. Thus X is locally Hamiltonian if and only if its flow consists of symplectomorphisms.

A locally Hamiltonian vector field may not be Hamiltonian. As a simple example, let $M = S^1 \times S^1$ equiped with the volume element ω such that $\pi^*\omega = dx \wedge dy$, where $\pi : \mathbb{R}^2 \to M$ is the universal covering projection. The smooth vector field

$$X = \pi_*(\frac{\partial}{\partial x})$$

is locally Hamiltonian, since locally $\pi^*(i_X\omega) = dy$. But if $j : S^1 \to M$ is the embedding $j(z) = (1, z)$, then $j^*(i_X\omega)$ is the natural generator of $H^1_{DR}(S^1) \cong \mathbb{R}$ and thus it is not exact. Therefore $i_X\omega$ is not exact.

Two elementary properties of Hamiltonian vector fields are the following.

Proposition 6.1. *The smooth function $H : M \to \mathbb{R}$ is a first integral of the Hamiltonian vector field X_H.*

Proof. Indeed $X_H(H) = dH(X_H) = \omega(X_H, X_H) = 0$. □

Proposition 6.2. *Let (M_1, ω_1) and (M_2, ω_2) be two symplectic manifolds. A diffeomorphism $f : M_1 \to M_2$ is symplectic if and only if $f_*(X_{H\circ f}) = X_H$ for every open set $U \subset M_2$ and smooth function $H : U \to \mathbb{R}$.*

Proof. The condition $X_H(f(p)) = f_{*p}(X_{H\circ f}(p))$ for every $p \in f^{-1}(U)$ is equivalent to

$$\omega_2(f(p))(X_H(f(p)), f_{*p}(v)) = \omega_2(f(p))(f_{*p}(X_{H\circ f}(p)), f_{*p}(v))$$

for every $v \in T_pM$, since ω_2 is non-degenerate and f is a diffeomorphism. Equivalently,

$$dH(f(p))(f_{*p}(v)) = (f^*\omega_2)(p)(X_{H\circ f}(p), v)$$

or

$$i_{X_{H\circ f}}\omega_1 = d(H \circ f) = f^*(dH) = i_{X_{H\circ f}}(f^*\omega_2)$$

on $f^{-1}(U)$. This is true, if f is symplectic. Conversely, if this holds, then for every $p \in M_1$ and u, $v \in T_pM_1$ there exists an open neighbourhood U of $f(p)$ in M_2 and a smooth function $H : U \to \mathbb{R}$ such that $u = X_{H\circ f}(p)$. So, $\omega_1(p)(u, v) = (f^*\omega_2)(p)(u, v)$ for every u, $v \in T_pM_1$. This means that $f^*\omega_2 = \omega_1$. $\square$

If (M, ω) is a symplectic manifold and F, $G \in C^\infty(M)$, then the smooth function

$$\{F, G\} = i_{X_G} i_{X_F}\omega \in C^\infty(M)$$

is called the *Poisson bracket* of F and G. From Proposition 6.2 we obtain the following.

Corollary 6.3. *Let (M_1, ω_1) and (M_2, ω_2) be two symplectic manifolds. A diffeomorphism $f : M_1 \to M_2$ is symplectic if and only if*

$$f^*\{F, G\} = \{f^*(F), f^*(G)\}$$

for every open set $U \subset M_2$ and F, $G \in C^\infty(U)$.

Proof. Suppose first that f is a symplectomorphism and F, $G \in C^\infty(U)$, where $U \subset M_2$ is an open set. Then

$$\{F, G\}(f(p)) = \omega_2(f(p))(X_F(f(p)), X_G(f(p)))$$

$$= \omega_2(f(p))(f_{*p}(X_{F\circ f}(p)), f_{*p}(X_{G\circ f}(p))) = (f^*\omega_2)(p)(X_{F\circ f}(p), X_{G\circ f}(p))$$

$$= \omega_1(p)(X_{F\circ f}(p), X_{G\circ f}(p)) = \{F \circ f, G \circ f\}(p).$$

Conversely, if $\{F, G\}(f(p)) = \{F \circ f, G \circ f\}(p)$ for every $p \in f^{-1}(U)$ and every open set $U \subset M_2$ and F, $G \in C^\infty(U)$, then

$$(f^*\omega_2)(p)(X_{F\circ f}(p), X_{G\circ f}(p)) = \omega_1(p)(X_{F\circ f}(p), X_{G\circ f}(p)).$$

But for every $p \in M_1$ and u, $v \in T_pM_1$ there exists an open set $V \subset M_1$ with $p \in V$ and F, $G \in C^\infty(f(V))$ such that $X_{F\circ f}(p) = u$ and $X_{G\circ f}(p) = v$. So $(f^*\omega_2)(p)(u, v) = \omega_1(p)(u, v)$. This means $f^*\omega_2 = \omega_1$. $\square$

Corollary 6.4. *Let X be a complete Hamiltonian vector field with flow $(\phi_t)_{t\in\mathbb{R}}$ on a symplectic manifold M. Then $\phi_t^*\{F, G\} = \{\phi_t^*(F), \phi_t^*(G)\}$ for every F,*

$G \in C^\infty(M)$. If X is not complete, the same is true on suitable open sets.

Corollary 6.5. *Let X be a complete Hamiltonian vector field with flow $(\phi_t)_{t\in\mathbb{R}}$ on a symplectic manifold (M,ω). Then*

$$X\{F,G\} = \{X(F),G\} + \{F,X(G)\}$$

for every F, $G \in C^\infty(M)$. If X is not complete, the same is true on suitable open sets.

Proof. From Corollary 6.4 we have

$$X\{F,G\} = \frac{d}{dt}\Big|_{t=0} \phi_t^*\{F,G\} = \frac{d}{dt}\Big|_{t=0} \{\phi_t^*(F),\phi_t^*(G)\} = \frac{d}{dt}\Big|_{t=0} \omega(X_{\phi_t^*(F)}, X_{\phi_t^*(G)})$$

$$= \omega(\frac{d}{dt}\Big|_{t=0} X_{\phi_t^*(F)}, X_G) + \omega(X_F, \frac{d}{dt}\Big|_{t=0} X_{\phi_t^*(G)}).$$

But

$$\tilde{\omega}(\frac{d}{dt}\Big|_{t=0} X_{\phi_t^*(F)}) = \frac{d}{dt}\Big|_{t=0} \tilde{\omega}(X_{\phi_t^*(F)}) = \frac{d}{dt}\Big|_{t=0} d\phi_t^*(F)$$

$$= d(\frac{d}{dt}\Big|_{t=0} \phi_t^*(F)) = dX(F) = \tilde{\omega}(X_{X(F)}),$$

which means that

$$\frac{d}{dt}\Big|_{t=0} X_{\phi_t^*(F)} = X_{X(F)}.$$

Consequently,

$$X\{F,G\} = \omega(X_{X(F)}, X_G) + \omega(X_F, X_{X(G)}) = \{X(F),G\} + \{F,X(G)\}. \qquad \square$$

It is obvious that for a symplectic manifold (M,ω) the Poisson bracket

$$\{,\} : C^\infty(M) \times C^\infty(M) \to C^\infty(M)$$

is bilinear and skew-symmetric. Form Corollary 6.5 follows now that it satisfies the Jacobi identity. Indeed, if F, G, $H \in C^\infty(M)$ then

$$\{F,G\} = (i_{X_F}\omega)(X_G) = dF(X_G) = X_G(F)$$

and thus $\{\{F,G\},H\} = X_H(\{F,G\})$. Consequently,

$$\{\{F,G\},H\} = \{X_H(F),G\} + \{F,X_H(G)\} = \{\{F,H\},G\} + \{F,\{G,H\}\}.$$

This is the Jacobi identity and so the $(C^\infty(M),\{,\})$ is a Lie algebra.

There is a Leibniz formula for the product of two smooth functions with respect to the Poisson bracket, because if F, G, $H \in C^\infty(M)$, then

$$\{F \cdot G, H\} = X_H(F \cdot G) = F \cdot X_H(G) + G \cdot X_H(F) = F \cdot \{G,H\} + G \cdot \{F,H\}.$$

Proposition 6.6. *Let X_H be a Hamiltonian vector field with flow ϕ_t on a symplectic manifold M. Then*

$$\frac{d}{dt}(F \circ \phi_t) = \{F \circ \phi_t, H\} = \{F, H\} \circ \phi_t$$

for every $F \in C^\infty(M)$.

Proof. By the chain rule, for every $p \in M$ we have

$$\frac{d}{dt}(F \circ \phi_t)(p) = (dF)(\phi_t(p))X_H(\phi_t(p)) = \{F, H\}(\phi_t(p))$$

$$= \{F \circ \phi_t, H \circ \phi_t\}(p) = \{F \circ \phi_t, H\}(p),$$

since H is a first integral of X_H. □

Corollary 6.7. *A smooth function $F : M \to \mathbb{R}$ on a symplectic manifold M is a first integral of a Hamiltonian vector field X_H on M if and only if $\{F, H\} = 0$.* □

Let $\mathfrak{sp}(M, \omega)$ and $\mathfrak{h}(M, \omega)$ denote the linear spaces of the symplectic and hamiltonian vector fields, respectively, of the symplectic manifold (M, ω). We shall conclude this section with a few remarks about these spaces.

Proposition 6.8. *If X, $Y \in \mathfrak{sp}(M, \omega)$, then $[X, Y] = -X_{\omega(X,Y)}$. In particular, $[X_F, X_G] = -X_{\{F,G\}}$ for every F, $G \in C^\infty(M)$.*

Proof. Indeed, $i_{[X,Y]} = [L_X, i_Y]$ and therefore

$$i_{[X,Y]}\omega = L_X(i_Y\omega) - i_Y(L_X\omega) = d(i_X i_Y \omega) + i_X(d(i_Y\omega)) - 0 =$$

$$d(\omega(Y, X)) + 0 = -i_{X_{\omega(X,Y)}}\omega.$$

Since ω is non-degenerate the result follows. □

Proposition 6.8 implies that $\mathfrak{sp}(M, \omega)$ and $\mathfrak{h}(M, \omega)$ are Lie subalgebras of the Lie algebra of smooth vector fields of M. Moreover, $\mathfrak{h}(M, \omega)$ is an ideal in $\mathfrak{sp}(M, \omega)$ since $[\mathfrak{sp}(M, \omega), \mathfrak{sp}(M, \omega)] \subset \mathfrak{h}(M, \omega)$. If M is connected, we have two obvious short exact sequences of Lie algebra homomorphisms

$$0 \to \mathbb{R} \to C^\infty(M) \xrightarrow{r} \mathfrak{h}(M, \omega) \to 0,$$

where $r(F) = -X_F$ for every $F \in C^\infty(M)$, and

$$0 \to \mathfrak{h}(M, \omega) \to \mathfrak{sp}(M, \omega) \to H^1_{DR}(M) \to 0,$$

which do not split in general. The first makes $(C^\infty(M), \{,\})$ a central extension of $\mathfrak{h}(M, \omega)$. In the second, we consider in $H^1_{DR}(M)$ the trivial Lie bracket.

Proposition 6.9. *Let (M, ω) be a compact, connected, symplectic $2n$-manifold and $\omega^n = \omega \wedge \omega \wedge ... \wedge \omega$ (n times).*

(a) If X, $Y \in \mathcal{X}(M)$, then $\omega(X,Y)\omega^n = -n \cdot i_X\omega \wedge i_Y\omega \wedge \omega^{n-1}$.
(b) If F, $G \in C^\infty(M)$, then

$$\int_M \{F,G\}\omega^n = 0.$$

(c) The set $C_0^\infty(M,\omega) = \{F \in C^\infty(M) : \int_M F\omega^n = 0\}$ is a Lie subalgebra of $(C^\infty(M), \{,\})$ and $C^\infty(M) = \mathbb{R} \oplus C_0^\infty(M,\omega)$.
(d) The short exact sequence of Lie algebra homomorphisms

$$0 \to \mathbb{R} \to C^\infty(M) \xrightarrow{r} \mathfrak{h}(M,\omega) \to 0$$

splits.

Proof. (a) Since

$$0 = i_X(i_Y\omega \wedge \omega^n) = \omega(X,Y)\omega^n - i_Y\omega \wedge i_X\omega^n$$

and $i_X\omega^n = n \cdot i_X\omega \wedge \omega^{n-1}$, we conclude that

$$\omega(X,Y)\omega^n = n \cdot i_Y\omega \wedge i_X\omega \wedge \omega^{n-1}.$$

(b) Using (a) and the fact that ω^{n-1} is closed, we have

$$\{F,G\}\omega^n = \omega(X_F,X_G)\omega^n = -n \cdot i_{X_F}\omega \wedge i_{X_G}\omega \wedge \omega^{n-1} =$$

$$-n \cdot dF \wedge dG \wedge \omega^{n-1} = -n \cdot d(FdG \wedge \omega^{n-1}).$$

The conclusion follows now from Stokes formula.

(c) From (b) follows immediately that $C_0^\infty(M,\omega)$ is a Lie subalgebra of $(C^\infty(M), \{,\})$. Moreover, every $F \in C^\infty(M)$ can be written as

$$F = \frac{1}{\mathrm{vol}(M)}\int_M F\omega^n + \left(F - \frac{1}{\mathrm{vol}(M)}\int_M F\omega^n\right).$$

(d) If we define $j : \mathfrak{h}(M,\omega) \to C^\infty(M)$ by

$$j(X_F) = -F + \frac{1}{\mathrm{vol}(M)}\int_M F\omega^n,$$

then $\{j(X_F), j(X_G)\} = \{F,G\} = j(-X_{\{F,G\}}) = j([X_F,X_G])$, by Proposition 6.8, and therefore j is a Lie algebra homomorphism. Obviously, $r \circ j = id$. □

2.7 The characteristic line bundle of a hypersurface

Let (M,ω) be a symplectic $2n$-manifold. If J is a compatible almost complex structure and g the corresponding Riemannian metric on M so that ω is given by the formula $\omega(u,v) = g(J(u),v)$ for all u, $v \in TM$, then for every smooth function $H : M \to \mathbb{R}$ the corresponding Hamiltonian vector field is $X_H = -J(\mathrm{grad}H)$, where $\mathrm{grad}H$ denotes the gradient vector field of H with respect to the Riemannian metric

g. As we know, the orbits of X_H lie on the level sets of H. If $c \in \mathbb{R}$ is a regular value of H and $S = H^{-1}(c) \neq \varnothing$, then S is a submanifold of M of codimension 1. An important observation is that the unparametrised orbits of X_H on S depend only on S and not on X_H, i.e. not on the Hamiltonian H. Note that $X_H|_S$ is a nowhere vanishing tangent vector field on S, because c is a regular value of H.

An embedded smooth submanifold of codimension 1 in M will be called a hypersurface in the sequel. Let S be a hypersurface in M such that $S = H_j^{-1}(c_j)$, where $c_j \in \mathbb{R}$ is a regular value of the smooth function $H_j : M \to \mathbb{R}$ for $j = 1, 2$. Since $\text{grad} H_j$is g-orthogonal to S for both $j = 1, 2$, there exists a smooth function $\lambda : S \to \mathbb{R} \setminus \{0\}$ such that $\text{grad} H_2(x) = \lambda(x) \text{grad} H_1(x)$ and therefore $X_{H_2}(x) = \lambda(x) X_{H_1}(x)$ for every $x \in S$. Let $\gamma : I \to S$ be an integral curve of X_{H_1} and let $h : I \to \mathbb{R}$ be the smooth function defined by

$$h(s) = \int_0^s \frac{1}{\lambda(\gamma(t))} dt.$$

Then $\sigma = \gamma \circ h^{-1}$ is an integral curve of X_{H_2}, because

$$\dot{\sigma} = \frac{1}{h' \circ h^{-1}} (\dot{\gamma} \circ h^{-1}) = (\lambda \circ \gamma \circ h^{-1}) \cdot X_{H_1} \circ (\gamma \circ h^{-1}) = X_{H_2} \circ (\gamma \circ h^{-1}) = X_{H_1} \circ \sigma.$$

This shows that the two Hamiltonian vector fields have the same unparametrised orbits on S.

The above imply that given a hypersurface S in M, there exists a (real) line bundle $\mathcal{L}_S \subset TS$ which gives the direction of every Hamiltonian vector field having S as a regular hypersurface of constant energy. Such a line bundle can be described without reference to any Hamiltonian function as follows. Since S has dimension $2n - 1$, the restriction of ω_x on T_xS is degenerate for every $x \in S$. The linear subspace

$$L_x = \{u \in T_xS : \omega_x(u, v) = 0 \text{ for every } v \in T_xS\}$$

of T_xS has dimension 1, because if $v \in T_xM \setminus T_xS$ is any non-zero vector, then $\omega_x(., v) : L_x \to \mathbb{R}$ is a linear isomorphism. The line bundle $\mathcal{L}_S$ with fiber L_x over $x \in S$ is called the *characteristic line bundle* of the hypersurface S. A characteristic of S is a leaf of the 1-dimensional foliation to which $\mathcal{L}_S$ is tangent. If $H : M \to \mathbb{R}$ is a smooth function and $c \in \mathbb{R}$ a regular value of H such that $S = H^{-1}(c)$, then obviously $X_H(x) \in L_x$ for every $x \in S$.

Lemma 7.1. *If S is a compact hypersurface in M such that $\mathcal{L}_S$ is orientable (or equivalently trivial), then there exists an open neighbourhood U of S and a smooth function $H : U \to \mathbb{R}$ such that $S = H^{-1}(c)$ for some regular value $c \in \mathbb{R}$ of H.*

Proof. Let $\mathcal{N}_S$ be the normal bundle of S in M with fiber

$$N_x = \{u \in T_xM : g(u, v) = 0 \text{ for every } v \in T_xS\}$$

over $x \in S$. If $u \in L_x$, then $\omega(u, v) = 0$ for every $v \in T_xS$ or equivalently $J(u) \in N_x$. Obviously, the almost complex structure J induces a vector bundle isomorphism $\mathcal{L}_S \cong \mathcal{N}_S$. Since $\mathcal{L}_S$ is assumed to be orientable, hence trivial, the same is true for

$\mathcal{N}_S$. Let s be a nowhere vanishing smooth section of $\mathcal{N}_S$ and let $\psi : (-\epsilon, \epsilon) \times S \to M$ be the smooth map defined by

$$\psi(t, x) = \exp_x(ts(x)),$$

where exp is the exponential map of the compatible Riemannian metric g and $\epsilon > 0$ is such that ψ is defined. Since S is assumed to be compact, there exists some $\epsilon > 0$ such that ψ maps $(-\epsilon, \epsilon) \times S$ diffeomorphically onto an open neighbourhood U of S. If now $H : U \to \mathbb{R}$ is the smooth function defined by $H(\psi(t, x)) = t$, then $S = H^{-1}(0)$ and 0 is a regular value of H. □

The proof of the preceding lemma motivates the following.

Definition 7.2. A parametrised family of hypersurfaces in M modelled on a compact hypersurface $S \subset M$ is a smooth diffeomorphism $\psi : I \times S \to U$, where $I \subset \mathbb{R}$ is an open interval containing zero and U is a relatively compact neighbourhood of S in M such that $\psi(0, x) = x$ for $x \in S$.

Example 7.3. Let $H : \mathbb{R}^{2n} \to \mathbb{R}$ be a smooth function such that 0 is a regular value of H and $S = H^{-1}(0) \neq \varnothing$ is compact. Let ϕ be the local flow of the euclidean gradient vector field of H. There exists some $\epsilon > 0$ such that $\phi|_{(-\epsilon,\epsilon)\times S}$ is a parametrised family of hypersurfaces in $\mathbb{R}^{2n}$ modelled on S. Instead of $\mathbb{R}^{2n}$ we could have taken any symplectic manifold and then ϕ would be the local flow of the gradient vector field of H with respect to a compatible Riemannian metric.

Summurizing, the above mean that for a compact hypersurface $S \subset M$ the characteristic line bundle $\mathcal{L}_S$ is orientable if and only if its normal bundle $\mathcal{N}_S$ in M is orientable if and only if S is orientable if and only if there exists a parametrised family of hypersurfaces in M modelled on S if and only if there exists a relatively compact open neighbourhood of S in M and a smooth function $H : U \to \mathbb{R}$ such that $S = H^{-1}(c)$ for some regular value $c \in \mathbb{R}$ of H.

Chapter 3

Examples of symplectic manifolds

3.1 The geometry of the tangent bundle

In this section we shall study the geometry of the tangent bundle of a Riemannian manifold. Its structure is useful in Riemannian Geometry and when studying mechanical problems within the framework of newtonian mechanics.

Let M be a smooth n-dimensional manifold and let $p : TM \to M$ be its tangent bundle. There exists a canonical subbundle V of $T(TM)$, which is just $V = \mathrm{Ker} p_*$, and is called the vertical subbundle. In other words, for each $u \in TM$ the fiber V_u is the tangent space to the fiber T_xM of TM at u, where $x = p(u)$.

It would be desirable to have a canonical complementary to V subbundle of $T(TM)$. Unfortunately, this is impossible. In order to construct a complementary subbundle, we must use a Riemmanian metric on M. So from now on we assume that M is a Riemannian manifold with metric g. Tne corresponding Levi-Civita connection ∇ induces the connection map $K : T(TM) \to TM$ which is defined as follows. Let $u \in T_xM$, $x \in M$. Let W be a normal neighbourhood of x in M, that is $W = \exp_x(U)$, where U is a star-shaped open neighbourhood of $0 \in T_xM$ and $\exp_x|_U : U \to W$ is a diffeomorphism. Let $\tau : p^{-1}(W) \to T_xM$ be the smooth map which sends each $v \in T_yM$, $y \in W$ to its parallel translation at x along the unique geodesic in W form y to x. For $w \in T_xM$, we let $R_{-w} : T_xM \to T_xM$ be the translation by the vector $-w$. The connection map $K_u : T_u(TM) \to T_xM$ is defined by

$$K_u(\xi) = (\exp_x \circ R_{-u} \circ \tau)_{*u}(\xi).$$

Obviously, this is a well-defined linear map. An alternative definition in terms of covariant differentiation is given by the following.

Proposition 1.1. *Let $z : (-\epsilon, \epsilon) \to TM$, $\epsilon > 0$, be a smooth curve such that $z(0) = u$ and $\dot{z}(0) = \xi$. If $\gamma = p \circ z : (-\epsilon, \epsilon) \to M$ and X is the smooth vector field along γ such that $z(t) = (\gamma(t), X(t)) \in T_{\gamma(t)}M$, $|t| < \epsilon$, then*

$$K_u(\xi) = \nabla_{\dot{\gamma}(0)} X.$$

Proof. The definition of the connection map and the chain rule imply that

$$K_u(\xi) = (\exp_x \circ R_{-u} \circ \tau)_{*z(0)}(\dot{z}(0)) = \frac{d}{dt}\Big|_{t=0} (\exp_x \circ R_{-u} \circ \tau \circ z).$$

Since $(\exp_x \circ R_{-u} \circ \tau \circ z)(t) = \exp_x(\tau(z(t)) - u)$, we get

$$K_u(\xi) = (\exp_x)_{*0}\big(\frac{d}{dt}\Big|_{t=0} \tau(X(t))\big) = \nabla_{\dot{\gamma}(0)} X$$

since $(\exp_x)_{*0}$ is the identity. □

The horizontal subbundle H of $T(TM)$ is now the one whose fiber at $u \in TM$ is $H_u = \mathrm{Ker} K_u$.

It is evident that horizontal curves in TM, that is smooth curves tangent to H, correspond to parallel vector fields along curves in M. To be more precise, given $u \in T_xM$ and $\gamma : (-\epsilon, \epsilon) \to M$ a smooth curve with $\gamma(0) = x$ and $\dot{\gamma}(0) = u$, let $X(t)$, $|t| < \epsilon$, be the parallel transport of u along γ. Let also $\sigma : (-\epsilon, \epsilon) \to TM$ be the smooth curve $\sigma(t) = (\gamma(t), X(t))$. Then $\sigma(0) = (x, u)$ and if $\xi = \dot{\sigma}(0)$, we have

$$u = \dot{\gamma}(0) = p_{*u}(\dot{\sigma}(0)) = p_{*u}(\xi).$$

This shows that $p_{*u}(H_u) = T_xM$, since $K_u(\xi) = \nabla_{\dot{\gamma}(0)} X = 0$.

Moreover, $p_{*u}|_{H_u} : H_u \to T_xM$ is an isomorphism. Indeed, let $\xi \in T_u(TM)$ be such that $p_{*u}(\xi) = 0$. There exists a vertical smooth curve $z : (-\epsilon, \epsilon) \to T_xM \subset TM$ such that $z(0) = u$ and $\dot{z}(0) = \xi$. Thus, $\gamma = p \circ z$ takes the constant value $\gamma(t) = x$ for all $|t| < \epsilon$ and therefore $(i_x \circ \tau)_{*u}(\xi) = \xi$, where $i_x : T_xM \hookrightarrow TM$ denotes the inclusion. Since $(\exp_x \circ R_{-u})_{*u}$ is an isomorphism, we conclude that if $\xi \in H_u$, then $\xi = 0$. The above argument also shows that $H_u \cap V_u = \{0\}$ and $K_u|_{V_u} : V_u \to T_xM$ is also an isomorphism. Hence $T_u(TM) = H_u \oplus V_u$ and the linear map $j_u : T_u(TM) \to T_xM \oplus T_xM$ given by

$$j_u(\xi) = (p_{*u}(\xi), K_u(\xi))$$

is a isomorphism.

If now $X \in T_xM$, the horizontal lift of X to $u \in TM$ is the unique vector $X^h \in H_u$ such that $p_{*u}(X^h) = X$. The vertical lift of X is the unique vector $X^v \in V_u$ such that $X^v(\tilde{f}) = X(f)$ for all smooth functions f, where $\tilde{f}$ is the smooth function on TM with $\tilde{f}(u) = u(f)$.

Using the above decomposition of $T(TM)$ as the Whitney direct sum of two subbundles, we can define a Riemannian metric $\langle , \rangle$ on TM such that H and V become orthogonal subbundles and the tangent bundle projection $p : TM \to M$ becomes a Riemannian submersion. This Riemannian metric is called the Sasaki metric and is defined by

$$\langle \xi, \zeta \rangle_u = g(p_{*u}(\xi), p_{*u}(\zeta)) + g(K_u(\xi), K_u(\zeta)).$$

It is worth to note that the geodesic vector field $G : TM \to T(TM)$ has a very simple expression under the isomorphism j_u, $u \in TM$. If γ_u denotes the geodesic with $\gamma_u(0) = x$ and $\dot{\gamma}_u(0) = u$, then

$$G(u) = \frac{d}{dt}\Big|_{t=0} \dot{\gamma}_u(t).$$

Since $\dot{\gamma}_u(t)$ is the parallel transport of u along γ_u, is follows that $p_{*u}(G(u)) = u$ and $K_u(G(u)) = 0$. Therefore, $j_u(G(u)) = (u, 0)$.

The following proposition will be used later.

Proposition 1.2. *Let $\overline{\nabla}$ denote the Levi-Civita connection of the Sasaki metric on TM. Then, $\overline{\nabla}_\xi G \in V$ for every $\xi \in H$.*

Proof. Let $u \in T_xM$, $x \in M$ and let $\{E_1(x), E_2(x), \cdots, E_n(x)\}$ be an orthonormal basis of T_xM. Let W be a normal neighbourhood of x. For every $y \in W$ there exists a unique geodecic arc in W joining y and x. Transporting parallely along this geodesic we obtain an orthonormal frame $\{E_1, E_2, \cdots, E_n\}$ on W such that $(\nabla_{E_i} E_j)(x) = 0$ for all $1 \leq i, j \leq n$. The set $\{E_1^h, E_2^h, \cdots, E_n^h\}$ of the horizontal lifts is orthonormal with respect to the Sasaki metric and spans $H|_{p^{-1}(W)}$. It suffices to prove that $\overline{\nabla}_{E_j^h} G \in V$ for every $1 \leq j \leq n$.

Since the geodesic vector field is horizontal and $j_u(G(u)) = (u, 0)$, we have

$$G(u) = \sum_{i=1}^{n} g(E_i(p(u)), u) E_i^h$$

and therefore

$$\overline{\nabla}_{E_j^h} G = \sum_{i=1}^{n} E_j^h(g(E_i(p(u)), u)) E_i^h + \sum_{i=1}^{n} g(E_i(p(u)), u) \overline{\nabla}_{E_j^h} E_i^h.$$

Since the tangent bundle projection $p : TM \to M$ is a Riemannian submersion we have

$$\overline{\nabla}_{E_j^h} E_i^h = \left(\nabla_{E_j} E_i\right)^h + \frac{1}{2}[E_j^h, E_i^h]^v$$

and so $(\overline{\nabla}_{E_j^h} E_i^h)(x)$ is vertical (recall that $[E_j^h, E_i^h]^v(x)$ depends only on the values $E_j^h(x)$ and $E_i^h(x)$). This implies that the second term in the above expression of $(\overline{\nabla}_{E_j^h} G)(x)$ is vertical. As far as the first term is concerned, for each $1 \leq j \leq n$ let $\gamma_j : (-\epsilon, \epsilon) \to M$, $\epsilon > 0$ be a smooth curve such that $\gamma_j(0) = x$ and $\dot{\gamma}_j(0) = E_j(x)$. Let $X_j(t)$, $|t| < \epsilon$ be the parallel transport of u along γ_j. Then

$$E_j^h(x)(g(E_i(p(u)), u)) = \frac{d}{dt}\bigg|_{t=0} g(E_i(\gamma_j(t)), X_j(t)) = g(\nabla_{E_j(x)}, u) = 0. \quad \square$$

The splitting j_u permits to define an almost complex structure J on TM by setting

$$J_u(\xi^h, \xi^v) = (-\xi^v, \xi^h),$$

where $j_u(\xi) = (\xi^h, \xi^v)$. Obviously, J interchanges H and V. Also, J_u preserves the Sasaki metric, because

$$\begin{aligned}\langle J_u(\xi), J_u(\zeta)\rangle &= g(p_{*u}(J_u(\xi)), p_{*u}(J_u(\zeta))) + g(K_u(J_u(\xi)), K_u(J_u(\zeta))) \\ &= g(-\xi^v, -\zeta^v) + g(\xi^h, \zeta^h) = \langle \xi, \zeta\rangle.\end{aligned}$$

Lemma 1.3. *If $\{E_1, E_2, \cdots, E_n\}$ is a local frame on M, then $[E_i^h, J(E_j^h)] = 0$ for all $1 \le i, j \le n$.*

Proof. Let $u \in T_xM$, $x \in M$ and $\gamma_j : (-\epsilon, \epsilon) \to M$, $\epsilon > 0$, be an integral curve of E_j through x. Let $X_j(t)$, $|t| < \epsilon$ be the parallel transport of u along γ_j. Let $Z_{ij} : (-\epsilon, \epsilon) \times (-\epsilon, \epsilon) \to TM$ be the smooth map defined by

$$Z_{ij}(t, s) = (\gamma_i(t), sE_j(\gamma_i(t)) + X_i(t)).$$

Then,

$$\frac{\partial Z_{ij}}{\partial s}(t, s) = J(E_j^h(Z_{ij}(t, s))),$$
$$\frac{\partial Z_{ij}}{\partial t}(0, s) = E_i^h(Z_{ij}(0, s))$$

and

$$\frac{\partial Z_{ij}}{\partial t}(t, 0) = E_i^h(Z_{ij}(t, 0)).$$

For every smooth function f defined locally on TM we have now

$$[E_i^h, J(E_j^h)](u)(f) = \frac{\partial}{\partial t}\left(\frac{\partial(f \circ Z_{ij})}{\partial s}\right)(0, 0) - \frac{\partial}{\partial s}\left(\frac{\partial(f \circ Z_{ij})}{\partial t}\right)(0, 0) = 0. \quad \square$$

Using the Sasaki metric and the almost complex structure J we can define a symplectic 2-form Ω on TM by the formula

$$\Omega_u(\xi, \zeta) = \langle J_u(\xi), \zeta\rangle = g(p_{*u}(\xi), K_u(\zeta)) - g(K_u(\xi), p_{*u}(\zeta)).$$

We shall use Proposition 1.2 and Lemma 1.3 to prove that Ω is precisely the symplectic 2-form which corresponds to the standard symplectic 2-form on T^*M under the Legendre transformation defined by the Riemannian metric g. Recall that this is the natural vector bundle isomorphism $\mathcal{L} : TM \to T^*M$ induced by g. Now we can compute for $u \in T_xM$, $x \in M$ and $\mathcal{L}(x, u) = (x, a)$ that

$$(\mathcal{L}^*\theta)_u(\xi) = \theta(\mathcal{L}_{*u}(\xi)) = a((\pi \circ \mathcal{L})_{*u}(\xi)) = a(p_{*u}(\xi)) = g(p_{*u}(\xi), u) = \langle \xi, G(u)\rangle,$$

where $\pi : T^*M \to M$ is the cotangent bundle projection and θ is the Liouville canonical 1-form on T^*M.

Proposition 1.4. $\Omega = \mathcal{L}^*(-d\theta)$.

Proof. We put $A = \mathcal{L}^*\theta$ for simplicity and observe that

$$dA(\xi, \zeta) = \xi(A(\zeta)) - \zeta(A(\xi)) - A([\xi, \zeta]) = \xi\langle \zeta, G\rangle - \zeta\langle \xi, G\rangle - \langle [\xi, \zeta], G\rangle$$

$$= \langle \overline{\nabla}_\xi \zeta, G\rangle + \langle \zeta, \overline{\nabla}_\xi G\rangle - \langle \overline{\nabla}_\zeta \xi, G\rangle - \langle \xi, \overline{\nabla}_\zeta G\rangle - \langle [\xi, \zeta], G\rangle = \langle \zeta, \overline{\nabla}_\xi G\rangle - \langle \xi, \overline{\nabla}_\zeta G\rangle,$$

since the Levi-Civita connection is symmetric.

Let now $\{E_1, E_2, \cdots, E_n\}$ be a local orthonormal frame on M. The set

$$\{E_1^h, E_2^h, \cdots, E_n^h, J(E_1^h), J(E_2^h), \cdots, J(E_n^h)\}$$

is a local orthonormal frame on TM. Since each $J(E_i^h)$ is vertical, so is also $[J(E_i^h), J(E_j^h)]$ for all $1 \leq i, j \leq n$. The fact that the geodesic vector field G is horizontal implies that $dA|_{V \times V} = 0$ from the above expression of dA, and Proposition 1.2 implies that $dA|_{H \times H} = 0$. Therefore it suffices to prove that $dA(E_i^h, J(E_j^h)) = -\Omega(E_i^h, J(E_j^h))$, $1 \leq i, j \leq n$. From Lemma 1.3 we get

$$dA(E_i^h, J(E_j^h)) = -J(E_j^h)(A(E_i^h)) = -J(E_j^h)(\langle E_i^h, G\rangle),$$

because $J(E_j^h)$ is vertical.

Let $u \in T_xM$, $x \in M$ and the frame $\{E_1, E_2, \cdots, E_n\}$ be defined on a normal neighbourhood of x as in the proof of Proposition 1.2. The curve $\gamma_j : (-\epsilon, \epsilon) \to TM$ defined by $\gamma_j(t) = (x, tE_j(x) + u)$, $|t| < \epsilon$, is an integral curve of $J(E_j^h)$ and so

$$(dA)_u(E_i^h, J(E_j^h)) = -\frac{d}{dt}\bigg|_{t=0} g(E_i(x), tE_j(x) + u) = -\delta_{ij}.$$

On the other hand, from the definition of Ω we have

$$\Omega_u(E_i^h, J(E_j^h)) = \langle J(E_i^h), J(E_j^h)\rangle = \langle E_i^h, E_j^h\rangle = \delta_{ij}. \quad \square$$

The tangent space $T_u(TM)$, for $u \in T_xM$, $x \in M$, can be described in terms of Jacobi fields along the geodesic γ_u with $\gamma_u(0) = x$ and $\dot{\gamma}_u(0) = u$. Let $\xi \in T_u(TM)$ and $z : (-\epsilon, \epsilon) \to TM$, $\epsilon > 0$, be a smooth curve such that $z(0) = u$ and $\dot{z}(0) = \xi$. If $(\phi_t)_{t \in \mathbb{R}}$ is the geodesic flow, i.e. the flow of G (assuming that the metric g is complete), then $F : (-\epsilon, \epsilon) \times \mathbb{R} \to M$ defined by $F(s,t) = p(\phi_t(z(s)))$ is a variation of the geodesic $\gamma_u(t) = p(\phi_t(u))$ through geodesics and therefore the variational field

$$I_\xi(t) = \frac{\partial F}{\partial s}(0, t)$$

is a Jacobi field along γ_u with initial conditions $I_\xi(0) = p_{*u}(\xi)$ and

$$\frac{DI_\xi}{dt}(0) = \frac{D}{dt}\left(\frac{\partial F}{\partial s}\right)(0,0) = \frac{D}{ds}\left(\frac{\partial F}{\partial t}\right)(0,0) = \frac{DX}{dt}(0,0) = K_u(\xi),$$

from Proposition 1.1, where $\phi_t(z(s)) = (p(\phi_t(z(s))), X(s,t))$.

If we denote the vector space of Jacobi fields along γ_u by $I(u)$, then the map $i_u : T_u(TM) \to I(u)$ defined by $i_u(\xi) = I_\xi$ is a linear isomorphism, because $\mathrm{Ker} i_u = H_u \cap V_u = \{0\}$ and both vector spaces have dimension $2n$.

The geodesic vector field G on TM is the Hamiltonian vector field of the kinetic energy $H(x, v) = \frac{1}{2} g_x(v, v)$ with respect to Ω. Indeed, let $z : (-\epsilon, \epsilon) \to TM$, $\epsilon > 0$, be a smooth curve such that $z(0) = u$ and $\dot{z}(0) = \xi$. Let $\gamma = p \circ z$ and X be the smooth vector field along γ with $z(t) = (\gamma(t), X(t))$. Then,

$$\Omega_u(G(u), \xi) = g(p_{*u}(G(u)), K_u(\xi)) = g(u, K_u(\xi))$$

and on the other hand

$$(dH)_u(\xi) = \frac{d}{dt}\bigg|_{t=0} \frac{1}{2} g(X(t), X(t)) = g(\nabla_{\dot{\gamma}(0)} X, X(0)) = g(K_u(\xi), u).$$

This shows that $dH = i_G\Omega$. Therefore, the geodesic flow leaves Ω invariant.

Proposition 1.5. *Let $u \in TM$, $\xi \in T_u(TM)$ and $t \in \mathbb{R}$. Then,*

$$(\phi_t)_{*u}(\xi) = (I_\xi(t), \frac{DI_\xi}{dt}(t)) \in T_{\phi_t(u)}(TM) \cong H_{\phi_t(u)} \oplus V_{\phi_t(u)}.$$

Proof. Using the above notations we have

$$I_\xi(t) = \frac{\partial}{\partial s}\Big|_{s=0} p(\phi_t(z(s))) = (p \circ \phi_t)_{*u}(\xi) = p_{*\phi_t(u)}((\phi_t)_{*u}(\xi))$$

and

$$\frac{DI_\xi}{dt}(t) = \frac{D}{dt}\left(\frac{\partial F}{\partial s}\right)(0,t) = \frac{D}{ds}\Big|_{s=0}\left(\frac{d}{dt}p(\phi_t(z(s)))\right)$$
$$= \frac{D}{ds}\Big|_{s=0}\phi_t(z(s)) = K_{\phi_t(u)}((\phi_t)_{*u}(\xi)).$$

It follows that

$$j_u((\phi_t)_{*u}(\xi)) = (I_\xi(t), \frac{DI_\xi}{dt}(t)). \quad \square$$

Corollary 1.6. $\Omega_u(\xi,\zeta) = g(-\frac{DI_\xi}{dt}(t), I_\zeta(t)) + g(I_\xi(t), \frac{DI_\zeta}{dt}(t))$ *for all $t \in \mathbb{R}$ (assuming that the metric g is complete).* $\square$

3.2 The manifold of geodesics

Let (M,g) be a complete Riemannian n-manifold. A unit speed geodesic $\gamma : \mathbb{R} \to M$ is called periodic of period $\ell > 0$ if $\gamma(t+\ell) = \gamma(t)$ for every $t \in \mathbb{R}$ and ℓ is the smallest positive real number with this property. In this case the length of γ is ℓ. If every geodesic of M is periodic of the same period ℓ, then (M,g) is called a C_ℓ-manifold and its metric a C_ℓ-metric. The geodesic flow of a C_ℓ-manifold is periodic and there exists a smooth free action of S^1 on the unit tangent bundle T^1M of M whose orbit space is smooth $(2n-2)$-manifold CM. Also the quotient map $q : T^1M \to CM$ is a principle S^1-bundle. The manifold CM is called the manifold of oriented geodesics of M.

Example 2.1. The sphere S^n, $n \geq 2$, equiped with the usual euclidean Riemannian metric is a $C_{2\pi}$-manifold. From the uniuqeness of geodesics follows that the orienred geodesics on S^n are in one-to-one correspondence with the oriented 2-dimensional linear subspaces of $\mathbb{R}^{n+1}$. Therefore, CS^n is diffeomorphic to $SO(n+1,\mathbb{R})/SO(2,\mathbb{R}) \times SO(n-1,\mathbb{R})$. The same space is the manifold of geodesics of the real projective space $\mathbb{R}P^n$ with its standard Riemannian metric which is a C_π-manifold and is doubly covered by S^n.

The manifold of geodesics of any C_ℓ-metric on S^2 can be determined from the honotopy exact sequence

$$\cdots \to \pi_1(S^1) \to \pi_1(T^1S^2) \to \pi_1(CS^2) \to \{1\}$$

of the fibration $q : T^1S^2 \to CS^2$. Recall that T^1S^2 is diffeomorphic to $\mathbb{R}P^3$ and so $\pi_1(T^1S^2) \cong \mathbb{Z}_2$. It follows that $\pi_1(CS^2)$ is either trivial or isomorphic to $\mathbb{Z}_2$. However, we shall show shortly that the manifold of geodesics carries a symplectic structure and is therefore orientable. Hence CS^2 is diffeomorphic to S^2.

The manifold of geodesics CM of a C_ℓ-manifold M can be given a natural symplectic structure. Recall from Proposition 1.4 that TM has a symplectic structure $\Omega = -dA$, where A is the pullback of the Liouville canonical 1-form on T^*M under the natural bundle isomorpsism $\mathcal{L} : TM \to T^*M$ defined by the Riemannian metric g. Let $\eta = A|_{T^1M}$. Recall also that $T^1M = H^{-1}(1/2)$, where $H : TM \to \mathbb{R}$ is the kinetic energy. Let $(\psi_t)_{t\in\mathbb{R}}$ be the smooth flow on TM defined by $\psi_t(u) = e^t u$. Its infinitesimal generator Y is the smooth vector field on TM which in local coordinates $(q^1, \ldots, q^n, v^1, \ldots, v^n)$ on TM is represented as

$$Y|_{\text{locally}} = \sum_{k=1}^{n} v^k \frac{\partial}{\partial v^k}.$$

Since $dH(Y) = 2H$, it follows that Y is transverse to T^1M.

In local coordinates we have

$$A|_{\text{locally}} = \sum_{i,j=1}^{n} g_{ij} v^j dq^i,$$

where $g = (g_{ij})$ is the local form of the Riemannian metric. Therefore,

$$dA = \sum_{i,j=1}^{n} g_{ij} dv^j \wedge dq^i + \sum_{i,j,k=1}^{n} \frac{\partial g_{ij}}{\partial q^k} v^j dq^k \wedge dq^i.$$

A simple calculation shows that $i_Y(dA) = A$ and so $d(i_Y\Omega) = \Omega$ or equivalently $L_Y\Omega = \Omega$. We compute now that

$$A \wedge (dA)^{n-1} = (-1)^n i_Y\Omega \wedge \Omega^{n-1} = \frac{(-1)^n}{n} i_Y \Omega^n.$$

Since Ω^n is a volume form on TM and Y is transverse to T^1M, we conclude that $\eta \wedge (d\eta)^{n-1}$ is a volume form on T^1M. Also, $i_G d\eta = 0$, because the geodesic vector field G is the Hamiltonian vector field of the kinetic energy, i.e. $i_G\Omega = dH$. Finally, $i_G\eta = 1$, because $A(G) = \langle G, G\rangle = 2H$, where $\langle , \rangle$ is the Sasaki metric on TM.

Since the geodesic flow leaves the symplectic 2-form Ω on TM invariant, if M is a C_ℓ-manifold, there exists a unique closed 2-form ω on CM such that $q^*\omega = -\Omega|_{T^1M}$, where $q : T^1M \to CM$ is the quotient map as above. Actually, the restriction of the geodesic flow on T^1M leaves η invariant, because $L_G\eta = di_G\eta + i_G d\eta = 0$. So, the subbundle $\mathrm{Ker}\eta$ of $T(T^1M)$ is also left invariant under the geodesic flow. Since $i_G\eta = 1$, we have a splitting $T(T^1M) = \langle G\rangle \oplus \mathrm{Ker}\eta$ and the derivative of the quotient map $q_{*u} : \mathrm{Ker}\eta_u \to T_{q(u)}CM$ is a linear isomorphism for every $u \in T^1M$.

Proposition 2.2. *The induced closed 2-form ω on CM is symplectic.*

Proof. Let $X \in T_{q(u)}CM$ be such that $\omega(X,.) = 0$. If $\tilde{X} = (q_{*u})^{-1}(X)$, then $-\Omega_u(\tilde{X},.) = (d\eta)_u(\tilde{X},.) = 0$ on $\mathrm{Ker}\eta_u$. Since

$$\Omega_u(G(u),\tilde{X}) = (i_G\Omega)_u(\tilde{X}) = (dH)_u(\tilde{X}) = 0$$

and

$$\Omega_u(Y(u),\tilde{X}) = (i_Y\Omega)_u(\tilde{X}) = -A_u(\tilde{X}) = -\eta_u(\tilde{X}) = 0,$$

it follows that $\Omega_u(\tilde{X},.) = 0$ on $T_u(TM)$ and therefore $\tilde{X} = 0$, because Ω is symplectic. Hence $X = 0$. □

The tangent space $T_{q(u)}CM$ can be described through Jacobi fields along the geodesic γ_u with $\gamma_u(0) = x$ and $\dot{\gamma}_u(0) = u$, where $u \in T^1_xM$. We shall use the notations of the previous section. Let $\xi \in T_u(T^1M)$ and let $z : (-\epsilon,\epsilon) \to T^1M$, $\epsilon > 0$, be a smooth curve such that $z(0) = u$ and $\dot{z}(0) = \xi$. If $\gamma = p \circ z$ and $z(t) = (\gamma(t), X(t))$, then $\|X(t)\| \in T^1M$ for all $|t| < \epsilon$. Therefore,

$$g(K_u(\xi), u) = g(\nabla_{\dot{\gamma}(0)}X, X(0)) = 0.$$

This implies that $T_u(T^1M) = \{\xi \in T_u(TM) : g(K_u(\xi), u) = 0\}$. Also we have

$$\begin{aligned} g(I_\xi(t),\dot{\gamma}_u(t)) &= g(\frac{DI_\xi}{dt}(0),\dot{\gamma}_u(0))t + g(I_\xi(0),\dot{\gamma}_u(0)) \\ &= g(K_u(\xi),u) + g(p_{*u}(\xi),v), \end{aligned}$$

for all $|t| < \epsilon$. Combining the above, we conclude that $g(I_\xi(t),\dot{\gamma}_u(t)) = g(p_{*u}(\xi),v)$ for every $\xi \in T_u(T^1M)$. Since $\mathrm{Ker}A_u = \{\xi \in T_u(TM) : g(p_{*u}(\xi),u) = 0\}$, it follows that $i_u(\mathrm{Ker}\eta_u)$ is the linear subspace $I^\perp(u)$ of $I(u)$ consisting of normal Jacobi fields along γ_u. Thus, the chain of linear isomorphisms

$$T_{q(u)}CM \overset{(q_{*u})^{-1}}{\longrightarrow} \mathrm{Ker}\eta_u \overset{i_u}{\longrightarrow} I^\perp(u)$$

gives a natural identification of $T_{q(u)}CM$ with $I^\perp(u)$.

3.3 Kähler manifolds

Let M be a complex manifold of complex dimension n. If $\phi : U \to \phi(U) \subset \mathbb{C}^n$ and $\psi : V \to \phi(V) \subset \mathbb{C}^n$ are two holomorphic charts with $U \cap V \neq \varnothing$, then $\psi \circ \phi^{-1} : \phi(U \cap V) \to \psi(U \cap V)$ is a biholomorphism and $(\psi \circ \phi^{-1})_{*z}$ is $\mathbb{C}$-linear for every $z \in \phi(U \cap V)$. This implies that multiplication by i in $\mathbb{C}^n$ lifts to a well defined almost complex structure $J : TM \to TM$ on M, where TM denotes the tangent bundle of M as a real smooth $2n$-manifold. If $\phi = (z^1, z^2, ..., z^n)$ is a system of holomorphic local coordinates on U and we write $z^j = x^j + iy^j$, $1 \le j \le n$, then

$$\{\frac{\partial}{\partial x^1}, ..., \frac{\partial}{\partial x^n}, \frac{\partial}{\partial y^1}, ..., \frac{\partial}{\partial y^n}\}$$

is a basis at each tangent space T_zM for $z \in U$ and

$$J(\frac{\partial}{\partial x^j}) = \frac{\partial}{\partial y^j}, \quad J(\frac{\partial}{\partial y^j}) = -\frac{\partial}{\partial x^j} \quad 1 \le j \le n.$$

A Riemannian metric g on M is called *hermitian* if it is J-invariant, that is $g_z(J_z(u), J_z(v)) = g_z(u, v)$ for every u, $v \in T_zM$, $z \in M$. For instance the euclidean Riemannian metric on $\mathbb{C}^n$ is hermitian. The proof of Proposition 3.7 of Chapter 2 shows that on a complex manifold there are always hermitian Riemannian metrics.

If g is a hermitian Riemannian metric, the smooth 2-form ω defined by

$$\omega(u, v) = g(J(u), v)$$

is non-degenerate and is called the *fundamental 2-form* of g. In this way each tangent space becomes a Kähler vector space and carries the positive definite hermitian product $h = g - i\omega$, which is J-sesquilinear. We extend h $\mathbb{C}$-bilinearly to the complexified tangent bundle $T_{\mathbb{C}}M = TM \otimes_{\mathbb{R}} \mathbb{C} = T'M \oplus T''M$, where $T'M$ is the eigenspace of the eigenvalue i and $T''M$ is the eigenspace of the eigenvalue $-i$ of J. So $h(J(u), v) = ih(u, v)$ and $h(u, J(v)) = -ih(u, v)$, but $h(iu, v) = ih(u, v) = h(u, iv)$.

Let now

$$\frac{\partial}{\partial z^j} = \frac{1}{2}(\frac{\partial}{\partial x^j} - i\frac{\partial}{\partial y^j}) \in T'M|_U \quad \text{and} \quad \frac{\partial}{\partial \bar{z}^j} = \frac{1}{2}(\frac{\partial}{\partial x^j} + i\frac{\partial}{\partial y^j}) \in T''M|_U.$$

We have

$$h(\frac{\partial}{\partial z^j}, \frac{\partial}{\partial z^k}) = \frac{1}{4}\Big[h(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}) - ih(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial y^k}) - ih(\frac{\partial}{\partial y^j}, \frac{\partial}{\partial x^k}) - h(\frac{\partial}{\partial y^j}, \frac{\partial}{\partial y^k})\Big] = 0$$

and similarly

$$h(\frac{\partial}{\partial \bar{z}^j}, \frac{\partial}{\partial \bar{z}^k}) = 0.$$

Also

$$\begin{aligned} h(\frac{\partial}{\partial z^j}, \frac{\partial}{\partial \bar{z}^k}) &= \frac{1}{4}\Big[h(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}) + ih(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial y^k}) - ih(\frac{\partial}{\partial y^j}, \frac{\partial}{\partial x^k}) + h(\frac{\partial}{\partial y^j}, \frac{\partial}{\partial y^k})\Big] \\ &= \frac{1}{2}\Big[h(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}) + ih(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial y^k})\Big] = \frac{1}{2}\Big[h(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}) + ih(\frac{\partial}{\partial x^j}, J(\frac{\partial}{\partial x^k}))\Big] \\ &= \frac{1}{2}\Big[h(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}) + h(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k})\Big] = h(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}). \end{aligned}$$

It follows that as a tensor h has a local expression

$$h = \sum_{j,k=1}^{n} h_{jk} dz^j \otimes d\bar{z}^k$$

where $h_{jk} = h(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k})$, $1 \leq j, k \leq n$, and $dz^j = dx^j + idy^j$, $d\bar{z}^j = dx^j - idy^j$.

In order to find a local expression of the fundamental 2-form ω we compute

$$\omega(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}) = -\mathrm{Im}h(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}) = -\mathrm{Im}h_{jk}$$

and

$$\omega(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial y^k}) = g(J(\frac{\partial}{\partial x^j}), \frac{\partial}{\partial y^k}) = g(\frac{\partial}{\partial y^j}, \frac{\partial}{\partial y^k}) = \mathrm{Re}h(\frac{\partial}{\partial y^j}, \frac{\partial}{\partial y^k})$$

$$= \mathrm{Re}h(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}) = \mathrm{Re}h_{jk}.$$

Therefore,

$$\omega(\frac{\partial}{\partial z^j}, \frac{\partial}{\partial \bar{z}^k}) = \frac{1}{4}\Big[\omega(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}) + i\omega(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial y^k}) - i\omega(\frac{\partial}{\partial y^j}, \frac{\partial}{\partial x^k}) + \omega(\frac{\partial}{\partial y^j}, \frac{\partial}{\partial y^k})\Big]$$

$$= \frac{1}{2}[-\mathrm{Im}h_{jk} + i\mathrm{Re}h_{jk}] = \frac{i}{2}h_{jk}.$$

Hence

$$\omega|_{\text{locally}} = \frac{i}{2}\sum_{j,k=1}^{n} h_{jk} dz^j \wedge d\bar{z}^k.$$

Note that this local expression agrees with the fact that ω is a real 2-form, because the matrix $(h_{jk})_{1\leq j,k\leq n}$ is hermitian and therefore

$$\bar{\omega} = -\frac{i}{2}\sum_{j,k=1}^{n} \bar{h}_{jk} d\bar{z}^j \wedge dz^k = \frac{i}{2}\sum_{j,k=1}^{n} h_{kj} dz^k \wedge d\bar{z}^j = \omega.$$

The fundamental 2-form ω need not be closed. For example, let $0 < \lambda < 1$ and $T : \mathbb{C}^n \setminus \{0\} \to \mathbb{C}^n \setminus \{0\}$ be the biholomorphism $T(z) = \lambda z$. Then T generates a free and properly discontinuous action of $\mathbb{Z}$ whose orbit space $M_\lambda = \mathbb{C}^n \setminus \{0\}/\mathbb{Z}$ inherits the structure of a complex manifold which makes the quotient map a holomorphic covering map. Writing each $z \in \mathbb{C}^n \setminus \{0\}$ in the form $z = r\zeta$, where $r > 0$ and $\zeta \in S^{2n-1}$, we see that the action is $k \cdot (r, \zeta) = (\lambda^k r, \zeta)$, $k \in \mathbb{Z}$. It follows that M_λ is diffeomorphic to $S^1 \times S^{2n-1}$, which in this way has a complex structure. However $H^2(S^1 \times S^{2n-1}; \mathbb{R}) = 0$, by Künneth's formula, and thus $S^1 \times S^{2n-1}$ admits no symplectic structure. Hence the fundamental 2-form of any hermitian Riemannian metric on $S^1 \times S^{2n-1}$ is not closed.

Recall that the exterior differential of ω is given by

$$d\omega(X, Y, Z) = X\omega(Y, Z) + Y\omega(Z, X) + Z\omega(X, Y)$$

$$-\omega([X, Y], Z) - \omega([Y, Z], X) - \omega([Z, X], Y)$$

for any smooth vector fields X, Y, Z on M. Let ∇ be the Levi-Civita connection of the hermitian Riemannian metric g. Then,

$$X\omega(Y, Z) = g(\nabla_X(JY), Z) + g(JY, \nabla_X Z)$$

$$= g((\nabla_X J)Y, Z) + g(J(\nabla_X Y), Z) + g(JY, \nabla_X Z)$$

$$= g((\nabla_X J)Y, Z) - g(\nabla_X Y, JZ) + g(JY, \nabla_X Z)$$

and

$$\omega([X, Y], Z) = g(J(\nabla_X Y), Z) - g(J(\nabla_Y X), Z) = -g(\nabla_X Y, JZ) + g(\nabla_Y X, JZ).$$

Permuting cyclically and substituting we get

$$d\omega(X, Y, Z) = g((\nabla_X J)Y, Z) + g((\nabla_Y J)Z, X) + g((\nabla_Z J)X, Y).$$

This implies that in case J is parallel with respect to ∇, i.e. $\nabla J = 0$, then $d\omega = 0$ and ω is symplectic. We shall prove that this sufficient condition is also necessary. We shall use the following.

Lemma 3.1. *Let M be a complex manifold with corresponding almost complex structure J and let g be a hermitian Riemannian metric on M with Levi-Civita connection ∇. Then, $J(\nabla_X J) = -(\nabla_X J)J$ and $\nabla_X J$ is skew-adjoint with respect to g for every smooth vector field X on M.*

Proof. For every smooth vector field Y on M we have

$$J(\nabla_X J)Y + (\nabla_X J)JY = J(\nabla_X(JY) - J(\nabla_X Y)) + \nabla_X(J^2Y) - J(\nabla_X(JY)) = 0.$$

This proves the first assertion. To prove the second, for every pair of smooth vector fields Y, Z on M we differentiate the equality $g(JY,Z) + g(Y,JZ) = 0$ in the direction of X to get

$$\begin{aligned} 0 = g(\nabla_X(JY),Z) &+ g(JY,\nabla_X Z) + g(\nabla_X Y,JZ) + g(Y,\nabla_X(JZ)) \\ &= g((\nabla_X J)Y,Z) + g(J(\nabla_X Y),Z) + g(JY,\nabla_X Z) \\ &+g(\nabla_X Y,JZ) + g(Y,(\nabla_X J)Z) + g(Y,J(\nabla_X Z)) \\ &= g((\nabla_X J)Y,Z) + g(Y,(\nabla_X J)Z), \end{aligned}$$

which means that $\nabla_X J$ is skew-adjoint with respect to g. □

Proposition 3.2. *Let M be a complex manifold with corresponding almost complex structure J and let g be a hermitian Riemannian metric on M with Levi-Civita connection ∇. The corresponding fundamental 2-form ω is closed if and only if $\nabla J = 0$.*

Proof. From the above only the converse needs proof. If X, Y, Z are smooth vector fields on M, from Lemma 3.1 we have

$$g((\nabla_X J)Y,JZ) = -g(J(\nabla_X J)Y,Z) = g((\nabla_X J)(JY),Z) = -g(JY,(\nabla_X J)Z).$$

So,

$$\begin{aligned} d\omega(X,Y,JZ)+d\omega(X,JY,Z) &= g((\nabla_X J)Y,JZ)+g((\nabla_Y J)(JZ),X)+g((\nabla_{JZ} J)X,Y) \\ &+g((\nabla_X J)(JY),Z) + g((\nabla_{JY} J)Z,X) + g((\nabla_Z J)X,JY) \\ &= 2g((\nabla_X J)Y,JZ) + g((\nabla_Y J)(JZ),X) + g((\nabla_{JY} J)Z,X) \\ &-g((\nabla_{JZ} J)Y,X) - g((\nabla_Z J)(JY),X) \\ &= 2g((\nabla_X J)Y,JZ) + g((\nabla_Y J)(JZ) + (\nabla_{JY} J)Z - (\nabla_{JZ} J)Y - (\nabla_Z J)(JY),X). \end{aligned}$$

However, again from Lemma 3.1 we have

$$\begin{aligned} &(\nabla_Y J)(JZ) + (\nabla_{JY} J)Z - (\nabla_{JZ} J)Y - (\nabla_Z J)(JY) \\ &= (\nabla_{JY} J)Z - J(\nabla_Y)Z + J(\nabla_Z J)Y - (\nabla_{JZ} J)Y \end{aligned}$$

$$= \nabla_{JY}(JZ) - J(\nabla_{JY}Z) - \nabla_{JZ}(JY) + J(\nabla_{JZ}Y) - J(\nabla_Y J)Z + J(\nabla_Z J)Y$$
$$= [JY, JZ] - J(\nabla_{JY}Z - \nabla_{JZ}Y + (\nabla_Y J)Z - (\nabla_Z J)Y)$$
$$= [JY, JZ] - J(\nabla_{JY}Z - \nabla_{JZ}Y + \nabla_Y(JZ) - J(\nabla_Y Z) - \nabla_Z(JY) + J(\nabla_Z Y))$$
$$= [JY, JZ] - J([Y, JZ] + [JY, Z] - J[Y, Z])$$
$$= [JY, JZ] - [Y, Z] - J[Y, JZ] - J[JY, Z].$$

The tensor $N^J(Y, Z) = [Y, Z] + J[Y, JZ] + J[JY, Z] - [JY, JZ]$ vanishes identically on M, because locally

$$N^J\left(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}\right) = N^J\left(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial y^k}\right) = N^J\left(\frac{\partial}{\partial y^j}, \frac{\partial}{\partial y^k}\right) = 0$$

for every $1 \leq j, k \leq n$. Cosequently,

$$d\omega(X, Y, JZ) + d\omega(X, JY, Z) = 2g((\nabla_X J)Y, JZ).$$

It is now obvious that if $d\omega = 0$, then $\nabla_X J = 0$ for every smooth vector field X on M, since $J : TM \to TM$ is a vector bundle automorphism. □

Remark 3.3. The tensor N^J which appeared in the proof of Proposition 6.2 is called the Nijenhuis tensor and can be defined if we have an almost complex structure J on a smooth manifold M. If J comes from a complex structure on M, then J is called integrable and $N^J = 0$. According to a famous theorem of Newlander and Nirenberg the converse also holds. More precisely, if J is an almost complex structure on a smooth manifold M and $N^J = 0$, then J is integrable and M admits a structure of complex manifold.

A hermitian Riemannian metric g on a complex manifold M is called *Kähler* if its corresponding fundamental 2-form is closed and therefore symplectic. A complex manifold is called Kähler if it admits a Kähler metric.

Example 3.4. The euclidean Riemannian metric on $\mathbb{C}^n$ is hermitian and its corresponding fundamental 2-form is the standard symplectic 2-form

$$\frac{i}{2}\sum_{j=1}^{n} dz^j \wedge d\bar{z}^j = \sum_{j=1}^{n} dx^j \wedge dy^j.$$

So $\mathbb{C}^n$ is a Kähler manifold.

Example 3.5. Let a_1, a_2,..., $a_{2n} \in \mathbb{C}^n$ be linearly independent over $\mathbb{R}$ and $L = a_1\mathbb{Z} + a_2\mathbb{Z} + \cdots + a_{2n}\mathbb{Z}$. The quotient group $\mathbb{C}^n/L$ is the orbit space of the holomorphic action of the group generated by the translations of $\mathbb{C}^n$ by a_j, $1 \leq j \leq n$ and topologically is a $2n$-torus. It inherits a unique complex structure such that the quotient map $p : \mathbb{C}^n \to \mathbb{C}^n/L$ becomes a holomorphic covering map and is called a complex torus. Since translations are isometries of the euclidean metric on $\mathbb{C}^n$, the latter induces a hermitian Riemannian metric on $\mathbb{C}^n/L$ so that p becomes a local isometry. If ω is the corresponding fundamental 2-form, then

$p^*\omega$ is the standard symplectic 2-form on $\mathbb{C}^n$. Since p is a local diffeomorphism and $p^*(d\omega) = d(p^*\omega) = 0$, it follows that ω is closed. Therefore the complex torus $\mathbb{C}^n/L$ is a Kähler manifold.

Example 3.6. The Fubini-Study metric on the complex projective space $\mathbb{C}P^n$ is by its definition a hermitian Riemannian metric, whose fundamental 2-form is closed by Mumford's criterion. Therefore $\mathbb{C}P^n$ is a Kähler manifold.

Example 3.7. Let $D \subset \mathbb{C}^n$ be a bounded, open, connected set and let

$$A^2(D) = \{f \in L^2(D) : f \text{ holomorphic}\}$$

where L^2 is considered with respect to the Lebesgue measure μ and equality of functions in $L^2(D)$ means equality almost everywhere. We equip $A^2(D)$ with the L^2-inner product and norm, and call it the *Bergman space* of D.

Let $A \subset D$ be a compact set and $0 < r < \inf\{\|z - w\| : z \in A, w \in \partial D\}$. Then $S(z,r) \subset A$ for every $z \in A$. For every $z \in A$ and $f \in A^2(D)$ the mean value property of holomorphic functions and Hölder's inequality imply that

$$|f(z)| = \frac{1}{\mu(S(z,r))}\left|\int_{S(z,r)} f d\mu\right| \leq \frac{1}{\mu(S(z,r))}(\mu(S(z,r)))^{1/2}\|f\|_{L^2(S(z,r))}$$

$$\leq (\mu(S(z,r)))^{-1/2}\|f\|_{L^2(D)} = \left(\frac{\pi^n}{\Gamma(n+1)}\right)^{-1/2} r^{-n}\|f\|_{L^2(D)}.$$

So there is a constant $c > 0$ depending only on A and n such that

$$|f(z)| \leq c\|f\|_{L^2(D)}$$

for every $z \in A$ and $f \in A^2(D)$. It follows easily from this inequality that $A^2(D)$ is a closed linear subspace of $L^2(D)$ and therefore is a separable Hilbert space itself. Also, if we take A to be a singleton $\{z\} \subset D$, it implies that the evaluation of $f \in A^2(D)$ at z is a continuous linear functional. From the Riesz representation theorem, there exists a unique $K_z \in A^2(D)$ such that

$$f(z) = \int_D f(\zeta)\overline{K_z(\zeta)}d\mu(\zeta)$$

for every $f \in A^2(D)$. The *Bergman kernel* is the function $K : D \times D \to \mathbb{C}$ with $K(z,\zeta) = \overline{K_z(\zeta)}$. Note that $\overline{K(\zeta,.)} = K_\zeta \in A^2(D)$ for every $\zeta \in D$ and thus

$$\overline{K(\zeta,z)} = \int_D \overline{K(\zeta,t)}K(z,t)d\mu(t) = K(z,\zeta).$$

In particular, the Bergman kernel $K(z,\zeta)$ is holomorphic with respect to z.

It is almost impossible to calculate the Bergman kernel explicitly, unless D is very symmetric. In general it can be constructed from any countable orthonormal basis of $A^2(D)$. Sometimes a suitable orthonormal basis can be chosen, which enables explicit calculation.

Let $\{\phi_j : j \in \mathbb{N}\}$ be an orthonormal basis of $A^2(D)$. For every $z \in$ we have

$$K_z(\zeta) = \sum_{j=1}^{\infty} \left(\int_D K_z(t)\overline{\phi_j(t)} d\mu(t) \right) \phi_j(\zeta) = \sum_{j=1}^{\infty} \overline{\phi_j(z)} \phi_j(\zeta)$$

and the convergence is uniform on compact subsets of D. Thus,

$$K(z, \zeta) = \sum_{j=1}^{\infty} \phi_j(z)\overline{\phi_j(\zeta)}.$$

The convergence is uniform on compact subsets of $D \times D$. In order to prove this it suffices to prove that

$$K(z, z) = \sum_{j=1}^{\infty} |\phi_j(z)|^2$$

uniformly on compact subsets of D, because $2|\phi_j(z)\overline{\phi_j(\zeta)}| \leq |\phi_j(z)|^2 + |\phi_j(\zeta)|^2$. Let $A \subset D$ be a compact set and $0 < r < \inf\{\|z - w\| : z \in A, w \in \partial D\}$. We put $A_r = \bigcup_{z \in A} S(z, r)$. As above, there is a constant $c > 0$ depending on n and r, hence on A, such that

$$|f(z)|^2 \leq c \int_{S(z,r)} |f|^2 d\mu \leq c \int_{A_r} |f|^2 d\mu$$

for every $z \in A_r$ and every $f \in A^2(D)$. Since

$$\sum_{j=1}^{\infty} \int_{A_r} |\phi_j(\zeta)|^2 d\mu(\zeta) = \int_{A_r} K(\zeta, \zeta) d\mu(\zeta)$$

it follows that

$$K(z, z) = \sum_{j=1}^{\infty} |\phi_j(z)|^2$$

uniformly on A_r, hence also on A.

Note that we have $K(z, z) \geq 0$ and the equality holds if and only if $\phi_j(z) = 0$ for all $j \in \mathbb{N}$. This would mean however that $f(z) = 0$ for every $f \in A^2(D)$. This contradiction shows that $K(z, z) > 0$ for every $z \in D$.

Let now h be the complex hermitian metric on D defined by

$$h = \sum_{j,k=1}^{n} \frac{\partial}{\partial z^j} \frac{\partial}{\partial \bar{z}^k} \log K(z, z) dz^j \otimes d\bar{z}^k.$$

The hermitian Riemannian metric $g = \mathrm{Re} h$ is called the *Bergman metric* on D. We shall prove that g is a Kähler metric. First we extend g to complex vector fields of M, that is smooth sections of $T_{\mathbb{C}}M$. Locally, a complex vector field has the form

$$\sum_{j=1}^{n} X_j' \frac{\partial}{\partial z^j} + \sum_{j=1}^{n} X_j'' \frac{\partial}{\partial \bar{z}^j}.$$

Let ∇ be the Levi-Civita connection of the Bergman metric. For convenience we shall use the notation

$$Z_j = \frac{\partial}{\partial z^j} \quad \text{and} \quad Z_{j^*} = \frac{\partial}{\partial \bar{z}^j}.$$

Let Γ^m_{jk}, $\Gamma^{m^*}_{jk}$, $\Gamma^m_{j^*k}$,..., $\Gamma^{m^*}_{j^*k^*}$ be defined by

$$\nabla_{Z_j} Z_k = \sum_{m=1}^{n} \Gamma^m_{jk} Z_m + \sum_{m=1}^{n} \Gamma^{m^*}_{jk} Z_{m^*}$$

and the similar equations for $\nabla_{Z_{j^*}} Z_k$, $\nabla_{Z_j} Z_{k^*}$ and $\nabla_{Z_{j^*}} Z_{k^*}$. Since ∇ is a symmetric connection, $\Gamma^m_{jk} = \Gamma^m_{kj}$ etc. Note that

$$g_{jk} = \mathrm{Re} h(\frac{\partial}{\partial z^j}, \frac{\partial}{\partial z^k}) = 0, \quad g_{j^*k^*} = \mathrm{Re} h(\frac{\partial}{\partial \bar{z}^j}, \frac{\partial}{\partial \bar{z}^k}) = 0$$

and

$$g_{jk^*} = \mathrm{Re} h(\frac{\partial}{\partial z^j}, \frac{\partial}{\partial \bar{z}^k}) = \frac{1}{2} \frac{\partial}{\partial z^j} \frac{\partial}{\partial \bar{z}^k} \log K(z, z).$$

If α, β, γ, δ are any indices, with or without a star, we have

$$\sum_{\delta} g_{\alpha\delta} \Gamma^{\delta}_{\beta\gamma} = \frac{1}{2}(Z_\beta g_{\alpha\gamma} + Z_\gamma g_{\alpha\beta} - Z_\alpha g_{\gamma\beta})$$

from which follows that $\Gamma^{m^*}_{jk} = \Gamma^m_{j^*k} = \Gamma^{m^*}_{jk^*} = \Gamma^m_{j^*k^*} = 0$ for all $1 \leq j, k, m \leq n$. Indeed, choosing α, β, γ to be non-starred indices we get $\Gamma^{m^*}_{jk} = 0$, and if we choose them all starred, then $\Gamma^m_{j^*k^*} = 0$. Choosing α, β, γ to be j^*, j^*, k, respectively, we get $\Gamma^m_{j^*k} = 0$, and j, j, k^*, respectively, we get $\Gamma^{m^*}_{jk^*} = 0$.

Since $\Gamma^{m^*}_{jk} = 0$, we have $\nabla_{Z_j} Z_k \in T'M$ and thus $J(\nabla_{Z_j} Z_k) = i\nabla_{Z_j} Z_k = \nabla_{Z_j}(JZ_k)$, which implies that $(\nabla_{Z_j} J)Z_k = 0$. Similarly, from $\Gamma^m_{j^*k^*} = 0$, we get $(\nabla_{Z_{j^*}} J)Z_{k^*} = 0$. Since $\Gamma^m_{j^*k} = \Gamma^{m^*}_{jk^*} = 0$ and ∇ is symmetric, it follows that $\Gamma^m_{j^*k} = \Gamma^{m^*}_{j^*k} = 0$ and so $\nabla_{Z_{j^*}} Z_k = 0$. Therefore,

$$\nabla_{Z_{j^*}}(JZ_k) = i\nabla_{Z_{j^*}} Z_k = 0 = J(\nabla_{Z_{j^*}} Z_k).$$

Similarly $\nabla_{Z_j} Z_{k^*} = 0$ and $\nabla_{Z_j}(JZ_{k^*}) = J(\nabla_{Z_j} Z_{k^*}) = 0$. By linearity, these show that $\nabla_X J = 0$ for every smooth vector field X. It follows from Proposition 3.2 that the Bergamn metric g is Kähler.

In the case of the unit ball $\mathbb{D}^{2n}$ the functions z^α, where α is a multiindex, form an orthonormal basis of $A^2(\mathbb{D}^{2n})$. Using this orthonormal basis one can calculate

$$K(z, \zeta) = \frac{n!}{\pi^n} \frac{1}{(1 - \langle z, \zeta \rangle)^{n+1}}$$

where $\langle , \rangle$ is the standard hermitian product on $\mathbb{C}^n$. An easy calculation shows that

$$h_{jk}(z) = \frac{\partial}{\partial z^j} \frac{\partial}{\partial \bar{z}^k} \log K(z, z) = \frac{n+1}{(1 - \|z\|^2)^2} [\bar{z}^j z^k + (1 - \|z\|^2)\delta_{jk}]$$

for $z = (z^1, z^2, ..., z^n)$.

In the particular case $n = 1$ this formula becomes

$$g_{11} = h_{11} = \frac{2}{(1-|z|^2)^2}.$$

So, the Bergman metric on the unit disc $\mathbb{D}^2$ coincides with the Poincaré hyperbolic metric.

3.4 Coadjoint orbits

Let G be a Lie group with Lie algera $\mathfrak{g}$ and identity element e. The action of G on itself by conjugation, i.e. $\psi_g(h) = ghg^{-1}$, $g \in G$, fixes e and induces the *adjoint linear representation* $\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g})$ defined by

$$\mathrm{Ad}_g(X) = (\psi_g)_{*e}(X) = \frac{d}{dt}\bigg|_{t=0} g(\exp tX)g^{-1}.$$

Example 4.1. The Lie group $SO(3,\mathbb{R})$ is compact, connected and its Lie algebra $\mathfrak{so}(3,\mathbb{R})$ is isomorphic to the Lie algebra of skew-symmetric linear maps of $\mathbb{R}^3$ with respect to the Lie bracket $[A,B] = AB - BA$, A, $B \in \mathbb{R}^{3\times 3}$.

On the other hand, the map $\widehat{\ } : \mathbb{R}^3 \to \mathfrak{so}(3,\mathbb{R})$ defined by

$$\hat{v} = \begin{pmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{pmatrix}$$

where $v = (v_1, v_2, v_3)$, is a linear isomorphism and $\hat{v} \cdot w = v \times w$, for every v, $w \in \mathbb{R}^3$. This actually characterizes $\widehat{\ }$. So we have

$$(\hat{u}\hat{v} - \hat{v}\hat{u})w = \hat{u}(v \times w) - \hat{v}(u \times w) = u \times (v \times w) - v \times (u \times w) = (u \times v) \times w = (\widehat{u \times v})w.$$

Thus, $\widehat{\ }$ is a Lie algebra isomorphism of the Lie algebra $(\mathbb{R}^3, \times)$ onto $\mathfrak{so}(3,\mathbb{R})$. Using this isomorphism we can describe the exponential map of $SO(3,\mathbb{R})$.

Let $w \in \mathbb{R}^3$, $w \neq 0$, and $\{e_1, e_2, e_3\}$ be an orthonormal basis of $\mathbb{R}^3$ such that $e_1 = w/\|w\|$. The matrix of $\hat{w}$ with respect to this basis is

$$\hat{w} = \|w\| \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

For $t \in \mathbb{R}$ let $\gamma(t)$ be the rotation around the axis determined by w through the angle $t\|w\|$, that is

$$\gamma(t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t\|w\| & -\sin t\|w\| \\ 0 & \sin t\|w\| & \cos t\|w\| \end{pmatrix}.$$

Then,

$$\dot{\gamma}(t) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\|w\|\sin t\|w\| & -\|w\|\cos t\|w\| \\ 0 & \|w\|\cos t\|w\| & -\|w\|\sin t\|w\| \end{pmatrix} = \gamma(t)\hat{w} = (L_{\gamma(t)})_{*I_3}(\hat{w}) = X_{\hat{w}}(\gamma(t)),$$

where $L_{\gamma(t)}$ denotes the left translation on $SO(3,\mathbb{R})$ by $\gamma(t)$ and $X_{\hat{w}}$ the left invariant vector field on $SO(3,\mathbb{R})$ corresponding to $\hat{w}$. In other words, γ is an integral curve of $X_{\hat{w}}$ with $\gamma(0) = I_3$. It follows that $\exp(t\hat{w}) = \gamma(t)$ for every $t \in \mathbb{R}$.

For every $A \in SO(3,\mathbb{R})$ and $v \in \mathbb{R}^3$ we have now

$$\mathrm{Ad}_A(\hat{v}) = \frac{d}{dt}\Big|_{t=0} A(\exp(t\hat{v}))A^{-1} = \frac{d}{dt}\Big|_{t=0} A\gamma(t)A^{-1} = A\gamma(0)\hat{v}A^{-1} = A\hat{v}A^{-1}.$$

Thus,

$$\mathrm{Ad}_A(\hat{v})w = A\hat{v}(A^{-1}w) = A(v \times A^{-1}w) = Av \times w$$

for every $w \in \mathbb{R}^3$, since $\det A = 1$. Hence $\mathrm{Ad}_A(\hat{v}) = \widehat{Av}$, and identifying $\mathbb{R}^3$ with $\mathfrak{so}(3,\mathbb{R})$ via $\widehat{\ }$ we conclude that $\mathrm{Ad}_A = A$.

Let now $\mathrm{ad} = (\mathrm{Ad})_{*e} : \mathfrak{g} \to T_e\mathrm{Aut}(\mathfrak{g}) \cong \mathrm{End}(\mathfrak{g})$, that is

$$\mathrm{ad}_X = (\mathrm{Ad})_{*e}(X) = \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{\exp(tX)}$$

for every $X \in \mathfrak{g}$. If we denote by X_L the left invariant vector field corresponding to X and $(\phi_t)_{t\in\mathbb{R}}$ its flow, then for every $Y \in \mathfrak{g} \cong T_0\mathfrak{g}$ we have

$$\mathrm{ad}_X(Y) = \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{\exp(tX)}(Y) = \frac{d}{dt}\Big|_{t=0} (\psi_{\exp(tX)})_{*e}(Y) =$$

$$\frac{d}{dt}\Big|_{t=0} (R_{\exp(-tX)} \circ L_{\exp(tX)})_{*e}(Y) = \frac{d}{dt}\Big|_{t=0} (R_{\exp(-tX)})_{*\exp(tX)} \circ (L_{\exp(tX)})_{*e}(Y) =$$

$$\frac{d}{dt}\Big|_{t=0} (R_{\exp(-tX)})_{*\exp(tX)}(Y_L(\exp(tX)) = \frac{d}{dt}\Big|_{t=0} (\phi_{-t})_{*\phi_t(e)}(Y_L(\phi_t(e)) = [X,Y],$$

since $\phi_t(g) = g\exp(tX) = R_{\exp(tX)}(g)$ for every $g \in G$, where R denotes right translation.

As usual, the adjoint representation induces a representation $\mathrm{Ad}^* : G \to \mathrm{Aut}(\mathfrak{g}^*)$ on the dual of the Lie algebra defined by $\mathrm{Ad}^*_g(a) = a \circ \mathrm{Ad}_{g^{-1}}$, $a \in \mathfrak{g}^*$, which is called the *coadjoint representation* of G.

Example 4.2. Continuing from Example 4.1, we shall describe the coadjoint representation of $SO(3,\mathbb{R})$. The transpose of the linear isomorphism $\widehat{\ }$ induces an isomorphism from $\mathfrak{so}(3,\mathbb{R})^*$ to $(\mathbb{R}^3)^*$ and the latter can be identified naturally with $\mathbb{R}^3$ via the euclidean inner product. The composition of these two isomorphisms gives a way to identify $\mathfrak{so}(3,\mathbb{R})^*$ with $\mathbb{R}^3$ and then, for every v, $w \in \mathbb{R}^3$ we have $\hat{v}^*(\hat{w}) = \langle v, w\rangle$, where $\hat{v}^*$ is the dual of $\hat{v}$ and $\langle , \rangle$ is the euclidean inner product. Now

$$\mathrm{Ad}^*_A(\hat{v}^*)(\hat{w}) = \hat{v}^*(\mathrm{Ad}_{A^{-1}}(\hat{w})) = \langle v, A^{-1}w\rangle = \langle Av, w\rangle,$$

for every $A \in SO(3,\mathbb{R})$, since the transpose of A is A^{-1}. This shows that $\mathrm{Ad}^*_A = A$ via the above identification. Note that the orbit of the point $\hat{v}^* \in \mathfrak{so}(3,\mathbb{R})^* \cong \mathbb{R}^3$ is the set $\{Av : A \in SO(3,\mathbb{R})\}$, which is the sphere of radius $\|v\|$ centered at 0.

The orbit $\mathcal{O}_\mu$ of $\mu \in \mathfrak{g}^*$ under the coadjoint representation is an immersed submanifold of $\mathfrak{g}^*$, since the action is smooth. If G_μ is the isotropy group of μ, then the map $\mathrm{Ad}^*_{\cdot}(\mu) : G/G_\mu \to \mathcal{O}_\mu$ taking the coset gG_μ to $\mu \circ \mathrm{Ad}_{g^{-1}}$ is a well defined, injective, smooth immersion of the homogeneous space G/G_μ onto $\mathcal{O}_\mu \subset \mathfrak{g}^*$. If the Lie group G is compact, then $\mathcal{O}_\mu$ is an embedded submanifold of $\mathfrak{g}^*$ and the above map an embedding. If however G is not compact, $\mathcal{O}_\mu$ may not be embedded.

Lemma 4.3. *If $\mu \in \mathfrak{g}^*$, then the tangent space of $\mathcal{O}_\mu$ is*

$$T_\mu \mathcal{O}_\mu = \{\mu \circ \mathrm{ad}_X : X \in \mathfrak{g}\}.$$

Proof. Let $\gamma : \mathbb{R} \to G$ be a smooth curve with $\dot{\gamma}(0) = X$. For instance, let $\gamma(t) = \exp(tX)$, in which case $\gamma(t)^{-1} = \exp(-tX)$. Then $\mu(t) = \mu \circ \mathrm{Ad}_{\gamma(t)^{-1}}$ is a smooth curve with values in $\mathcal{O}_\mu \subset \mathfrak{g}^*$ and $\mu(0) = \mu$. If $Y \in \mathfrak{g}$, then $\mu(t)(Y) = \mu(\mathrm{Ad}_{\gamma(t)^{-1}}(Y))$ for every $t \in \mathbb{R}$ and defferentiating at 0 we get

$$\mu'(0)(Y) = \mu(\mathrm{ad}_{(-X)}(Y)) = -\mu(\mathrm{ad}_X(Y)),$$

taking into account the natural identification $T_\mu \mathfrak{g}^* \cong \mathfrak{g}^*$. □

Example 4.4. In the case of the Lie group $SO(3, \mathbb{R})$, for every v, $w \in \mathbb{R}^3 \cong \mathfrak{so}(3, \mathbb{R})$ and $\mu \in \mathbb{R}^3 \cong \mathfrak{so}(3, \mathbb{R})^*$ we have

$$\mu(\mathrm{ad}_{\hat{v}}(\hat{w})) = \langle \mu, v \times w \rangle = \langle \mu \times v, w \rangle.$$

It follows that $T_\mu \mathcal{O}_\mu = \{\mu \times v : v \in \mathbb{R}^3\}$, which is indeed the orthogonal plane to μ, i.e. the tangent plane of the sphere of center 0 and radius $\|\mu\|$ at μ.

The proof of Lemma 4.3 shows that for every $X \in \mathfrak{g}$, the fundamental vector field $X_{\mathfrak{g}^*}$ of the coadjoint action induced by X is given by the formula

$$X_{\mathfrak{g}^*}(\mu) = \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}^*_{\exp(tX)}(\mu) = -\mu \circ \mathrm{ad}_X.$$

Obviously, $T_\mu \mathcal{O}_\mu = \{X_{\mathfrak{g}^*}(\mu) : X \in \mathfrak{g}\}$. Note that if X, $X' \in \mathfrak{g}$ are such that $X_{\mathfrak{g}^*}(\mu) = X'_{\mathfrak{g}^*}(\mu)$, then

$$-\mu([X, Y]) = X_{\mathfrak{g}^*}(\mu)(Y) = X'_{\mathfrak{g}^*}(\mu)(Y) = -\mu([X', Y])$$

for every $Y \in \mathfrak{g}$. So there is a well defined 2-form ω^- on the coadjoint orbit $\mathcal{O} = \mathcal{O}_\mu$ such that

$$\omega^-_\mu(X_{\mathfrak{g}^*}(\mu), Y_{\mathfrak{g}^*}(\mu)) = -\mu([X, Y])$$

for every $\mu \in \mathcal{O}$ and X, $Y \in \mathfrak{g}$. We call ω^- the *Kirillov 2-form* on $\mathcal{O}$.

The Kirillov 2-form ω^- is non-degenerate, because if $\omega^-_\mu(X_{\mathfrak{g}^*}(\mu), Y_{\mathfrak{g}^*}(\mu)) = 0$ for every $Y_{\mathfrak{g}^*}(\mu) \in T_\mu \mathcal{O}$, then $X_{\mathfrak{g}^*}(\mu)(Y) = -\mu([X, Y]) = 0$ for every $Y \in \mathfrak{g}$. This means $X_{\mathfrak{g}^*}(\mu) = 0$. In order to prove that ω^- is symplectic, it remains to show that it is closed. For this we shall need a series of lemmas.

First note that $\mathrm{Ad}_g[X,Y] = [\mathrm{Ad}_g(X), \mathrm{Ad}_g(Y)]$ for every X, $Y \in \mathfrak{g}$ and $g \in G$.

Lemma 4.5. *$(\mathrm{Ad}_g(X))_{\mathfrak{g}^*} = \mathrm{Ad}_g^* \circ X_{\mathfrak{g}^*} \circ \mathrm{Ad}_{g^{-1}}^*$ for every $X \in \mathfrak{g}$ and $g \in G$.*

Proof. Let $\gamma : \mathbb{R} \to G$ be a smooth curve $\dot\gamma(0) = X$. For instance $\gamma(t) = \exp(tX)$, and then

$$\mathrm{Ad}_g(X) = \frac{d}{dt}\Big|_{t=0} g\gamma(t)g^{-1}.$$

Therefore,

$$(\mathrm{Ad}_g(X))_{\mathfrak{g}^*}(\mu) = \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}^*_{g\gamma(t)g^{-1}}(\mu) = \frac{d}{dt}\Big|_{t=0} (\mathrm{Ad}_g^* \circ \mathrm{Ad}^*_{\gamma(t)} \circ \mathrm{Ad}^*_{g^{-1}})(\mu)$$
$$= (\mathrm{Ad}_g^* \circ X_{\mathfrak{g}^*} \circ \mathrm{Ad}^*_{g^{-1}})(\mu). \quad \square$$

Lemma 4.6. *The Kirillov 2-form is Ad^*-invariant.*

Proof. Let $\mu \in \mathfrak{g}^*$ and $\nu = \mathrm{Ad}_g^*(\mu)$, $g \in G$. By Lemma 4.5,

$$(\mathrm{Ad}_g(X))_{\mathfrak{g}^*}(\nu) = \mathrm{Ad}_g^*(X_{\mathfrak{g}^*}(\mu)).$$

Thus, for every X, $Y \in \mathfrak{g}$ we have

$$((\mathrm{Ad}_g^*)^*\omega^-)_\mu(X_{\mathfrak{g}^*}(\mu), Y_{\mathfrak{g}^*}(\mu)) = \omega_\nu^-((\mathrm{Ad}_g(X))_{\mathfrak{g}^*}(\nu), (\mathrm{Ad}_g(Y))_{\mathfrak{g}^*}(\nu))$$
$$= -\nu([\mathrm{Ad}_g(X), \mathrm{Ad}_g(Y)]) = -\nu(\mathrm{Ad}_g[X,Y]) = -\mu([X,Y])$$
$$= \omega_\mu^-(X_{\mathfrak{g}^*}(\mu), Y_{\mathfrak{g}^*}(\mu)). \quad \square$$

For every $\nu \in \mathfrak{g}^*$ we have a well defined 1-form ν_L on G such that

$$(\nu_L)_g = \nu \circ (L_{g^{-1}})_{*g} \in T_g^*G.$$

Moreover, ν_L is left invariant, because for every $h \in G$ we have

$$(L_h^*\nu_L)_g = (\nu_L)_{L_h(g)} \circ (L_h)_{*g} = \nu \circ ((L_{g^{-1}h^{-1}})_{*hg} \circ (L_h)_{*g})$$
$$= \nu \circ (L_{g^{-1}h^{-1}} \circ L_h)_{*g} = \nu_L(g).$$

Obviously, $i_{X_L}\nu_L$ is constant and equal to $\nu(X)$ for every $X \in \mathfrak{g}$.

Let $\nu \in \mathcal{O}$ and $\phi_\nu : G \to \mathcal{O}$ be the submersion $\phi_\nu(g) = \mathrm{Ad}_g^*(\nu)$. The 2-form $\sigma = \phi_\nu^*\omega^-$ on G is left invariant, because

$$L_g^*\sigma = (\phi_\nu \circ L_g)^*\omega^- = (\mathrm{Ad}_g^* \circ \phi_\nu)^*\omega^- = \phi_\nu^*((\mathrm{Ad}_g^*)^*\omega^-) = \phi_\nu^*\omega^- = \sigma$$

for every $g \in G$, since ω^- is Ad^*-invariant and $\phi_\nu \circ L_g = \mathrm{Ad}_g^* \circ \phi_\nu$.

Lemma 4.7. *For every X, $Y \in \mathfrak{g}$ we have $\sigma(X_L, Y_L) = -\nu_L([X_L, Y_L])$.*

Proof. First we observe that

$$(\phi_\nu^*\omega^-)_e(X,Y) = \omega_\nu^-((\phi_\nu)_{*e}(X), (\phi_\nu)_{*e}(Y)) = \omega_\nu^-(X_{\mathfrak{g}^*}(\nu), Y_{\mathfrak{g}^*}(\nu)) = -\nu([X,Y]).$$

Therefore,

$$\sigma(X_L, Y_L)(e) = (\phi_\nu^* \omega^-)_e(X, Y) = -\nu([X, Y]) = -\nu_L([X_L, Y_L])(e).$$

Since the smooth functions $\sigma(X_L, Y_L)$, $-\nu_L([X_L, Y_L]) : G \to \mathbb{R}$ are left invariant and take the same value at e, they must be identical. □

Note that

$$(d\nu_L)(X_L, Y_L) = X_L(\nu_L(Y_L)) - Y_L(\nu_L(X_L)) - \nu_L([X_L, Y_L]) = -\nu_L([X_L, Y_L]),$$

since the functions $\nu_L(Y_L) = i_{Y_L}\nu_L$ and $\nu_L(X_L) = i_{X_L}\nu_L$ are constant.

Lemma 4.8. *The 2-form σ is exact and $\sigma = d\nu_L$.*

Proof. Since σ is left invariant, for any two smooth vector fields X, Y on G we have

$$\begin{aligned}
\sigma(X, Y)(g) &= (L_{g^{-1}}^* \sigma)_g(X(g), Y(g)) = \sigma_e((L_{g^{-1}})_{*g}(X(g)), (L_{g^{-1}})_{*g}(Y(g))) \\
&= \sigma(X'_L, Y'_L)(e) \qquad (\text{setting } X' = (L_{g^{-1}})_{*g}(X(g)) \text{ and similarly for } Y') \\
&= (d\nu_L)(X'_L, Y'_L)(e) \qquad (\text{by Lemma 4.7}) \\
&= (d\nu_L)_g((L_g)_{*e}(X'), (L_g)_{*e}(Y')) \qquad (\text{since } \nu_L \text{ is left invariant}) \\
&= (d\nu_L)_g(X(g), Y(g)) = (d\nu_L)(X, Y)(g). \qquad \square
\end{aligned}$$

Proposition 4.9. *The Kirillov 2-form ω^- on $\mathcal{O}$ is closed and therefore symplectic.*

Proof. By Lemma 4.8, $d(\phi_\nu^* \omega^-) = d\sigma = d(d\nu_L)) = 0$. Hence $\phi_\nu^*(d\omega^-) = 0$. But ϕ_ν^* is injective, since ϕ_ν is a submersion. It follows that $d\omega^- = 0$. □

Corollary 4.10. *Every orbit of the coadjoint action of a Lie group G on its dual Lie algebra $\mathfrak{g}^*$ has even dimension.* □

We shall end this section with a couple of illustrating examples.

Example 4.11. As we saw in Example 4.2, if $\mu \in \mathfrak{so}(3, \mathbb{R})^* \cong \mathbb{R}^3$, then $\mathcal{O}_\mu$ is the sphere centered at 0 with radius $\|\mu\|$. Let v, $w \in \mathfrak{so}(3, \mathbb{R}) \cong \mathbb{R}^3$. Then $v_{\mathbb{R}^3} = \mu \times v \in T_\mu \mathcal{O}_\mu$ and $w_{\mathbb{R}^3} = \mu \times w \in T_\mu \mathcal{O}_\mu$. Hence the Kirillov 2-form on $\mathcal{O}_\mu$ is given by the formula

$$\omega_\mu^-(v_{\mathbb{R}^3}, w_{\mathbb{R}^3}) = -\langle \mu, v \times w \rangle.$$

Since $\mathcal{O}_\mu$ is a sphere, its area element is given by the formula

$$dA(v, w) = \langle N, v \times w \rangle,$$

where N is the outer unit normal vector. It follows that

$$dA(\mu \times v, \mu \times w) = \langle \frac{1}{\|\mu\|}\mu, (\mu \times v) \times (\mu \times w) \rangle = \langle \frac{1}{\|\mu\|}\mu, \langle \mu, \mu \times w \rangle v - \langle v, \mu \times w \rangle \mu \rangle$$

$$= -\|\mu\|\langle v, \mu \times w\rangle = \|\mu\|\langle \mu, v \times w\rangle,$$

where we have used the property $(a \times b) \times c = \langle a, c\rangle b - \langle b, c\rangle a$ of the vector product in $\mathbb{R}^3$. This shows that

$$\omega^- = -\frac{1}{\|\mu\|} dA.$$

Example 4.12. The connected Lie group of the orientation preserving affine transformations of $\mathbb{R}$ is represented as a group of matrices by

$$G = \{\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} : a > 0, b \in \mathbb{R}\}.$$

Its Lie algebra is

$$\mathfrak{g} = \{\begin{pmatrix} x & y \\ 0 & 0 \end{pmatrix} : x, y \in \mathbb{R}\} \cong \mathbb{R}^2$$

with Lie bracket $[A, B] = AB - BA$. The exponential map is computed as follows. Let x, $y \in \mathbb{R}$ with $x \neq 0$. Let $\gamma : \mathbb{R} \to G$ be the smooth curve defined by

$$\gamma(t) = \begin{pmatrix} e^{tx} & \frac{y}{x}(e^{tx} - 1) \\ 0 & 1 \end{pmatrix}.$$

Then $\gamma(0) = I_2$ and

$$\gamma'(t) = \begin{pmatrix} xe^{tx} & ye^{tx} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} e^{tx} & \frac{y}{x}(e^{tx} - 1) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x & y \\ 0 & 0 \end{pmatrix} = (L_{\gamma(t)})_{*I_2} \begin{pmatrix} x & y \\ 0 & 0 \end{pmatrix}.$$

This shows that if

$$W = \begin{pmatrix} x & y \\ 0 & 0 \end{pmatrix},$$

then $\exp(tW) = \gamma(t)$. If now

$$A = \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix},$$

then

$$\mathrm{Ad}_A(W) = \frac{d}{dt}\Big|_{t=0} A\gamma(t)A^{-1} = AWA^{-1} = \begin{pmatrix} x & ay - bx \\ 0 & 0 \end{pmatrix}.$$

So the orbit of W under the adjoint action is the line

$$\{\begin{pmatrix} x & t \\ 0 & 0 \end{pmatrix} : t \in \mathbb{R}\},$$

when $x \neq 0$. If $x = 0$ and $y > 0$, it is the upper half-line, and if $y < 0$ it is the lower half-line. We see now that the adjoint representation cannot be equivalent (in any sense) to the coadjoint representation, since the orbits of the latter have even dimension. In the sequel, we shall find the orbits of the coadjoint representation and the Kirillov 2-form.

Let $\mu \in \mathfrak{g}^* \cong \mathbb{R}^2$, the isomorphism given by euclidean inner product. This means that if $\mu = (\alpha^*, \beta^*)$, then $\mu(W) = \alpha^* x + \beta^* y$. So,

$$\mathrm{Ad}^*_{A^{-1}}(\mu)(W) = \mu(\mathrm{Ad}_A(W)) = \alpha^* x + \beta^*(ay - bx) = (\alpha^* - \beta^* b)x + \beta^* ay$$

which implies that $\mathrm{Ad}^*_{A^{-1}}(\alpha^*, \beta^*) = (\alpha^* - \beta^* b, \beta^* a)$. Therefore,

$$\mathcal{O}_\mu = \{(\alpha^* - \beta^* b, \beta^* a) : a > 0, b \in \mathbb{R}\}$$

for $\mu = (\alpha^*, \beta^*)$, as before.

If $\beta^* \neq 0$, then for $\beta^* > 0$ the coadjoint orbit $\mathcal{O}_\mu$ is the open upper half plane and for $\beta^* < 0$ it is the open lower half plane. For $\beta^* = 0$ we have $\mathcal{O}_\mu = \{(\alpha^*, 0)\}$. In the case $\beta^* \neq 0$, the Kirillov 2-form ω^- on $\mathcal{O}_\mu$ satisfies

$$\omega^-_\mu((W_1)_{\mathfrak{g}^*}, (W_2)_{\mathfrak{g}^*}) = -\mu([W_1, W_2]) = -\mu \begin{pmatrix} 0 \\ x_1 y_2 - x_2 y_1 \end{pmatrix} = \beta^*(x_2 y_1 - x_1 y_2),$$

where

$$W_j = \begin{pmatrix} x_j & y_j \\ 0 & 0 \end{pmatrix}, \quad j = 1, 2,$$

since

$$[W_1, W_2] = \begin{pmatrix} 0 & x_1 y_2 - x_2 y_1 \\ 0 & 0 \end{pmatrix}.$$

Consequently, $\omega^- = -\beta^* d\alpha^* \wedge d\beta^*$.

3.5 Homogeneous symplectic manifolds

Let $\mathfrak{g}$ be a (real) Lie algebra and let $\Lambda^k(\mathfrak{g})$, $k \geq 0$, denote the vector space of all skew-symmetric covariant k-tensors on $\mathfrak{g}$. For every $k \geq 0$, let $\delta : \Lambda^k(\mathfrak{g}) \to \Lambda^{k+1}(\mathfrak{g})$ be the linear map defined by

$$(\delta\omega)(X_0, X_1, ..., X_k) = \sum_{i<j} (-1)^{i+j} \omega([X_i, X_j], X_0, ..., \hat{X}_i, ..., \hat{X}_j, ..., X_k).$$

For $k = 0$, we have $\delta = 0$, and for $k = 1$ we have

$$(\delta\omega)(X_0, X_1) = -\omega([X_0, X_1]).$$

A standard computation shows that $\delta \circ \delta = 0$. Let $Z^k(\mathfrak{g}) = \Lambda^k(\mathfrak{g}) \cap \mathrm{Ker}\delta$ and $B^k(\mathfrak{g}) = \Lambda^k(\mathfrak{g}) \cap \mathrm{Im}\delta$. The quotient $H^k(\mathfrak{g}) = Z^k(\mathfrak{g})/B^k(\mathfrak{g})$ is called the Lie algebra k-cohomology of $\mathfrak{g}$. Obviously, $H^0(\mathfrak{g}) = \{0\}$.

Let G be a Lie group with Lie algebra $\mathfrak{g}$. Then $\Lambda^1(\mathfrak{g}) = \mathfrak{g}^*$. For every $k \geq 0$ the space $\Lambda^k(\mathfrak{g})$ can be identified in the obvious way with the vector space of left invariant differential k-forms on G. The adjoint representation induces a left action Ad^* on $\Lambda^k(\mathfrak{g})$, which for $k = 1$ is just the coadjoint representation. So $\mathrm{Ad}^*_g(\theta) = (\mathrm{Ad}_{g^{-1}})^*\theta$, for $\theta \in \Lambda^k(\mathfrak{g})$. It is evident that $Z^k(\mathfrak{g})$ is an Ad^*-invariant subspace.

Lemma 5.1. *Let G be a Lie group with Lie algebra $\mathfrak{g}$. If $H^1(\mathfrak{g}) = \{0\}$ and $H^2(\mathfrak{g}) = \{0\}$, then the* Ad*- *actions of G on $\mathfrak{g}^*$ and $Z^2(\mathfrak{g})$ are isomorphic.*

Proof. Since $H^2(\mathfrak{g}) = \{0\}$, for every $\theta \in Z^2(\mathfrak{g})$ there exists $\mu \in \mathfrak{g}^*$ such that $\delta\mu = \theta$. On the other hand, since $H^1(\mathfrak{g}) = \{0\}$, we have $Z^1(\mathfrak{g}) = B^1(\mathfrak{g}) = \delta(\Lambda^0(\mathfrak{g})) = 0$. Thus, if $\delta\mu = 0$, then $\mu = 0$. It follows that $\delta : \mathfrak{g}^* \to Z^2(\mathfrak{g})$ is an isomorphism. It is obvious that δ is Ad*-equivariant. □

Let now (M, ω) be a symplectic manifold, G a Lie group and $\phi : G \times M \to M$ a smooth, symplectic action. Let $\phi_g = \phi(g, .)$ and $\phi^p = \phi(., p)$ for $g \in G$ and $p \in M$. Then, $\phi_g \circ \phi^p = \phi^p \circ L_g$ and $\phi^{\phi_g(p)} = \phi^p \circ R_g$. The closed 2-form $(\phi^p)^*\omega$ on G is left invariant, because

$$(L_g)^*((\phi^p)^*\omega) = (\phi^p \circ L_g)^*\omega = (\phi_g \circ \phi^p)^*\omega = (\phi^p)^*((\phi_g)^*\omega) = (\phi^p)^*\omega,$$

since ϕ_g is a symplectomorphism. Let $\Psi : M \to Z^2(\mathfrak{g})$ be the smooth map defined by

$$\Psi(p) = ((\phi^p)^*\omega)_e.$$

Since $(\phi^p)^*\omega$ is left invariant, Ψ is equivariant. Indeed,

$$((\phi^{\phi_g(p)})^*\omega)_e = ((\phi^p \circ R_g)^*\omega)_e = (R_g^*((\phi^p)^*\omega))_e = (\mathrm{Ad}_{g^{-1}})^*((\phi^p)^*\omega)_e.$$

In case the action is transitive, then $\Psi(M)$ is precisely one orbit in $Z^2(\mathfrak{g})$. Recall that if the action is transitive, then M is diffeomorphic to the homogeneous space G/H, where H is the isotropy group of any point of M.

Proposition 5.2. *Let G be a Lie group with Lie algebra $\mathfrak{g}$. If $H^1(\mathfrak{g}) = \{0\}$ and $H^2(\mathfrak{g}) = \{0\}$, then for every $\theta \in Z^2(\mathfrak{g})$ there exists a homogenous symplectic G-manifold M such that $\Psi(M) = \mathcal{O}_\theta$, where $\mathcal{O}_\theta$ is the orbit of θ under the* Ad*-*action of G on $Z^2(\mathfrak{g})$.*

Proof. Let θ_L denote the left invariant 2-form on G defined by θ. According to Lemma 5.1, there exists a unique $\mu \in \mathfrak{g}^*$ such that $\delta\mu = \theta$. Let G_μ be the isotropy group of μ under the coadjoint representation. Then,

$$G_\mu = \{g \in G : \mathrm{Ad}_g^*\mu = \mu\} = \{g \in G : (L_{g^{-1}} \circ R_g)^*\mu_L = \mu_L\}$$
$$= \{g \in G : (R_g)^*\mu_L = \mu_L\}.$$

If $X \in \mathfrak{g}$, the flow $(\psi_t)_{t\in\mathbb{R}}$ of X_L is given by the formula

$$\psi_t(g) = g\exp(tX) = R_{\exp(tX)}(g).$$

It follows that the Lie algebra of G_μ is $\mathfrak{g}_\mu = \{X \in \mathfrak{g} : L_{X_L}\mu_L = 0\}$. But $i_{X_L}\mu_L$ is a constant function, and therefore $L_{X_L}\mu_L = i_{X_L}(d\mu_L) = i_{X_L}\theta_L$. So,

$$\mathfrak{g}_\mu = \{X \in \mathfrak{g} : i_{X_L}\theta_L = 0\}.$$

Let H_θ be the connected component of G_μ wich contains e. The Lie algebra of H_θ is $\mathfrak{g}_\mu$, and of course H_θ is a closed Lie subgroup of G. Recall that the homogeneous space $M = G/H_\theta$ becomes a smooth manifold in a unique way such that the quotient map $\pi : G \to M$ is smooth and it has local cross sections. In particular π is a submersion.

Every $g \in G$ is contained in the domain U of a chart $(U, x^1, x^2, ..., x^m)$ of G, where $m = \dim G$, which maps U diffeomorphically onto $\mathbb{R}^m$, such that $(\pi(U), x^1, x^2, ..., x^n)$ is a chart on M, where $m - n = \dim H_\theta$, and in these local coordinates $\pi(x^1, x^2, ..., x^m) = (x^1, x^2, ..., x^n)$. There are smooth functions $a_{ij} : U \to \mathbb{R}$, $1 \le i < j \le m$, such that

$$\theta_L|_U = \sum_{1 \le i < j \le m} a_{ij} dx^i \wedge dx^j.$$

The tangent space of the submanifold gH_θ of G at g has basis $\{\frac{\partial}{\partial x^{n+1}}, ..., \frac{\partial}{\partial x^m}\}$. Since $\tilde{\theta}_L(\frac{\partial}{\partial x^l}) = 0$ for $n < l \le m$, it follows that

$$\sum_{l<j} a_{lj} dx^j - \sum_{i<l} a_{il} dx^i = 0.$$

Therefore $a_{ij} = 0$, when $i > n$ or $j > n$, which means that

$$\theta_L|_U = \sum_{1 \le i < j \le n} a_{ij} dx^i \wedge dx^j$$

and then

$$0 = d(\theta_L) = \sum_{1 \le i < j \le n} \left(\sum_{l=1}^{m} \frac{\partial a_{ij}}{\partial x^l} dx^l \right) \wedge dx^i \wedge dx^j.$$

Hence $\frac{\partial a_{ij}}{\partial x^l} = 0$ for $n < l \le m$, and so the functions a_{ij} do not depend on the coordinates $x^{n+1}, ..., x^m$. This implies that θ_L descends to a well defined 2-form $\bar{\theta}_L$ on $\pi(U)$ given by the same formula. It is standard and easy, but somewhat tedious, to show that there is a well defined closed 2-form ω on M such that $\pi^*\omega = \theta_L$ and $\omega|_{\pi(U)} = \bar{\theta}_L$. Moreover, ω is non-degenerate, because $\tilde{\omega}_{\pi(g)}(\pi_{*g}(v)) = 0$ if and only if $(\tilde{\theta}_L)_g(v) = 0$ or equivalently $v \in T_g(gH_\theta)$, that is $\pi_*(v) = 0$. So we have so far shown that (M, ω) is a symplectic manifold.

Let $\phi : G \times M \to M$ be the natural transitive left action of G on M so that $\phi_g(hH_\theta) = (gh)H_\theta$, $g \in G$. Then $\phi_g \circ \pi = \pi \circ L_g$ and therefore

$$\pi^*(\phi_g^*\omega) = (\phi_g \circ \pi)^*\omega = (\pi \circ L_g)^*\omega = L_g^*(\pi^*\omega) = L_g^*\theta_L = \theta_L = \pi^*\omega.$$

Since π is a submersion, π^* is injective in the level of forms, and hence $\phi_g^*\omega = \omega$. This shows that the action is symplectic. In order to complete the proof, it remains to show that $\Psi(M) = \mathcal{O}_\theta$. If $p = gH_\theta \in M$, then $\phi^p = \pi \circ R_g$, and for every X, $Y \in \mathfrak{g}$ we have

$$\Psi(p)(X, Y) = ((\phi^p)^*\omega)_e(X, Y) = \omega_{\pi(g)}(\pi_{*g}((R_g)_{*e}(X)), \pi_{*g}((R_g)_{*e}(Y)))$$

$$= (\pi^*\omega)_g((R_g)_{*e}(X), (R_g)_{*e}(Y)) = (\theta_L)_g((R_g)_{*e}(X), (R_g)_{*e}(Y))$$

$$= \theta((L_{g^{-1}} \circ R_g)_{*e}(X), (L_{g^{-1}} \circ R_g)_{*e}(X)) = \mathrm{Ad}_g^*(\theta)(X, Y).$$

In other words $\Psi(gH_\theta) = \mathrm{Ad}_g^*(\theta)$ for every $g \in G$ and therefore $\Psi(M) = \mathcal{O}_\theta$. $\square$

Let again $\theta \in Z^2(\mathfrak{g})$ and suppose that (M, ω) is a symplectic manifold, on which the Lie group G with Lie algebra $\mathfrak{g}$ acts transitively, symplectically and such that $\Psi(M) = \mathcal{O}_\theta$. Then M is diffeomorphic to the homogeneous space G/H, where H is the isotropy group of any point of M and necessarily $\theta_L = (\phi^p)^*\omega$, where $p \in M$. The Lie algebra of H is

$$\mathfrak{h} = \{X \in \mathfrak{g} : i_X\theta = 0\},$$

If H_θ is the connected component of H which contains e, then $\Psi(G/H_\theta) = \mathcal{O}_\theta$, as the proof of Proposition 6.2 shows. The homogeneous space G/H_θ is a covering space of M. These show that if M amd N are two homogeneous, symplectic G-manifolds with $\Psi(M) = \Psi(N)$, then N is a covering space of M or vice versa.

Summarizing the results of this section, we have proved the following.

Theorem 5.3. (Kostant-Souriau) *Let G be a Lie group with Lie algebra $\mathfrak{g}$ such that $H^1(\mathfrak{g}) = \{0\}$ and $H^2(\mathfrak{g}) = \{0\}$. Then, up to covering spaces, the homogeneous, symplectic G-manifolds are in one-to-one, onto correspondence with the coadjoint orbits in $\mathfrak{g}^*$.* $\square$

Before we end this section, we need to make some remarks about the assumptions in the Kostant-Souriau theorem. If G is a compact, connected Lie group with Lie algebra $\mathfrak{g}$, then $H^k(\mathfrak{g})$ is isomorphic to the k-th deRham cohomology, and so to the k-th real singular cohomology $H^k(G;\mathbb{R})$ of G for every $k \geq 0$. Moreover, in this case the condition $H^1(G;\mathbb{R}) = 0$ implies that $H^2(G;\mathbb{R}) = 0$ also. For example, the special orthogonal group $SO(3,\mathbb{R})$ is a compact, connected Lie group and is diffeomorphic to the 3-dimensional real projective space $\mathbb{R}P^3$. Therefore, its Lie algebra $\mathfrak{so}(3,\mathbb{R})$ satisfies the assumptions of the Kostant-Souriau theorem.

3.6 Poisson manifolds

In this section we shall describe an algebraic foundation of mechanics. A *Poisson algebra* is a triple $(\mathcal{A}, \{,\}, \cdot)$, where the pair $(\mathcal{A}, \{,\})$ is Lie algebra, while at the same time $\mathcal{A}$ is a commutative ring with a unit element and multiplication $\cdot$, such that we have a Leibniz formula

$$\{f, g \cdot h\} = h \cdot \{f, g\} + g \cdot \{f, h\}$$

for every f, g, $h \in \mathcal{A}$. From section 6 of chapter 2 follows that if (M, ω) is a symplectic manifold, then $(C^\infty(M), \{,\}, \cdot)$ is a Poisson algebra, where $\{,\}$ is the Poisson bracket with respect to ω and $\cdot$ is the usual multiplication of functions. A map $\phi : \mathcal{A} \to \mathcal{B}$ of Poisson algebras is called a homomorphism if is a Lie algebra homomorphism and a homomorphism of commutative rings with unit element.

The Leibniz formula says that for every $f \in \mathcal{A}$ the linear map $\mathrm{ad}_f : \mathcal{A} \to \mathcal{A}$ with $\mathrm{ad}_f(g) = \{g, f\}$ is a derivation. It is called the *Hamiltonian derivation* defined by

f. An $f \in \mathcal{A}$ is called a *Casimir element* if $\{f, g\} = 0$ for every $g \in \mathcal{A}$. For example, the unit $1 \in \mathcal{A}$ is a Casimir element, since

$$\{f, 1\} = \{f, 1 \cdot 1\} = 1 \cdot \{f, 1\} + 1 \cdot \{f, 1\} = 2\{f, 1\} = 0$$

for every $f \in \mathcal{A}$. A Poisson algebra $\mathcal{A}$ is called non-degenerate if every Casimir element of $\mathcal{A}$ is of the form $t \cdot 1$, $t \in \mathbb{R}$.

A *Poisson manifold* is a smooth manifold M together with a Poisson structure on the ring of smooth functions $C^\infty(M)$. So the Poisson structure on M is completely determined by the Lie-Poisson bracket $\{,\}$ on $C^\infty(M)$. If $(U, x^1, x^2, ..., x^n)$ is a chart on M, since ad_f is a derivation of $C^\infty(M)$, it is a smooth vector field on M. So,

$$\mathrm{ad}_f|_U = \sum_{k=1}^{n} \{x^k, f\} \frac{\partial}{\partial x^k}.$$

For every f, $g \in C^\infty(M)$ we have

$$\{g, f\}|_U = \sum_{k=1}^{n} \{x^k, f\} \frac{\partial g}{\partial x^k} = -\sum_{k=1}^{n} \{f, x^k\} \frac{\partial g}{\partial x^k} = \sum_{j,k=1}^{n} \{x^k, x^j\} \frac{\partial f}{\partial x^j} \cdot \frac{\partial g}{\partial x^k}.$$

It follows that the Poisson structure on M is determined by a contravariant, skew-symmetric 2-tensor W, which is called the structural tensor of the Poisson structure. For every $p \in M$, the skew-symmetric, bilinear form $W_p : T_p^*M \times T_p^*M \to \mathbb{R}$ is determined by the structural matrix $(\{x^j, x^k\})_{1 \leq j,k \leq n}$. Its rank is called the rank of the Poisson structure at p.

Proposition 6.1. *The Poisson structure of a Poisson manifold M is defined by a symplectic structure on M if and only if the structural matrix is invertible at every point of M.*

Proof. Let (M, ω) be a symplectic manifold and $\{,\}$ be the corresponding Poisson bracket. Then the Poisson tensor is given by $W(df, dg) = \omega(X_f, X_g)$, where X_f and X_g are the Hamiltonian vector fields with Hamiltonian functions f, $g \in C^\infty(M)$, respectively. Let f be such that $W(df, dg) = 0$ for every $g \in C^\infty(M)$. Since T_p^*M is generated by $\{(dg)(p) : g \in C^\infty(M)\}$ for every $p \in M$ and ω is non-degenerate, T_pM is generated by $\{X_g(p) : g \in C^\infty(M)\}$. It follows now that $X_f(p) = 0$ for every $p \in M$. Therefore, $df = 0$ on M. This shows that the structural matrix is invertible. If M is connected, the Poisson structure is also non-degenerate. For the converse, let M be a Poisson manifold, such that the structural matrix W is everywhere invertible. For $f \in C^\infty(M)$ put $X_f = \mathrm{ad}_f$. We define

$$\omega(X_f, X_g) = \{f, g\} = W(df, dg) = df(X_g).$$

Since T_p^*M is generated by $\{(dg)(p) : g \in C^\infty(M)\}$ and W is invertible, it follows that ω is a non-degenerate 2-form and it remains to show that ω is closed. For this, we observe first that

$$[X_f, X_g](h) = X_f(X_g(h)) - X_g(X_f(h)) = X_f(\{h, g\}) - X_g(\{h, f\})$$

$$= \{\{h,g\},f\} - \{\{h,f\},g\} = -\{h,\{f,g\}\} = -X_{\{f,g\}}(h)$$

for every $h \in C^\infty(M)$. Consequently,

$$d\omega(X_f, X_g, X_h) = X_f(\omega(X_g, X_h)) - X_g(\omega(X_f, X_h)) + X_h(\omega(X_f, X_g))$$

$$-\omega([X_f, X_g], X_h) + \omega([X_f, X_h], X_g) - \omega([X_g, X_h], X_f)$$

$$= 2[\{\{f,g\},h\} + \{\{g,h\},f\} + \{\{h,f\},g\}] = 0. \quad \square$$

Example 6.2. Let $(\mathfrak{g}, [,])$ be a (real) Lie algebra of finite dimension n and $\mathfrak{g}^*$ be its dual. Since $\mathfrak{g}$ has finite demension, the double dual $\mathfrak{g}^{**}$ is naturally isomorphic to $\mathfrak{g}$, and so their elements can be identified. For f, $g \in C^\infty(\mathfrak{g}^*)$ let $\{f,g\} \in C^\infty(\mathfrak{g}^*)$ be defined by

$$\{f,g\}(\mu) = \mu[df(\mu), dg(\mu)]$$

for $\mu \in \mathfrak{g}^*$. It is obvious that the bracket $\{,\}$ is bilinear and skew-symmetric. Moreover, the Leibniz formula holds, since it holds for d. In order to have a Poisson manifold, it remains to verify the Jacobi identity. If $\{x_1, x_2, ..., x_n\}$ is a basis of $\mathfrak{g}$, then x_1, x_2,...,x_n can be considered as (global) coordinate functions on $\mathfrak{g}$. If f, $g \in C^\infty(\mathfrak{g}^*)$, then

$$\{f,g\} = \sum_{i,j=1}^{n} \{x_i, x_j\} \frac{\partial f}{\partial x_i} \cdot \frac{\partial g}{\partial x_j}.$$

Since $\{x_i, x_j\}(\mu) = \mu[dx_i(\mu), dx_j(\mu)] = \mu[x_i, x_j]$ for every $\mu \in \mathfrak{g}^*$, it follows from the Jacobi identity on $\mathfrak{g}$, that it also holds for $\{,\}$ on the the set $\{x_1, x_2, ..., x_n\}$. In general, if f, g, $h \in C^\infty(\mathfrak{g}^*)$ note first that

$$\sum_{k=1}^{n} \{x_k, x_i\}\{\frac{\partial f}{\partial x_k}, x_j\} = \sum_{k,l=1}^{n} \{x_k, x_i\}\{x_l, x_j\} \frac{\partial^2 f}{\partial x_l \partial x_k}$$

$$= \sum_{k,l=1}^{n} \{x_l, x_j\}\{x_k, x_i\} \frac{\partial^2 f}{\partial x_k \partial x_l} = \sum_{k=1}^{n} \{x_k, x_j\}\{\frac{\partial f}{\partial x_k}, x_i\}.$$

Now we compute

$$\{\{f,g\},h\} = \sum_{i,j,k=1}^{n} \{\{x_i, x_j\}, x_k\} \frac{\partial f}{\partial x_i} \cdot \frac{\partial g}{\partial x_j} \cdot \frac{\partial h}{\partial x_k}$$

$$+ \sum_{i,j,k=1}^{n} \{x_i, x_j\}\{\frac{\partial f}{\partial x_i}, x_k\} \frac{\partial h}{\partial x_k} \cdot \frac{\partial g}{\partial x_j} + \sum_{i,j,k=1}^{n} \{x_i, x_j\}\{\frac{\partial g}{\partial x_j}, x_k\} \frac{\partial h}{\partial x_k} \cdot \frac{\partial f}{\partial x_i}.$$

Similarly,

$$\{\{g,h\},f\} = \sum_{i,j,k=1}^{n} \{\{x_j, x_k\}, x_i\} \frac{\partial f}{\partial x_i} \cdot \frac{\partial g}{\partial x_j} \cdot \frac{\partial h}{\partial x_k}$$

$$+ \sum_{i,j,k=1}^{n} \{x_j, x_k\}\{\frac{\partial g}{\partial x_j}, x_i\} \frac{\partial h}{\partial x_k} \cdot \frac{\partial f}{\partial x_i} + \sum_{i,j,k=1}^{n} \{x_j, x_k\}\{\frac{\partial h}{\partial x_k}, x_i\} \frac{\partial g}{\partial x_j} \cdot \frac{\partial f}{\partial x_i}$$

and

$$\{\{h,f\},g\} = \sum_{i,j,k=1}^{n} \{\{x_k,x_i\},x_j\}\frac{\partial f}{\partial x_i}\cdot\frac{\partial g}{\partial x_j}\cdot\frac{\partial h}{\partial x_k}$$

$$+\sum_{i,j,k=1}^{n}\{x_k,x_i\}\{\frac{\partial h}{\partial x_k},x_j\}\frac{\partial g}{\partial x_j}\cdot\frac{\partial f}{\partial x_i}+\sum_{i,j,k=1}^{n}\{x_k,x_i\}\{\frac{\partial f}{\partial x_i},x_j\}\frac{\partial h}{\partial x_k}\cdot\frac{\partial g}{\partial x_j}.$$

Summing up we get

$$\{\{f,g\},h\}+\{\{g,h\},f\}+\{\{h,f\},g\}=$$

$$\sum_{i,j,k=1}^{n}\Big(\{\{x_i,x_j\},x_k\}+\{\{x_j,x_k\},x_i\}+\{\{x_k,x_i\},x_j\}\Big)\frac{\partial f}{\partial x_i}\cdot\frac{\partial g}{\partial x_j}\cdot\frac{\partial h}{\partial x_k}=0.$$

In this way $\mathfrak{g}^*$ becomes a Poisson manifold.

If M_1 and M_2 are two Poisson manifolds, a smooth map $h : M_1 \to M_2$ is called *Poisson* if $h^* : C^\infty(M_2) \to C^\infty(M_1)$ is a homomorphism of Poisson algebras.

Let M be a Poisson manifold. For every $f \in C^\infty(M)$, the smooth vector field X_f corresponding to the Hamiltonian derivation $\mathrm{ad}_f = \{.,f\}$ is called the *Hamiltonian vector field* of f. This definition agrees with the definition of section 4 in case M is symplectic.

Proposition 6.3. *Let M be a Poisson manifold and X_f be a Hamiltonian vector field on M with Hamiltonian function $f \in C^\infty(M)$. Let $\phi : D \to M$ be the flow of X_f, where $D \subset \mathbb{R}\times M$ is an open neighbourhood of $\{0\}\times M$.*

(i) If $g \in C^\infty(M)$, then

$$\frac{d}{dt}(g\circ\phi_t) = \{g,f\}\circ\phi_t = \{g\circ\phi_t, f\}.$$

(ii) $f\circ\phi_t = f$

(iii) The flow of the Hamiltonian vector field X_f consists of Poisson maps.

Proof. (i) If $p \in M$, then on the one hand

$$\frac{d}{dt}(g\circ\phi_t) = X_f(g)(\phi_t(p)) = \{g,f\}(\phi_t(p))$$

and on the other hand

$$\frac{d}{dt}(g\circ\phi_t) = g_{*\phi_t(p)}((\phi_t)_{*p}(X_f(p))) = (g\circ\phi_t)_{*p}(X_f(p)) = X_f(g\circ\phi_t)(p) = \{g\circ\phi_t, f\}.$$

(ii) This is obvious from (i) taking $g = f$.

(iii) Let $g_1,\, g_2 \in C^\infty(M)$ and let $g \in C^\infty(D)$ be defined by

$$g(t,p) = \{g_1\circ\phi_t, g_2\circ\phi_t\}(p) - \{g_1,g_2\}(\phi_t(p)).$$

From (i) and the Jacobi identity we have

$$\frac{\partial g}{\partial t} = \{\frac{d}{dt}(g_1\circ\phi_t), g_2\circ\phi_t\}(p) + \{g_1\circ\phi_t, \frac{d}{dt}(g_2\circ\phi_t)\}(p) - \frac{d}{dt}\{g_1,g_2\}(\phi_t(p)) =$$

$$\{\{g_1 \circ \phi_t, f\}, g_2 \circ \phi_t\} + \{g_1 \circ \phi_t, \{g_2 \circ \phi_t, f\}\} - \{\{g_1, g_2\} \circ \phi_t, f\} =$$
$$\{g_t, f\} = X_f(g_t),$$

where as usual $g_t = g(t, .)$, and $g(0,p) = 0$. By uniqueness of solutions of ordinary differential equations, we must necessarily have $g(t,p) = 0$ for all t such that $(t,p) \in D$. □

If $h : M_1 \to M_2$ is a Poisson map of Poisson manifolds and $f \in C^\infty(M_2)$, then $h_{*p}(X_{h^*(f)}(p)) = X_f(h(p))$ for every $p \in M_1$. Therefore, h transforms integral curves of $X_{h^*(f)}$ in M_1 to integral curves of X_f on M_2.

If M is a Poisson manifold and $N \subset M$ is an immersed submanifold, then N is called a *Poisson submanifold* if the inclusion $i : N \hookrightarrow M$ is a Poisson map. On every Poisson manifold M one can define an equivalence relation $\sim$ by setting $p \sim q$ if and only if there is a piecewise smooth curve from p to q whose smooth parts are pieces of integral curves of Hamiltonian vector fields of M. The equivalence classes are called the *symplectic leaves* of the Poisson structure of M. We shall prove that the symplectic leaves are immersed submanifolds and carry a unique symplectic structure so that the become Poisson submanifolds of M.

Let $p \in M$ and f_1, f_2,...,$f_k \in C^\infty(M)$ be such that the set $\{X_{f_1}(p), ..., X_{f_k}(p)\}$ is a basis of $\mathrm{Im}\tilde{W}_p$, where W_p is the Poisson tensor and $\tilde{W}_p : T_p^*M \to T_p^{**}M \cong T_pM$ is the induced linear map. In other words, $X_{f_j}(p) = \tilde{W}_p(df_j(p))$. There exists some $\epsilon > 0$ and an open neighbourhood U of p such that the flow ϕ_j of X_{f_j} is defined on $(-\epsilon, \epsilon) \times U$ for every $1 \leq j \leq k$. Taking a smaller $\epsilon > 0$, we may assume that

$$\Phi_p(t_1, t_2, ..., t_k) = (\phi_{1,t_1} \circ \phi_{2,t_2} \circ ... \circ \phi_{k,t_k})(p)$$

is defined for $|t_j| < \epsilon$, $1 \leq j \leq k$. Obviously, Φ_p is smooth and

$$(\Phi_p)_{*0}(\frac{\partial}{\partial t_j}) = X_{f_j}(p)$$

for $1 \leq j \leq k$. So, $(\Phi_p)_{*0}$ is a monomorphism and from the inverse function theorem there exists an open neighbourhood V_p of 0 in $\mathbb{R}^k$ such that $\Phi_p : V_p \to M$ is an embedding. Note also that $\mathrm{Im}(\Phi_p)_{*0} = \mathrm{Im}\tilde{W}_p$.

Lemma 6.4. *There exists an open neighbourhood V_p of 0 in $\mathbb{R}^k$ such that $\mathrm{Im}(\Phi_p)_{*t} = \mathrm{Im}\tilde{W}_{\Phi_p(t)}$ for every $t = (t_1, t_2, ..., t_k) \in V_p$.*

Proof. We have

$$(\Phi_p)_{*t}(\frac{\partial}{\partial t_j}) = \big((\phi_{1,t_1})_* \circ ... \circ (\phi_{j-1,t_{j-1}})_* \circ X_{f_j} \circ \phi_{j+1,t_{j+1}} \circ ... \phi_{k,t_k}\big)(p)$$

$$= X_{h_j}(\Phi_p(t)) \in \mathrm{Im}\tilde{W}_{\Phi_p(t)},$$

where $h_j = f_j \circ (\phi_{1,t_1} \circ ... \circ \phi_{j-1,t_{j-1}})^{-1}$. Therefore, $\mathrm{Im}(\Phi_p)_{*t} \leq \mathrm{Im}\tilde{W}_{\Phi_p(t)}$. However, $\dim \mathrm{Im}(\Phi_p)_{*t} = \dim \mathrm{Im}(\Phi_p)_{*0} = \dim \mathrm{Im}\tilde{W}_p = \dim \mathrm{Im}\tilde{W}_{\Phi_p(t)}$, since the flows of Hamiltonian vector fields consist of Poisson maps, for $t \in V_p$ such that $\Phi_p : V_p \to M$

is an embedding. □

If $q \in \Phi_p(V_p)$ and $\Phi_q : V_q \to M$ is an embedding constructed as Φ_p from functions g_1, g_2,...,$g_k \in C^\infty$, then there is an open neighbourhood V_0 of 0 in $\mathbb{R}^k$ such that Φ_q maps V_0 diffeomorphically onto an open subset of $\Phi_p(V_p)$, from the inverse function theorem.

Theorem 6.5. (Symplectic Stratification) *In a Poisson manifold M every symplectic leaf $S \subset M$ is an immersed submanifold and $T_pS = \mathrm{Im}\tilde{W}_p$ for every $p \in S$. Moreover, S has a unique symplectic structure such that S is a Poisson submanifold of M.*

Proof. Using the above notations, the family of all pairs $(\Phi_p(V_p), \Phi_p^{-1})$, $p \in S$, constructed from functions f_1, f_2,...,$f_k \in C^\infty(M)$ such that $\{X_{f_1}(p), X_{f_2}(p), ..., X_{f_k}(p)\}$ is a basis of $\mathrm{Im}\tilde{W}_p$, is a smooth atlas for S. Indeed, let p, $q \in S$ and $y \in \Phi_p(V_p) \cap \Phi_q(V_q)$. From the last remark, shrinking V_y we may assume that $\Phi_y(V_y) \subset \Phi_p(V_p) \cap \Phi_q(V_q)$ and Φ_y is an embedding of V_y similtaneously into $\Phi_p(V_p)$ and $\Phi_q(V_q)$. Therefore, S is an immersed Poisson submanifold of M and $T_pS = \mathrm{Im}\tilde{W}_p$, from Lemma 6.4. By Proposition 6.1, it remains to show that the structural matrix of S is invertible at every point $p \in S$. Let $f \in C^\infty(M)$ be such that $\{f, g\}(p) = 0$ for every $g \in C^\infty(M)$. Then $df(p)(X_g(p)) = X_g(f)(p) = 0$ for every $g \in C^\infty(M)$, which implies that $d(f|_S)(p) = df(p)|_{T_pS} = 0$. This shows that the structural matrix of S at p is invertible. □

Example 6.6. Let $(\mathfrak{g}, [,])$ be the Lie algebra of a Lie group G and $\mathfrak{g}^*$ be its dual. If $f \in C^\infty(\mathfrak{g}^*)$, the Hamiltonian vector field X_f with respect to the Poisson structure on $\mathfrak{g}^*$ defined in Example 6.2 satisfies

$$X_f(\mu)(g) = \{g, f\}(\mu) = \mu([dg(\mu), df(\mu)]) = -(\mu \circ \mathrm{ad}_{df(\mu)})(dg(\mu))$$

for every $g \in C^\infty(\mathfrak{g}^*)$ and $\mu \in \mathfrak{g}^*$, where we have identified $\mathfrak{g}^{**}$ with $\mathfrak{g}$. Thus, $X_f(\mu) = -(\mathrm{ad}_{df(\mu)})^*$ for every $\mu \in \mathfrak{g}^*$ and X_f is precisely a fundamental vector field of the coadjoint representation of G. It follows that the symplectic leaves in $\mathfrak{g}^*$ are the coadjoint orbits. Moreover, the restricted Poisson structure on each coadjoint orbit coincides with the Kirillov symplectic structure.

Chapter 4

Symmetries and integrability

4.1 Symplectic group actions

Let M be a smooth manifold, G a Lie group with Lie algebra $\mathfrak{g}$ and $\phi : G \times M \to M$ be a smooth group action. If $X \in \mathfrak{g}$, the fundamental vector field $\phi_*(X) \in \mathcal{X}(M)$ of the action which corresponds to X is the infinitesimal generator of the flow $\phi_X : \mathbb{R} \times M \to M$ defined by $\phi_X(t,p) = \phi(\exp(tX), p)$. Note that for $g \in G$ the transformed vector field $(\phi_g)_*(\phi_*(X))$ is the fundamental vector field $\phi_*(\mathrm{Ad}_g(X))$, that is

$$(\phi_g)_{*p}(\phi_*(X)(p)) = \phi_*(\mathrm{Ad}_g(X))(\phi_g(p))$$

for every $p \in M$. Indeed,

$$\phi_*(\mathrm{Ad}_g(X))(\phi_g(p)) = \frac{d}{dt}\Big|_{t=0} \phi^{\phi_g(p)}(\exp(t\mathrm{Ad}_g(X))) =$$

$$(\phi^{\phi_g(p)})_{*e}(\frac{d}{dt}\Big|_{t=0} \exp(t\mathrm{Ad}_g(X))) = (\phi^{\phi_g(p)})_{*e}(\mathrm{Ad}_g(X)) =$$

$$\frac{d}{dt}\Big|_{t=0} \phi(g\exp(tX)g^{-1}, \phi(g,p)) = \frac{d}{dt}\Big|_{t=0} \phi(g\exp(tX), p) =$$

$$\frac{d}{dt}\Big|_{t=0} (\phi^p \circ L_g)(\exp(tX)) = \frac{d}{dt}\Big|_{t=0} (\phi_g \circ \phi^p)(\exp(tX)) =$$

$$(\phi_g)_{*p}((\phi^p)_{*e}(X)) = (\phi_g)_{*p}(\phi_*(X)(p)).$$

Lemma 1.1. *The linear map $\phi_* : \mathfrak{g} \to \mathcal{X}(M)$ is an anti-homomorphism of Lie algebras, meaning that $\phi_*([X,Y]) = -[\phi_*(X), \phi_*(Y)]$ for every X, $Y \in \mathfrak{g}$.*

Proof. If $p \in M$, then we compute

$$[\phi_*(X), \phi_*(Y)](p) = \frac{d}{dt}\Big|_{t=0} (\phi_{\exp(-tX)})_{*\phi_{\exp(tX)}(p)}(\phi_*(Y))(\phi_{\exp(tX)}(p))) =$$

$$\frac{d}{dt}\Big|_{t=0} \phi_*(\mathrm{Ad}_{\exp(-tX)}(Y))(p) = \phi_*(-\mathrm{ad}_X(Y))(p) = -\phi_*([X,Y]). \quad \square$$

Although ϕ_* is an anti-homomorphism of Lie algebras, it follows that $\phi_*(\mathfrak{g})$ is a Lie subalgebra of $\mathcal{X}(M)$ of finite dimension.

Definition 1.2. Let (M, ω) be a symplectic manifold and G a Lie group. A smooth group action $\phi : G \times M \to M$ is called *symplectic* if $\phi_g = \phi(g, .) : M \to M$ is a symplectomorphism for every $g \in G$.

If ϕ is symplectic, then $\phi_*(\mathfrak{g}) \subset \mathfrak{sp}(M, \omega)$, and therefore

$$\phi_*([\mathfrak{g}, \mathfrak{g}]) \subset [\mathfrak{sp}(M, \omega), \mathfrak{sp}(M, \omega)] \subset \mathfrak{h}(M, \omega),$$

by Proposition 6.8 in chapter 2. If $H_\phi : \mathfrak{g} \to H^1_{DR}(M)$ is the linear map defined by $H_\phi(X) = [i_{\phi_*(X)}\omega]$, then $X \in \text{Ker } H_\phi$ if and only if $\phi_*(X)$ is a Hamiltonian vector field, and $[\mathfrak{g}, \mathfrak{g}] \subset \text{Ker } H_\phi$.

Definition 1.3. A symplectic group action ϕ is called Hamiltonian if $H_\phi = 0$.

Thus, if $H^1(M; \mathbb{R}) = \{0\}$, then every symplectic group action on M is Hamiltonian. In particular, every symplectic group action on a simply connected symplectic manifold is Hamiltonian. Also if the Lie algebra $\mathfrak{g}$ of G is perfect, meaning that $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$, then every symplectic group action of G is Hamiltonian. This happens for example in the case $G = SO(3, \mathbb{R})$, because $\mathfrak{so}(3, \mathbb{R})$ is isomorphic to the Lie algebra $(\mathbb{R}, \times)$, which is obviously perfect.

If ϕ is a Hamiltonian group action, in general there is no canonical way to choose a Hamiltonian function for $\phi_*(X)$, since adding a constant to a Hamiltonian function yields a new Hamiltonian function. If there is a linear map $\rho : \mathfrak{g} \to C^\infty(M)$ such that $\rho(X)$ is a Hamiltonian function for $\phi_*(X)$ for every $X \in \mathfrak{g}$, there is a smooth map $\mu : M \to \mathfrak{g}^*$ defined by $\mu(p)(X) = \rho(X)(p)$.

Examples 1.4. (a) Let M be a smooth manifold, G a Lie group with Lie algebra $\mathfrak{g}$ and $\phi : G \times M \to M$ a smooth group action. Then, ϕ is covered by a group action $\tilde{\phi}$ of G on T^*M defined by $\tilde{\phi}(g, a) = a \circ (\phi_{g^{-1}})_{*\phi_g(\pi(a))}$, where $\pi : T^*M \to M$ is the cotangent bundle projection. Since $\pi \circ \tilde{\phi}_g = \phi_g \circ \pi$, differentiating we get

$$\pi_{*\tilde{\phi}_g(a)} \circ (\tilde{\phi}_g)_{*a} = (\phi_g)_{*\pi(a)} \circ \pi_{*a}$$

for every $a \in T^*M$ and $g \in G$. The Liouville 1-form θ on T^*M remains invariant under the action of G, because

$$((\tilde{\phi}_g)^*\theta)_a = \theta_{\tilde{\phi}_g(a)} \circ (\tilde{\phi}_g)_{*a} = a \circ (\phi_g^{-1})_{*\phi_g(\pi(a))} \circ (\phi_g)_{*\pi(a)} \circ \pi_{*a} = a \circ \pi_{*a} = \theta_a.$$

Consequently, the action of G on T^*M is symplectic with respect to the canonical symplectic structure $\omega = -d\theta$. Moreover, it is Hamiltonian, because

$$0 = L_{\tilde{\phi}_*(X)}\theta = i_{\tilde{\phi}_*(X)}(d\theta) + d(i_{\tilde{\phi}_*(X)}\theta)$$

and therefore $i_{\tilde{\phi}_*(X)}\omega = d(i_{\tilde{\phi}_*(X)}\theta)$. Here we have a linear map $\rho : \mathfrak{g} \to C^\infty(T^*M)$ defined by $\rho(X) = i_{\tilde{\phi}_*(X)}\theta$ and $\mu : T^*M \to \mathfrak{g}^*$ is given by the formula

$$\mu(a)(X) = \theta_a(\tilde{\phi}_*(X)).$$

(b) Let G be a Lie group with Lie algebra $\mathfrak{g}$ and $\mathcal{O}$ be a coadjoint orbit. The symplectic Kirillov 2-form ω^- is Ad^*-invariant, by Lemma 4.6 of chapter 3, and so the natural action of G on $\mathcal{O}$ is symplectic. Recall that

$$\omega^-_\nu(X_{\mathfrak{g}^*}(\nu), Y_{\mathfrak{g}^*}(\nu)) = -\nu([X,Y]) = (\nu \circ \mathrm{ad}_Y)(X) = -Y_{\mathfrak{g}^*}(\nu)(X) = -X(Y_{\mathfrak{g}^*}(\nu))$$

for every X, $Y \in \mathfrak{g}$ and $\nu \in \mathcal{O}$, having identified $\mathfrak{g}^{**}$ with $\mathfrak{g}$. If now $\rho_X \in C^\infty(\mathfrak{g}^*)$ is the (linear) function defined by $\rho_X(\nu) = -\nu(X)$, then $d\rho_X(\nu) = -X$ (again we identify $\mathfrak{g}^{**}$ with $\mathfrak{g}$). It follows that $i_{X_{\mathfrak{g}^*}}\omega^- = d\rho_X$, which shows that the action of G on $\mathcal{O}$ is Hamiltonian.

Let $\phi : G \times M \to M$ be a Hamiltonian group action of the Lie group G with Lie algebra $\mathfrak{g}$ on a connected, symplectic manifold (M, ω). We assume that we have a linear lift $\rho : \mathfrak{g} \to C^\infty(M)$ such that $\phi_*(X) = X_{\rho(X)}$ for every $X \in \mathfrak{g}$. We shall study the possibility to change ρ to a new lift which is also a Lie algebra homomorphism. From Proposition 6.8 of chapter 2 and Lemma 1.1 we have

$$X_{\{\rho(X_0),\rho(X_1)\}} = -[X_{\rho(X_0)}, X_{\rho(X_1)}] = \phi_*([X_0, X_1]) = X_{\rho([X_0,X_1])},$$

for every X_0, $X_1 \in \mathfrak{g}$. Since M is connected, there exists $c(X_0, X_1) \in \mathbb{R}$ such that

$$\{\rho(X_0), \rho(X_1)\} = \rho([X_0, X_1]) + c(X_0, X_1).$$

Obviously, $c : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ is a skew-symmetric, bilinear form. Moreover, $\delta c = 0$, from the Jacobi identity and the linearity of ρ. Hence $c \in Z^2(\mathfrak{g})$. If $\tilde{\rho} : \mathfrak{g} \to C^\infty(M)$ is another linear lift and $\sigma = \tilde{\rho} - \rho$, then $\sigma \in \mathfrak{g}^*$ and

$$\{\tilde{\rho}(X_0), \tilde{\rho}(X_1)\} = \{\rho(X_0), \rho(X_1)\} = \rho([X_0, X_1]) + c(X_0, X_1) =$$

$$\tilde{\rho}([X_0, X_1]) + c(X_0, X_1) - \sigma([X_0, X_1]).$$

Hence, $\tilde{c}(X_0, X_1) - c(X_0, X_1) = -\sigma([X_0, X_1]) = (\delta\sigma)(X_0, X_1)$. We conclude that there is a choice of $\tilde{\rho}$ such that $\tilde{c} = 0$ if and only if $[c] = 0$ in $H^2(\mathfrak{g})$. Thus, in case $H^2(\mathfrak{g}) = \{0\}$, we can always select a linear lift $\rho : \mathfrak{g} \to C^\infty(M)$ which is a Lie algebra homomorphism.

Examples 1.5. (a) Let M be a smooth manifold, G a Lie group with Lie algebra $\mathfrak{g}$ and $\phi : G \times M \to M$ a smooth group action. As we saw in Example 1.4(a), the covering action $\tilde{\phi}$ on T^*M is Hamiltonian and $\rho : \mathfrak{g} \to C^\infty(T^*M)$ is given by the formula $\rho(X) = i_{\tilde{\phi}_*(X)}\theta$, where θ is the invariant Liouville 1-form. Then,

$$c(X_0, X_1) = -d\theta(\tilde{\phi}_*(X_0), \tilde{\phi}_*(X_1)) - \theta(\tilde{\phi}_*([X_0, X_1]) =$$

$$-L_{\tilde{\phi}_*(X_0)}\rho(X_1) + L_{\tilde{\phi}_*(X_1)}\rho(X_0) + \theta([\tilde{\phi}_*(X_0), \tilde{\phi}_*(X_1)]) - \theta(\tilde{\phi}_*([X_0, X_1]) =$$

$$-\{\rho(X_1), \rho(X_0)\} + \{\rho(X_0), \rho(X_1)\} - 2\theta(\tilde{\phi}_*([X_0, X_1]) = 2c(X_0, X_1)$$

and hence $c = 0$.

(b) If G is a Lie group with Lie algebra $\mathfrak{g}$ and $\mathcal{O}$ is a coadjoint orbit, then $\rho(X)(\nu) = -\nu(X)$ for every $X \in \mathfrak{g}$ and $\nu \in \mathcal{O} \subset \mathfrak{g}^*$, as we saw in Example 1.4(b). Therefore, $c = 0$, from the definition of the Kirillov 2-form.

(c) We shall now describe a simple example, where $[c]$ is a non-zero element of $H^2(\mathfrak{g})$. Let $G = (\mathbb{R}^2, +)$, in which case $\mathfrak{g} = \mathbb{R}^2$ with trivial Lie bracket. Let $M = \mathbb{R}^2$ endowed with the euclidean area 2-form $dx \wedge dy$. Let G act on M by translations. The action is symplectic and if $X = (a, b) \in \mathfrak{g}$, then

$$\phi_*(X) = a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y},$$

which is Hamiltonian with Hamiltonian function $\rho(X)(x, y) = ay - bx$. Then,

$$c((a_0, b_0), (a_1, b_1)) = a_0 b_1 - a_1 b_0$$

and therefore $[c] = c \neq 0$.

Definition 1.6. Let M be a symplectic manifold and G be a Lie group with Lie algebra $\mathfrak{g}$. A Hamiltonian group action $\phi : G \times M \to M$ is called *Poisson* (or *strongly Hamiltonian*) if there is a lift $\rho : \mathfrak{g} \to C^\infty(M)$ which is a Lie algebra homomorphism.

We conclude this section with a couple of criteria giving sufficient conditions for a symplectic group action to be Poisson.

Theorem 1.7. *Let (M, ω) be a compact, connected, symplectic $2n$-manifold and G a Lie group with Lie algebra $\mathfrak{g}$. Then, every Hamiltonian group action $\phi : G \times M \to M$ is Poisson.*

Proof. Recall from Proposition 6.9 of Chapter 2 that $C^\infty(M) = \mathbb{R} \oplus C_0^\infty(M, \omega)$. If $X \in \mathfrak{g}$ and $F \in C^\infty(M)$ is a Hamiltonian function of $\phi_*(X)$, we define

$$\rho(X) = F - \frac{1}{\mathrm{vol}(M)} \int_M F\omega^n,$$

where $\omega^n = \omega \wedge \omega \wedge ... \wedge \omega$ n-times. Then ρ is a linear lift. Let X_0, $X_1 \in \mathfrak{g}$ and F_0, $F_1 \in C^\infty(M)$ be Hamiltonian functions of $\phi_*(X_0)$ and $\phi_*(X_1)$, respectively. From Proposition 6.8 of Chapter 2 and Lemma 1.1 we have

$$X_{\{F_0,F_1\}} = -[X_{F_0}, X_{F_1}] = -[\phi_*(X_0), \phi_*(X_1)] = \phi_*([X_0, X_1]),$$

and therefore

$$\rho([X_0, X_1]) = \{F_0, F_1\} - \frac{1}{\mathrm{vol}(M)} \int_M \{F_0, F_1\}\omega^n =$$

$$\{\rho(X_0), \rho(X_1)\} - 0 = \{\rho(X_0), \rho(X_1)\},$$

from Proposition 6.9(b) of Chapter 2. □

Theorem 1.8. *Let G be a Lie group with Lie algebra $\mathfrak{g}$. If $H^1(\mathfrak{g}) = \{0\}$ and $H^2(\mathfrak{g}) = \{0\}$, then every symplectic group action of G is Poisson.*

Proof. Let (M, ω) be a symplectic manifold and $\phi : G \times M \to M$ be a symplectic action of G. The condition $H^1(\mathfrak{g}) = \{0\}$ is equivalent to $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$. This implies that $\phi_*(\mathfrak{g}) \subset \mathfrak{h}(M, \omega)$, which means that the action is Hamiltonian. Since $H^2(\mathfrak{g}) = \{0\}$, the discussion preceeding the Examples 1.5 shows that the action is Poisson. □

We shall end this section with a final remark concerning the existence of invariant almost complex structures. Let (M, ω) be a symplectic manifold and G be a Lie group acting smoothly and symplectically on M. If G is compact (or more generally the action is proper), there exists a G-invariant Riemannian metric on M. Starting with such a Riemannian metric, one can repeat the second part of the proof of Proposition 3.7 of Chapter 1 to construct a G-invariant almost complex structure J on M which is compatible with ω. Obviously, the corresponding compatible Riemannian metric g on M given by the formula $g_x(u, v) = -\omega(J(u), v)$, for u, $v \in T_xM$, $x \in M$, is also G-invariant.

4.2 Momentum maps

Let (M, ω) be a connected, symplectic manifold, G be a Lie group with Lie algebra $\mathfrak{g}$ and $\phi : G \times M \to M$ be a Poisson action.

Definition 2.1. A *momentum map* for ϕ is a smooth map $\mu : M \to \mathfrak{g}^*$ such that $\rho : \mathfrak{g} \to C^\infty(M)$ defined by $\rho(X)(p) = \mu(p)(X)$ for $X \in \mathfrak{g}$ and $p \in M$ satisfies
(i) $\phi_*(X) = X_{\rho(X)}$, and
(ii) $\{\rho(X), \rho(Y)\} = \rho([X, Y])$ for every X, $Y \in \mathfrak{g}$.

From the point of view of dynamical systems, one reason to study momentum maps is the following. If $H : M \to \mathbb{R}$ is a G-invariant, smooth function, then μ is constant along the integral curves of the Hamiltonian vector field X_H. Indeed, for every $X \in \mathfrak{g}$ we have

$$L_{X_H}\rho(X) = \{\rho(X), H\} = -\{H, \rho(X)\} = -L_{\phi_*(X)}H = 0.$$

Theorem 2.2. *If G is a connected Lie group, then a momentum map $\mu : M \to \mathfrak{g}^*$ is G-equivariant with respect to the coadjoint action on $\mathfrak{g}^*$.*

Proof. The momentum map μ is G-equivariant when $\mu(\phi_g(p)) = \mu(p) \circ \mathrm{Ad}_{g^{-1}}$ or equivalently

$$\rho(X)(\phi_g(p)) = \rho(\mathrm{Ad}_{g^{-1}}(X))(p)$$

for every $X \in \mathfrak{g}$, $g \in G$ and $p \in M$. Observe that if this is true for two elements g_1, $g_2 \in G$ and for every $X \in \mathfrak{g}$ and $p \in M$, then this is also true for the element g_1g_2. Recall that since G is connected, if V is any connected, open neighbourhood of the identity $e \in G$ with $V = V^{-1}$, then $G = \bigcup_{n=1}^{\infty} V^n$, where $V^n = V \cdot \ldots \cdot V$, n-times.

It follows that it suffices to prove the above equality for $g = \exp tY$ for $Y \in \mathfrak{g}$ and $t \in \mathbb{R}$. In other words, it suffices to show that

$$\rho(X)(\phi_{\exp(tY)}(p)) = \rho(\mathrm{Ad}_{\exp(-tY)}(X))(p)$$

for every X, $Y \in \mathfrak{g}$, $p \in M$ and $t \in \mathbb{R}$. As this is true for $t = 0$, we need only show that the two sides have equal derivatives with respect to t. The derivative of the left hand side is

$$\frac{d}{dt}\rho(X)(\phi_{\exp(tY)}(p)) = d\rho(X)(\phi_{\exp(tY)}(p))(\frac{d}{dt}\phi_{\exp(tY)}(p)) =$$

$$\omega(\phi_*(X)(\phi_{\exp(tY)}(p)), \phi_*(Y)(\phi_{\exp(tY)}(p))) =$$

$$\omega((\phi_g)_{*p}(\phi_*(\mathrm{Ad}_{\exp(-tY)}(X))(p)), (\phi_g)_{*p}(\phi_*(\mathrm{Ad}_{\exp(-tY)}(Y))(p))) =$$

$$\omega(\phi_*(\mathrm{Ad}_{\exp(-tY)}(X))(p), \phi_*(\mathrm{Ad}_{\exp(-tY)}(Y))(p)) =$$

$$\omega(\phi_*(\mathrm{Ad}_{\exp(-tY)}(X))(p), \phi_*(Y)(p)),$$

since $\mathrm{Ad}_{\exp(-tY)}(Y) = Y$, the action is symplectic and using the remarks in the beginning of section 1. The derivative of the right hand side is

$$\frac{d}{dt}\rho(\mathrm{Ad}_{\exp(-tY)}(X))(p) = \rho(\frac{d}{dt}\mathrm{Ad}_{\exp(-tY)}(X))(p) =$$

$$\rho(\mathrm{ad}_{(-Y)}(\mathrm{Ad}_{\exp(-tY)}(X)))(p) = \rho([-Y, \mathrm{Ad}_{\exp(-tY)}(X)])(p) =$$

$$\{\rho(\mathrm{Ad}_{\exp(-tY)}(X)), \rho(Y)\}(p) = \omega(\phi_*(\mathrm{Ad}_{\exp(-tY)}(X))(p), \phi_*(Y)(p)). \quad \square$$

In general, for every $X \in \mathfrak{g}$ and $g \in G$ the smooth function

$$(\phi_g)^*(\rho(X)) - \rho(\mathrm{Ad}_{g^{-1}}(X)) : M \to \mathbb{R}$$

has differential

$$d((\phi_g)^*(\rho(X)) - \rho(\mathrm{Ad}_{g^{-1}}(X))) = (\phi_g)^*(d\rho(X)) - d\rho(\mathrm{Ad}_{g^{-1}}(X)) =$$

$$(\phi_g)^*(\tilde{\omega}^{-1}(\phi_*(X))) - \tilde{\omega}^{-1}(\phi_*(\mathrm{Ad}_{g^{-1}}(X))) = \tilde{\omega}^{-1}((\phi_{g^{-1}})_*\phi_*(X) - \phi_*(\mathrm{Ad}_{g^{-1}}(X))) = 0,$$

because the group action is symplectic and using the remarks in the beginning of section 1. Since M is connected, it is constant and so we have a function $c : G \to \mathfrak{g}^*$ defined by

$$c(g) = (\phi_g)^*(\rho(X)) - \rho(\mathrm{Ad}_{g^{-1}}(X)) = \mu(\phi_g(p)) - \mathrm{Ad}^*_g(\mu(p))$$

for any $p \in M$. If now g_0, $g_1 \in G$, then

$$c(g_0 g_1) = \mu(\phi_{g_0}(\phi_{g_1}(p))) - \mathrm{Ad}^*_{g_0}(\mathrm{Ad}^*_{g_1}(\mu(p))) =$$

$$\mu(\phi_{g_0}(\phi_{g_1}(p))) - \mathrm{Ad}^*_{g_0}(\mu(\phi_{g_1}(p))) + \mathrm{Ad}^*_{g_0}(\mu(\phi_{g_1}(p))) - \mathrm{Ad}^*_{g_0}(\mathrm{Ad}^*_{g_1}(\mu(p))) =$$

$$c(g_0) + \mathrm{Ad}^*_{g_0}(\mu(\phi_{g_1}(p)) - \mathrm{Ad}^*_{g_1}(\mu(p))) = c(g_0) + \mathrm{Ad}^*_{g_0}(c(g_1)).$$

This means that c is a 1-cocycle with respect to the group cohomology of G^δ with coefficients in the G-module $\mathfrak{g}^*$, with respect to the coadjoint action, where G^δ denotes G made discrete. If μ' is another momentum map, there exists a constant $a \in \mathfrak{g}^*$ such that $\mu' = \mu + a$. The corresponding cocycle c' is given by the formula

$$c'(g) = \mu(\phi_g(p)) + a - \mathrm{Ad}_g^*(\mu(p)) - \mathrm{Ad}_g^*(a) = (c - \delta a)(g)$$

where δ denotes the coboundary operator in group cohomology. Thus, the cohomology class $[c] \in H^1(G^\delta; \mathfrak{g}^*)$ does not depend on the choice of the momentum map but only on the group action.

Proposition 2.3. *If $H^1(G^\delta; \mathfrak{g}^*) = \{0\}$, there exists a G-equivariant momentum map.*

Proof. Let μ be any momentum map with corresponding 1-cocycle c. There exists $a \in \mathfrak{g}^*$ such that $c = \delta a$, that is $c(g) = \mathrm{Ad}_g^*(a) - a$ for every $g \in G$. Then $\mu + a$ is a G-equivariant momentum map, because

$$\mu(\phi_g(p)) + a = c(g) + \mathrm{Ad}_g^*(\mu(p)) + a = \mathrm{Ad}_g^*(a) - a + \mathrm{Ad}_g^*(\mu(p)) + a = \mathrm{Ad}_g^*(\mu + a)(p)$$

for every $g \in G$ and $p \in M$. □

Examples 2.4. (a) Let $\phi : G \times M \to M$ be a smooth action of the Lie group G with Lie algebra $\mathfrak{g}$ on the smooth manifold M and $\tilde{\phi} : G \times T^*M \to T^*M$ be the lifted action on the cotangent bundle. As we saw in Examples 1.4(a) and 1.5(a), the action of G on T^*M is Poisson and actually the Liouville 1-form θ on T^*M is G-invariant. The momentum map $\mu : T^*M \to \mathfrak{g}^*$ is given by the formula

$$\mu(a)(X) = \theta_a(\tilde{\phi}_*(X)(a))$$

for $X \in \mathfrak{g}$ and $a \in T^*M$, and is G-equivariant, because θ is G-invariant. Indeed,

$$\mu(\tilde{\phi}_g(a))(X) = \theta_{\tilde{\phi}(a)}(\tilde{\phi}_*(X)(\tilde{\phi}_g(a))) = ((\tilde{\phi}_{g^{-1}})^*\theta)_{\tilde{\phi}(a)}(\tilde{\phi}_*(X)(\tilde{\phi}_g(a))) =$$

$$\theta_a((\tilde{\phi}_{g^{-1}})_{*\tilde{\phi}_g(a)}(\tilde{\phi}_*(X)(\tilde{\phi}_g(a)))) = \theta_a(\tilde{\phi}_*(\mathrm{Ad}_{g^{-1}}(X))(a)) = \mu(a)(\mathrm{Ad}_{g^{-1}}(X)),$$

for every $g \in G$.

In the case of the 3-dimensional euclidean space $\mathbb{R}^3$ we have $T^*\mathbb{R}^3 \cong \mathbb{R}^3 \times \mathbb{R}^3$, where the isomorphism is defined by the euclidean inner product $\langle , \rangle$, identifying thus $T^*\mathbb{R}^3$ with $T\mathbb{R}^3$. The Liouville 1-form is given by the formula

$$\theta_{(q,p)}(v, w) = \langle v, p \rangle.$$

The natural action of $SO(3, \mathbb{R})$ on $\mathbb{R}^3$ is covered by the action $\tilde{\phi}$ such that

$$\tilde{\phi}_A(q, p)(v) = \langle p, A^{-1}v \rangle = \langle Ap, v \rangle$$

for every $v \in T_q\mathbb{R}^3$ and $A \in SO(3, \mathbb{R})$. Therefore, $\tilde{\phi}_A(q, p) = (Aq, Ap)$ for every $(q, p) \in T^*\mathbb{R}^3$ and $A \in SO(3, \mathbb{R})$. If now $v \in \mathbb{R}^3 \cong \mathfrak{so}(3, \mathbb{R})$, the corresponding fundamental vector field of the action satisfies

$$\tilde{\phi}_*(v)(q, p) = (\hat{v}q, \hat{v}p) = (v \times q, v \times p).$$

It follows that the momentum map satisfies

$$\mu(q,p)(v) = \langle v \times q, p\rangle = \langle q \times p, v\rangle$$

for every $v \in \mathbb{R}^3$. Consequently, the momentum map is the angular momentum

$$\mu(q,p) = q \times p.$$

Suppose now that we have a system of n particles in $\mathbb{R}^3$. The configuration space is $\mathbb{R}^{3n}$. The additive group $\mathbb{R}^3$ acts on $\mathbb{R}^{3n}$ by translations, that is

$$\phi_x(q^1, q^2, ..., q^n) = (q^1 + x, q^2 + x, ..., q^n + x)$$

for every $x \in \mathbb{R}^3$. The lifted action on $T^*\mathbb{R}^{3n} \cong \mathbb{R}^{3n} \times \mathbb{R}^{3n}$ is

$$\tilde{\phi}_x(q^1, q^2, ..., q^n, p_1, p_2, ..., p_n) = (q^1 - x, q^2 - x, ..., q^n - x, p_1, p_2, ..., p_n).$$

If now $X \in \mathbb{R}^3$, the corresponding fundamental vector field of the action is

$$\tilde{\phi}_*(X)(q^1, q^2, ..., q^n, p_1, p_2, ..., p_n) = (-X, -X, ..., -X, 0, 0, ..., 0).$$

Hence the momentum map $\mu : T^*\mathbb{R}^{3n} \to \mathbb{R}^3$ satisfies

$$\mu(q^1, q^2, ..., q^n, p_1, p_2, ..., p_n)(X) = \sum_{j=1}^{n} \langle -X, p_j\rangle = \langle X, -\sum_{j=1}^{n} p_j\rangle.$$

In other words, the momentum map in this case is the total linear momentum

$$\mu(q^1, q^2, ..., q^n, p_1, p_2, ..., p_n) = -\sum_{j=1}^{n} p_j.$$

This example justifies the use of the term momentum map.

(b) Let G be a Lie group with Lie algebra $\mathfrak{g}$ and $\mathcal{O} \subset \mathfrak{g}^*$ be a coadjoint orbit. As we saw in Examples 1.4(b) and 1.5(b), the transitive action of G on $\mathcal{O}$ is Poisson with momentum map $\mu : \mathcal{O} \to \mathfrak{g}^*$ given by the formula $\mu(\nu) = -\nu$ for every $\nu \in \mathcal{O}$. In other words, the momentum map is minus the inclusion of $\mathcal{O}$ in $\mathfrak{g}^*$, which is of course G-equivariant.

(c) Let h be the usual hermitian product and ω the standard symplectic 2-form in $\mathbb{C}^n$ defined by the formula $\omega(v,w) = \mathrm{Re}h(Jv,w)$, where $J : \mathbb{C}^n \to \mathbb{C}^n$ is multiplication by i. The natural group action $\phi : U(n) \times \mathbb{C}^n \to \mathbb{C}^n$ preserves h (by definition) and is symplectic, since the elements of $U(n)$ commute with J. If $X \in \mathfrak{u}(n)$, the corresponding fundamental vector field is $\phi_*(X)(z) = Xz$ and therefore

$$(i_{\phi_*(X)}\omega)_z(v) = \omega(Xz, v) = \mathrm{Re}h(JXz, v)$$

for every $v \in T_z\mathbb{C}^n$ and $z \in \mathbb{C}^n$. Let now $\rho(X) : \mathbb{C}^n \to \mathbb{R}$ be the smooth function defined by

$$\rho(X)(z) = \frac{i}{2}h(Xz, z)$$

which takes indeed real values since

$$h(Xz,z)=h(z,\bar{X}^t z)=h(z,-Xz)=-\overline{h(Xz,z)},$$

because $X\in\mathfrak{u}(n)$. Observe that

$$h(X(z+v),z+v)-h(Xz,z)=h(Xv,v)+h(Xz,v)-\overline{h(Xz,v)}$$

and

$$\lim_{v\to 0}\frac{h(Xv,v)}{\|v\|}=0.$$

It follows that

$$d\rho(X)(z)v=\frac{i}{2}[h(Xz,v)-\overline{h(Xz,v)}]=\mathrm{Re}(ih(Xz,v))=(i_{\phi_*(X)}\omega)_z(v)$$

for every $v\in T_z\mathbb{C}^n$ and $z\in\mathbb{C}^n$. This means that the action is Hamiltonian. Moreover, it is Poisson because for every X, $Y\in\mathfrak{u}(n)$ we have

$$\rho([X,Y])(z)=\rho(XY-YX)(z)=\frac{i}{2}[h(XYz,z)-h(YXz,z)]=$$

$$\frac{i}{2}[h(Yz,-Xz)-h(Xz,-Yz)]=\frac{i}{2}[-\overline{h(Xz,Yz)}+h(Xz,Yz)]=$$

$$\mathrm{Re}(ih(Xz,Yz))=\omega_z(Xz,Yz)=\{\rho(X),\rho(Y)\}(z).$$

In accordance to Theorem 2.2, the corresponding momentum map $\mu:\mathbb{C}^n\to\mathfrak{u}(n)^*$ is indeed $U(n)$-equivariant since

$$\mu(Az)(X)=\frac{i}{2}h(XAz,Az)=\frac{i}{2}h(\bar{A}^tXAz,z)=\frac{i}{2}h(A^{-1}XAz,z)=$$

$$\frac{i}{2}h(\mathrm{Ad}_{A^{-1}}(X)z,z)=(\mu(z)\circ\mathrm{Ad}_{A^{-1}})(X)$$

for every $z\in\mathbb{C}^n$, $A\in U(n)$ and $X\in\mathfrak{u}(n)$, because $\mathrm{Ad}_{A^{-1}}(X)=A^{-1}XA$.

(d) On $\mathbb{C}^{n\times n}$ we consider the inner product

$$\langle A,B\rangle=\frac{1}{2}\mathrm{Tr}(AB^*+BA^*),$$

where $A^*=\bar{A}^t$, and the corresponding symplectic form $\omega(A,B)=\langle iA,B\rangle$. Note that since AB^*+BA^* is hermitian, it has real eigenvalues and $\langle,\rangle$ is a euclidean inner product. The action ϕ of $U(n)$ on $\mathbb{C}^{n\times n}$ by conjugation is isometric and symplectic. If $X\in\mathfrak{u}(n)$, the corresponding fundamental vector field of the action at $A\in\mathbb{C}^{n\times n}$ is

$$\phi_*(X)(A)=\frac{d}{dt}\bigg|_{t=0}(\exp tX)A(\exp tX)^*=XA+AX^*=[X,A].$$

Let $\mu:\mathbb{C}^{n\times n}\to\mathfrak{u}(n)^*\cong\mathfrak{u}(n)$ be the smooth map defined by

$$\mu(A)=-\frac{1}{2}i[A,A^*],$$

where the identification of $\mathfrak{u}(n)^*$ with $\mathfrak{u}(n)$ is made through the restriction of the above inner product to $\mathfrak{u}(n)$. It is easy to see that μ is $U(n)$-equivariant. We shall show that μ is a momentum map. Let $\rho : \mathfrak{u}(n) \to C^\infty(\mathbb{C}^{n\times n})$ be the corresponding map defined by the formula

$$\rho(X)(A) = \mu(A)(X) = -\langle \frac{1}{2} i[A, A^*], X\rangle.$$

For every $A \in \mathbb{C}^{n\times n}$ and $H \in \mathbb{C}^{n\times n}$, we have

$$d\rho(X)(A)H = \langle \frac{1}{2} i([A^*, H] - [A, H^*]), X\rangle = \frac{1}{4}\mathrm{Tr}(i([A^*, H] - [A, H^*])(X^* - X)) =$$

$$\frac{1}{2}\mathrm{Tr}(iX([A, H^*] - [A^*, H])) = \frac{1}{2}\mathrm{Tr}(i(XAH^* - XH^*A - XA^*H + XHA^*)) =$$

$$\frac{1}{2}\mathrm{Tr}(i([X, A]H^* - H[X, A^*])) = \langle i[X, A], H\rangle = \omega(\phi_*(X)(A), H).$$

In other words $\phi_*(X) = X_{\rho(X)}$. Moreover, the action is Poisson, because for every X, $Y \in \mathfrak{u}(n)$ and $A \in \mathbb{C}^{n\times n}$ we have

$$\{\rho(X), \rho(Y)\}(A) = \langle i[X, A], [Y, A]\rangle = \frac{1}{2}\mathrm{Tr}(i[A, A^*][X, Y]) = \rho([X, Y])(A).$$

4.3 Symplectic reduction

Let (M, ω) be a connected, symplectic manifold, G a Lie group with Lie algebra $\mathfrak{g}$ and $\phi : G \times M \to M$ a symplectic action. In general, the orbit space $G\backslash M$ of the action may not be a smooth manifold (not even a Hausdorff space). Even in the case it is, it may not admit any symplectic structure, as for instance it may be odd dimensional. If the action is Poisson and there is a G-equivariant momentum map $\mu : M \to \mathfrak{g}^*$, there exists a well defined continuous map $\tilde{\mu} : G\backslash M \to G\backslash \mathfrak{g}^*$. Under certain circumstances, the level sets $\tilde{\mu}^{-1}(\mathcal{O}_a)$, $a \in \mathfrak{g}^*$, can be given a symplectic structure in a natural way. It is easy to see that the inclusion $j : \mu^{-1}(a) \hookrightarrow \mu^{-1}(\mathcal{O}_a)$ induces a continuous bijection $j_\# : G_a\backslash\mu^{-1}(a) \to G\backslash\mu^{-1}(\mathcal{O}_a)$. In certain cases, $j_\#$ is a homeomorphism or even a diffeomorphism of smooth manifolds. For example, if the action of G on M is free and proper and a is a regular value of μ, then $\mu^{-1}(\mathcal{O}_a)$ is a smooth submanifold of M and so are $G\backslash\mu^{-1}(\mathcal{O}_a)$ and $G_a\backslash\mu^{-1}(a)$. Moreover, in this case $j_\#$ is a diffeomorphism. In particular, these are true if G is compact and the action is free.

Definition 3.1. Let P, Q be two smooth manifolds and $f : P \to Q$ be a smooth map. A point $q \in Q$ is called a *clean* (or *weakly regular*) value of f if $f^{-1}(q)$ is an embedded smooth submanifold of M and $T_p f^{-1}(q) = \mathrm{Ker} f_{*p}$ for every $p \in f^{-1}(q)$.

Obviously, a regular value is always clean, but the converse is not true. For example, $(0, 0) \in \mathbb{R}^2$ is a clean, but not regular, value of the smooth function $f : \mathbb{R}^3 \to \mathbb{R}^2$ with $f(x, y, z) = (z^2, z)$.

Theorem 3.2. *Let* (M,ω) *be a symplectic manifold,* G *be a Lie group with Lie algebra* $\mathfrak{g}$ *and* $\phi : G \times M \to M$ *be a Poisson action with a* G*-equivariant momentum map* $\mu : M \to \mathfrak{g}^*$*. Let* $a \in \mathfrak{g}^*$ *be a clean value of* μ *such that the orbit space* $M_a = G_a\backslash\mu^{-1}(a)$ *is a smooth manifold and the quotient map* $\pi_a : \mu^{-1}(a) \to M_a$ *is a smooth submersion, where* G_a *is the isotropy group of* a *with respect to the coadjoint action. Then there exists a unique symplectic* 2*-form* ω_a *on* M_a *auch that* $\pi_a^*\omega_a = \omega|_{\mu^{-1}(a)}$.

Proof. First note that $\mu^{-1}(a)$ is indeed G_a-invariant, since μ is G-equivariant. Evidently, $\tilde{\omega}_a = \omega|_{\mu^{-1}(a)}$ is closed and G_a-invariant, because the action is symplectic. So there exists a unique 2-form ω_a on M_a such that $\pi_a^*\omega_a = \tilde{\omega}_a$. Since π_a is a submersion and $\tilde{\omega}_a$ is closed, so is ω_a. It remains to show that ω_a is non-degenerate.

Observe that for any $p \in \mu^{-1}(a)$ we have

$$(T_pGp)^\perp = \{v \in T_pM : \omega_p(\phi_*(X)(p), v) = 0 \text{ for every } X \in \mathfrak{g}\} =$$

$$\{v \in T_pM : \mu_{*p}(v)(X) = 0 \text{ for every } X \in \mathfrak{g}\} =$$

$$\mathrm{Ker}\mu_{*p} = T_p\mu^{-1}(a),$$

because T_pGp is generated by the values at p of the fundamental vector fields of the action. On the other hand, $\mu^{-1}(a) \cap Gp = G_ap$, since μ is G-equivariant, and therefore $T_pG_ap \subset T_p\mu^{-1}(a) \cap T_pGp$. Actually, we have equality. To see this, let $v \in T_p\mu^{-1}(a) \cap T_pGp$. There exists $X \in \mathfrak{g}^*$ such that $v = \phi_*(X)(p)$ and since $T_p\mu^{-1}(a) = \mathrm{Ker}\mu_{*p}$, we have

$$0 = \mu_{*p}(v)(X) = \left.\frac{d}{dt}\right|_{t=0} \mu(\phi_{\exp(tX)}(p)) = \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}^*_{\exp(tX)}(\mu(p)) = (\mathrm{Ad}^*)_*(X)(a).$$

This means $X \in \mathfrak{g}_a$, the Lie algebra of G_a, or in other words $v \in T_pG_ap$.

It follows that $T_pG_ap = T_p\mu^{-1}(a) \cap (T_p\mu^{-1}(a))^\perp$. Suppose now that $v \in T_p\mu^{-1}(a)$ is such that $\tilde{\omega}_a(v, w) = 0$ for every $w \in T_p\mu^{-1}(a)$. Then $v \in (T_p\mu^{-1}(a))^\perp$ and so $v \in T_pG_ap$. Hence $(\pi_a)_{*p}(v) = 0$. This proves that ω_a is non-degenerate. □

Under the assumptions of Theorem 3.2 let $H \in C^\infty(M)$ be G-invariant. As we observed in the beginning of section 3.2, the momentum map μ is constant along the integral curves of the Hamiltonian vector field X_H, which is obviously G-invariant, since the action is symplectic. Thus, X_H is tangent to $\mu^{-1}(a)$ and is G_a-invariant. Let $H_a \in C^\infty(M_a)$ be defined by $H_a \circ \pi_a = H$ and X_{H_a} be the corresponding Hamiltonian vector field on the symplectic manifold (M_a, ω_a). Then $(\pi_a)_*X_H = X_{H_a}$, because for every $p \in \mu^{-1}(a)$ and $v \in T_p\mu^{-1}(a)$ we have

$$(\omega_a)_{\pi_a(p)}((\pi_a)_{*p}(X_H(p)), (\pi_a)_{*p}(v)) = ((\pi_a)^*\omega_a)_p(X_H(p), v) = \omega_p(X_H(p), v) =$$

$$dH(p)(v) = dH_a(\pi_a(p))((\pi_a)_{*p}(v)) = (\omega_a)_{\pi_a(p)}(X_{H_a}(\pi_a(p)), (\pi_a)_{*p}(v)).$$

The Hamiltonian vector field X_{H_a} is called the *reduced Hamiltonian vector field.* This is a geometric way to use the symmetry group G of X_H in order to reduce the number of differential equations we have to solve, if we want to find its integral curves.

Examples 3.3. (a) Let M be a symplectic manifold and $H \in C^\infty(M)$ be such that the Hamiltonian vector field X_H is complete. Its flow is a Poisson group action of $\mathbb{R}$ on M with momentum map H itself. Since $\mathbb{R}$ is abelian, the coadjoint action is trivial. If now $a \in \mathbb{R}$ is a clean value of H, then according to Theorem 3.2 the orbit space $\mathbb{R}\backslash H^{-1}(a)$ has a natural symplectic structure.

(b) Let $SO(3,\mathbb{R})$ act on $T^*\mathbb{R}^3 \cong \mathbb{R}^3 \times \mathbb{R}^3$ as in the Example 2.4(a). As we saw, the momentum map $\mu : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ is the angular momentum

$$\mu(q,p) = q \times p.$$

The Jacobian matrix of μ at (q,p) is $(-\hat{p}, \hat{q})$, and so every non-zero $v \in \mathbb{R}^3$ is a regular value of μ. The isotropy group of v is the group of rotations of $\mathbb{R}^3$ around the axis generated by v, hence isomorphic to S^1. Thus, the orbit space $S^1\backslash\mu^{-1}(v)$ has a symplectic structure.

(c) Let $\phi : S^1 \times \mathbb{C}^{n+1} \to \mathbb{C}^{n+1}$ be the action with $\phi(e^{it}, z) = e^{it}z$. As we saw in the Example 2.4(c), the action is Poisson with respect to the standard symplectic structure of $\mathbb{C}^{n+1}$ and the momentum map $\mu : \mathbb{C}^{n+1} \to \mathfrak{u}(1)^* = (i\mathbb{R})^* \cong \mathbb{R}$ is given by the formula

$$\mu(z) = \frac{i}{2}h(iz,z) = -\frac{1}{2}|z|^2.$$

Now $a = -\dfrac{1}{2}$ is a regular value of μ and $\mu^{-1}(a) = S^{2n+1}$. Since S^1 is abelian, we conclude that $\mathbb{C}P^n = S^1\backslash S^{2n+1}$ has a symplectic 2-form. It is clear from the definitions that this is exactly the fundamental symplectic 2-form of the Fubini-Study metric.

(d) Let M be a symplectic $2n$-manifold and H_1,...,$H_k \in C^\infty(M)$ such that the Hamiltonian vector fields X_{H_1},...,X_{H_k} are complete. If $\{H_i, H_j\} = 0$ for every i, $j = 1, 2, ..., k$, then their flows commute and define a Poisson action of $\mathbb{R}^k$ on M with momentum map $\mu = (H_1, ..., H_k) : M \to \mathbb{R}^k$. Since $\mathbb{R}^k$ is abelian, we get a symplectic structure on the orbit space $\mathbb{R}^k\backslash\mu^{-1}(a)$ for every clean value $a \in \mathbb{R}^k$ of μ. In the next section we shall examine this situation in further detail when $k = n$.

4.4 Completely integrable Hamiltonian systems

Let (M,ω) be a connected, symplectic $2n$-manifold and $H_1 \in C^\infty(M)$. The triple (H_1, M, ω) is called a *completely integrable Hamiltonian system* if there are H_2,...,$H_n \in C^\infty(M)$ such that $\{H_i, H_j\} = 0$ for every $1 \le i, j \le n$ and the differential 1-forms dH_1, dH_2,...,dH_n are linearly independent on a dense open set $D \subset M$. In this section we shall always assume that we have such a system.

For every $p \in M$ the set $\{X_{H_1}(p), X_{H_2}(p), ..., X_{H_n}(p)\}$ generates an isotropic linear subspace of T_pM. If $p \in D$, then it is a basis of a Lagrangian subspace of T_pM. If $f = (H_1, H_2, ..., H_n) : M \to \mathbb{R}^n$, then $f|_D$ is a smooth submersion and so the connected components of the fibers $f^{-1}(y) \cap D$, $y \in \mathbb{R}^n$, are the leaves of a foliation of D by Lagrangian submanifolds, because $f_{*p}(X_{H_i}(p)) = 0$ for every $1 \le i \le n$.

Suppose that the Hamiltonian vector fields X_{H_1}, X_{H_2},..., X_{H_n} are complete. Since their flows commute, they define a Poisson group action $\phi : \mathbb{R}^n \times M \to M$ with fundamental vector fields X_{H_1}, X_{H_2},..., X_{H_n} and momentum map f. Let $y \in \mathbb{R}^n$ be a regular value of f. Then $f^{-1}(y) \subset D$ is a $\mathbb{R}^n$-invariant, regular n-dimensional submanifold of M. The vector fields X_{H_1}, X_{H_2},..., X_{H_n} are tangent to $f^{-1}(y)$, and since they are linearly independent at every point of $f^{-1}(y)$, every orbit in $f^{-1}(y)$ is an open subset of $f^{-1}(y)$. This implies that every connected component N of $f^{-1}(y)$ is an orbit of the action. Thus, N is diffeomorphic to the homogenenous space $\mathbb{R}^n/\Gamma_p$, where Γ_p is the isotropy group of p. Note that Γ_p does not depend on p, but only on N, since $\mathbb{R}^n$ is abelian. Also, Γ_p is a 0-dimensional closed subgroup of $\mathbb{R}^n$ and therefore is discrete. The discrete subgroups of $\mathbb{R}^n$ are described as follows.

Lemma 4.1. *Let $\Gamma \leq \mathbb{R}^n$ be a non-trivial discrete subgroup. Then Γ is a lattice, that is there exist $1 \leq k \leq n$ and linearly independent vectors v_1,...,v_k such that*

$$\Gamma = \mathbb{Z}v_1 + ... + \mathbb{Z}v_k.$$

Proof. Let $u_1 \in \Gamma \setminus \{0\}$. Since Γ is discrete, there exists $\lambda > 0$ such that $\lambda u_1 \in \Gamma$ and $\Gamma \cap (-\lambda, \lambda)u_1 = \{0\}$. Let $v_1 = \lambda u_1$ and so $\Gamma \cap \mathbb{R}v_1 = \mathbb{Z}v_1$. If $\Gamma = \mathbb{Z}v_1$, then $k = 1$ and we have finished. Suppose that $\Gamma \neq \mathbb{Z}v_1$ and $u_2 \in \Gamma \setminus \mathbb{Z}v_1$ be such that $\Gamma \cap \mathbb{R}u_2 = \mathbb{Z}u_2$. Then, v_1 and u_2 are linearly independent. Let

$$P(v_1, u_2) = \{t_1v_1 + t_2u_2 : t_1, t_2 \in [0, 1]\}$$

be the parallelogram generated by v_1 and u_2. The set $\Gamma \cap P(v_1, u_2)$ is finite, because Γ is discrete. So there exists $v_2 \in P(v_1, u_2)$ such that $\Gamma \cap P(v_1, v_2) = \{0, v_1, v_2, v_1+v_2\}$. It follows now that

$$\Gamma \cap (\mathbb{R}v_1 \oplus \mathbb{R}v_2) = \mathbb{Z}v_1 + \mathbb{Z}v_2,$$

because if there exist t_1, $t_2 \in \mathbb{R} \setminus \mathbb{Z}$ such that $t_1v_1 + t_2v_2 \in \Gamma$, then

$$(t_1 - [t_1])v_1 + (t_2 - [t_2])v_2 \in \Gamma \cap P(v_1, v_2),$$

contradiction. If $\Gamma = \mathbb{Z}v_1 + \mathbb{Z}v_2$, then $k = 2$ and we have finished. If not, then we proceed inductively using the same argument repeatedly, replacing the parallelograms with parallelopipeds etc. Since $\mathbb{R}^n$ has finite dimension, we end up with linearly independent vectors v_1,...,v_k such that $\Gamma = \mathbb{Z}v_1 + ... + \mathbb{Z}v_k$. □

Corollary 4.2. *Let $\Gamma \leq \mathbb{R}^n$ be a non-trivial discrete subgroup. Then there exists $1 \leq k \leq n$ such that the homogenenous space $\mathbb{R}^n/\Gamma$ is diffeomorphic to $T^k \times \mathbb{R}^{n-k}$. If $\mathbb{R}^n/\Gamma$ is compact, then $k = n$ and $\mathbb{R}^n/\Gamma$ is diffeomorphic to the n-torus T^n.*

Proof. From Lemma 4.1 there exist $1 \leq k \leq n$ and linearly independent vectors v_1,...,v_k such that $\Gamma = \mathbb{Z}v_1 + ... + \mathbb{Z}v_k$. We complete to a basis $\{v_1, ..., v_k, v_{k+1}, ..., v_n\}$ of $\mathbb{R}^n$ and consider the linear isomorphism $T : \mathbb{R}^n \to \mathbb{R}^n$ with $T(v_j) = e_j$, $1 \leq j \leq n$. Then, $T(\Gamma) = \mathbb{Z}^k \times \{0\}$, and so T imduces a diffeomorphism $\tilde{T} : \mathbb{R}/\Gamma \to T^k \times \mathbb{R}^{n-k}$. The rest is obvious. □

Note that if N is compact then the restrictions of the Hamiltonian vector fields X_{H_1}, X_{H_2},..., X_{H_n} to N are automatically complete. So, we have arrived at the following.

Theorem 4.3. (Arnold-Liouville) *Let $y \in \mathbb{R}^n$ be a regular value of f and N be a connected component of $f^{-1}(y)$.*

(i) If N is compact, then it is diffeomorphic to the n-torus T^n.

(ii) If N is not compact and X_{H_1}, X_{H_2},..., X_{H_n} are complete, then N is diffeomorphic to $T^k \times \mathbb{R}^{n-k}$ for some $1 \leq k \leq n$. □

It is not hard now to describe the flow of the Hamiltonian vector field X_{H_1} on N. Let $p \in N$ and $\tilde{\phi^p} : \mathbb{R}^n/\Gamma \to N$ be the diffeomorphism which is induced by $\phi^p = \phi(.,p) : \mathbb{R}^n \to N$. Let $(\psi_t)_{t\in\mathbb{R}}$ be the flow of X_{H_1} on N and $\tilde{\psi_t} = (\tilde{\phi^p})^{-1} \circ \psi_t \circ \tilde{\phi^p}$, $t \in \mathbb{R}$, be the conjugate flow on $\mathbb{R}^n/\Gamma$. Then

$$\tilde{\psi_t}([t_1, ..., t_n]) = (\tilde{\phi^p})^{-1}(\psi_t(\phi((t_1, ..., t_n), p))) =$$

$$(\tilde{\phi^p})^{-1}(\phi((t + t_1, t_2, ..., t_n), p)) = [t + t_1, t_2, ..., t_n].$$

In other words, $\psi_t([v]) = [v + te_1]$ for every $v \in \mathbb{R}^n$ and $t \in \mathbb{R}$. Using the notations of the proof of Corollary 4.2, let $T(e_1) = (\nu_1, ..., \nu_n)$ and let $\chi_t = \tilde{T} \circ \tilde{\psi_t} \circ \tilde{T}^{-1}$, $t \in \mathbb{R}$, be the conjugate flow on $T^k \times \mathbb{R}^{n-k}$. Then,

$$\chi_t(e^{2\pi i t_1}, ..., e^{2\pi i t_k}, t_{k+1}, ..., t_n) = \tilde{T}(\tilde{\psi_t}([t_1 v_1 + ... + t_n v_n])) = \tilde{T}([te_1 + t_1 v_1 + ... + t_n v_n]).$$

Since

$$T(te_1 + t_1 v_1 + ... + t_n v_n) = tT(e_1) + t_1 e_1 + ... + t_n e_n = (t_1 + t\nu_1, ..., t_n + t\nu_n)$$

it follows that

$$\chi_t(e^{2\pi i t_1}, ..., e^{2\pi i t_k}, t_{k+1}, ..., t_n) = (e^{2\pi i(t_1 + t\nu_1)}, ..., e^{2\pi i(t_k + t\nu_k)}, t_{k+1} + t\nu_{k+1}, ..., t_n + t\nu_n).$$

This shows that the flow of X_{H_1} on N is smoothly conjugate to a linear flow on $T^k \times \mathbb{R}^{n-k}$. In case N is compact, then $k = n$ and the real numbers ν_1,...,ν_n are called the *frequences of the flow* on N. As is well known, if they are linearly independent over $\mathbb{Q}$, then the flow on N is uniquely ergodic and every orbit is dense in N.

In the rest of this section we shall study more closely the case of a compact connected component N of $f^{-1}(y)$, where $y \in \mathbb{R}^n$ is a regular value of f. We are mainly interested in the structure of (M, ω) around N. Since N is compact, there exist an open neighbourhood U of y and a ϕ-invariant neighbourhood V of N such that $\overline{V}$ is compact, $f(V) = U$ and $f|_V : V \to U$ is a submersion with compact fibers. Therefore, $f|_V$ is a locally trivial fibration with Lagrangian fibers diffeomorphic to the n-torus T^n. Shrinking U to an open neighbourhood of y diffeomorphic to $\mathbb{R}^n$, we get an open neighbourhood V of N diffeomorphic to $U \times N$ and so to $\mathbb{R}^n \times T^n$.

The orbits of the restriction of the Poisson group action ϕ on V are the fibers $(f|_V)^{-1}(q)$, $q \in U$, and the isotropy group of a point on $(f|_V)^{-1}(q)$ depends only on q, since $\mathbb{R}^n$ is abelian. We shall show first that the isotropy groups vary smoothly with q. Let $t_0 \in \Gamma_p \setminus \{0\}$, where $p \in N$, and let $s : U \to V$ be a smooth section, that is $f \circ s = id$. Identifying a small open neighbourhood B of p in N with $\mathbb{R}^n$, there exist an open neighbourhood W of t_0 in $\Gamma_p \setminus \{0\}$ and an open neighbourhood U_y of y in U such that

$$pr(\phi(t, s(q))) - s(q) \in B$$

for every $t \in W$ and $q \in U_y$, where $pr : V \to N$ is the projection. The smooth map $G : W \times U_y \to B$ with

$$G(t, q) = pr(\phi(t, s(q))) - s(q)$$

is thus well defined and 0 is a regular value of $G(., y)$. From the Implicit Function Theorem there exist an open neighbourhood U'_y of y and a smooth map $h : U'_y \to \mathbb{R}^n$ such that $G(h(q), q) = 0$ for every $q \in U'_y$ and $h(y) = t_0$. In other words,

$$\phi(h(q), s(q)) = s(q)$$

for every $q \in U'_y$. Varying now t_0 in a basis of the lattice Γ_p, we conclude that there exists smooth functions v_1,...,$v_n : U \to \mathbb{R}^n$ such that

$$\Gamma_p = \mathbb{Z}v_1(f(p)) + ... + \mathbb{Z}v_n(f(p))$$

for every $p \in V$, shrinking U and V appropriately.

Let now Y_i be the infinitesimal generator of the smooth flow $\phi^i : \mathbb{R} \times V \to V$ with $\phi^i(t, p) = \phi(tv_i(f(p)), p)$. Then,

$$Y_i(p) = \sum_{j=1}^{n} v_{i,j}(f(p))X_{H_j}(p),$$

where $v_i = (v_{i,1}, ..., v_{i,n})$. Obviously, the flow ϕ^i is periodic with period 1, the vector fields Y_1,...,Y_1 are linearly independent and $[Y_i, Y_j] = 0$, because $[X_{H_i}, X_{H_j}] = 0$ and X_{H_1},...,X_{H_n} are tangent to the fibers of $f|_V$. This means that there is a well defined group action of the n-torus T^n on V with fundamental vector fields Y_1,...,Y_n, whose orbits are the fibers of $f|_V$. We shall show that this action is Poisson and and we shall construct a momentum map. First note that since we have selected U to be contractible and V is diffeomorphic to $U \times N$, the inclusion $N \subset V$ induces an isomorphism in cohomology. It follows that $\omega|_V$ is exact since $\omega|_N = 0$, because N is Lagrangian. Let η be a smooth 1-form on V such that $\omega|_V = -d\eta$ and for $1 \leq i \leq n$ let $g_i : V \to \mathbb{R}$ be the smooth function defined by

$$g_i(p) = \int_0^1 (i_{Y_i}\eta)(\phi^i(t, p))dt.$$

Since T^n is abelian, it suffices to show that $Y_i = X_{g_i}$ for all $1 \leq i \leq n$, that is

$$\omega(Y_i(p), Z(p)) = dg_i(p)(Z(p))$$

for every $Z(p) \in T_pM$ and $p \in V$. Then, $(g_1, ..., g_n)$ will be a momentum map.

Let $Z(p) \in T_pM$ and let Z be an extension to a smooth vector field on V which is invariant by the action of T^n, that is $[Y_i, Z] = 0$ for every $1 \leq i \leq n$. Differentiating g_i we get

$$dg_i(p)(Z(p)) = \int_0^1 d(i_{Y_i}\eta)(\phi^i(t, p))(Z(\phi^i(t, p)))dt,$$

since Z is ϕ^i-invariant. But

$$d(i_{Y_i}\eta)(Z) = L_Z(i_{Y_i}\eta) = i_{Y_i}(L_Z\eta) + \eta([Z, Y_i]) = (L_Z\eta)(Y_i) =$$
$$d(i_Z\eta)(Y_i) + i_Z(d\eta)(Y_i) = d(i_Z\eta)(Y_i) + \omega(Y_i, Z)$$

and

$$\int_0^1 d(i_Z\eta)(Y_i)(\phi^i(t,p))dt = (i_z\eta)(\phi^i(1,p)) - (i_z\eta)(\phi^i(0,p)) = 0,$$

since the flow ϕ^i is periodic with period 1. Consequently,

$$dg_i(p)(Z(p)) = \int_0^1 \omega(Y_i, Z)(\phi^i(t,p))dt$$

and it suffices to show that $\omega(Y_i, Z)$ is ϕ^i-invariant. Indeed, since $[Y_i, Z] = 0$, we have $L_{Y_i}(\omega(Y_i, Z)) = i_{Y_i}(L_{Y_i}\omega)(Z)$ and

$$i_{Y_i}(L_{Y_i}\omega) = i_{Y_i}(d(i_{Y_i}\omega)) = \sum_{j=1}^n Y_i(v_{i,j} \circ f)dH_j - \sum_{j=1}^n Y_i(H_j)d(v_{i,j} \circ f) = 0,$$

because $Y_i(v_{i,j} \circ f) = Y_i(H_j) = 0$, since H_j and $v_{i,j} \circ f$, $1 \leq j \leq n$, are constant along the orbits of Y_i.

Since now $(g_1, ..., g_n)$ is a momentum map of the action of T^n, it is constant on the fibers of $f|_V$ and so it is a function of $f(p)$, $p \in V$. We shall henceforth consider $(g_1, ..., g_n)$ as a function defined on U. Its rank at every point is n because

$$dg_1 \wedge ... \wedge dg_n = \frac{1}{n!} i_{Y_1}...i_{Y_n}(\omega \wedge ... \wedge \omega).$$

Considering local coordinates θ_1,...θ_n on N around a point $p \in N$, the smooth map $g = (g_1, ..., g_n, \theta_1, ..., \theta_n) : V \to \mathbb{R}^{2n}$ defines local coordinates in a small neighbourhood of p in M. Moreover, since $(g_1, ..., g_n)$ is a momentum map of the n-torus action and the action on the fibers is simply translation, we have

$$(g^{-1})^*\omega = \sum_{i=1}^n d\theta_i \wedge dg_i + \sum_{i<j} a_{ij} dg_i \wedge dg_j$$

for some smooth functions a_{ij}, $1 \leq i < j \leq n$. The fact that ω is closed implies that the 2-form

$$\alpha = \sum_{i<j} a_{ij} dg_i \wedge dg_j$$

is also closed, which means that a_{ij} does not depend on θ_1,...,θ_n. Having chosen U contractible, there exists a smooth 1-form β such that $\alpha = -d\beta$. So, there are smooth functions β_1,...,β_n of g_1,..,g_n such that

$$\beta = \sum_{i=1}^n \beta_i dg_i.$$

Putting now $\psi_i = \theta_i - \beta_i$ we get

$$(g^{-1})^*\omega = \sum_{i=1}^n d\psi_i \wedge dg_i.$$

The local coordinates $(g_1, ..., g_n, \psi_1, ..., \psi_n)$ are called *action angle coordinates*.

4.5 Hamiltonian torus actions

Let (M, ω) be a connected, symplectic manifold, G a Lie group with Lie algebra $\mathfrak{g}$ and $\phi : G \times M \to M$ a symplectic action. For $X \in \mathfrak{g}$ let G_X denote the closure in G of the one-parameter subgroup of G generated by X. The fixed point set $\mathrm{Fix}_{G_X}(M)$ of the restricted action of G_X on M coincides with the zeros of the fundamental vector field $\phi_*(X)$. If the action of G is Hamiltonian, the zeros of X are precisely the critical points of the Hamiltonian function of X. So the set of the critical points of the Hamiltonian of X coincides with $\mathrm{Fix}_{G_X}(M)$, which is a submanifold of M, if G_X is compact.

Lemma 5.1. *Let (M, ω) be a connected, symplectic manifold, T^m the m-torus and $\phi : T^m \times M \to M$ a symplectic action. If K is closed subgroup of T^m, then $\mathrm{Fix}_K(M)$ is a T^m-invariant symplectic submanifold of M.*

Proof. It is obvious that $\mathrm{Fix}_K(M)$ is T^m-invariant, because T^m is abelian. Let J be a T^m-invariant almost complex structure on M which is compatible with ω and $\langle , \rangle$ be the corresponding T^m-invariant Riemannian metric. For every $p \in \mathrm{Fix}_K(M)$ and $h \in K$, the linear map $(\phi_h)_{*p} : T_pM \to T_pM$ is symplectic and preserves J and $\langle , \rangle$ at p. Therefore, $\mathrm{Ker}(id - (\phi_h)_{*p})$ is a symplectic linear subspace of T_pM and so is

$$\mathrm{Fix}_K(T_pM) = \bigcap_{h \in K} \mathrm{Ker}(id - (\phi_h)_{*p}).$$

There is some $\delta > 0$ such that the exponential map $\exp_p$ at p, with respect to the Levi-Civita connection of $\langle , \rangle$, maps the open ball of radius δ in T_pM centered at $0 \in T_pM$ diffeomorphically onto an open neighbourhood V of p in M. Since ϕ_h is a Riemannian isometry for every $h \in K$, it commutes with the exponential map. This implies that

$$\phi_h(\exp_p(u)) = \exp_p((\phi_h)_{*p}(u))$$

for every $u \in T_pM$ with $\|u\| < \delta$. It follows that

$$V \cap \mathrm{Fix}_K(M) = \exp_p(\mathrm{Fix}_K(T_pM)).$$

This proves that $\mathrm{Fix}_K(M)$ is a submanifold of M whose tangent space at p is $\mathrm{Fix}_K(T_pM)$ and is therefore symplectic. □

In the case of a Hamiltonian torus action the Hamiltonians of the fundamental vector fields of the action are Morse-Bott functions. This observation is due to M. Atiyah. As we shall see later, it results in constraints on the structure of momentum maps.

Proposition 5.2. *Let (M, ω) be a connected, symplectic manifold, T^m the m-torus with Lie algebra $\mathfrak{t}_m = (i\mathbb{R})^m$ and $\phi : T^m \times M \to M$ a Hamiltonian action. Then, for every $X \in \mathfrak{t}_m$ the Hamiltonian function $H : M \to \mathbb{R}$ of $\phi_*(X)$ is a Morse-Bott function whose critical submanifolds are symplectic. Moreover, all the Morse indices are even.*

Proof. Let p be a critical point of H and let $(\phi_t)_{t\in\mathbb{R}}$ denote the flow of $\phi_*(X)$. As in the proof of Lemma 5.1 we fix a T^m-invariant almost complex structure J on M which is compatible with ω and a corresponding T^m-invariant Riemannian metric $\langle,\rangle$. Then $(\phi_t)_{*p}$ is a unitary automorphism of the Kähler vector space T_pM for every $t \in \mathbb{R}$. Therefore,

$$\hat{\phi}_p = \frac{d}{dt}\bigg|_{t=0} (\phi_t)_{*p}$$

is a skew-hermitian endomorphism of T_pM and $(\phi_t)_{*p} = e^{t\hat{\phi}_p}$ for all $t \in \mathbb{R}$. It follows that

$$\mathrm{Ker}\hat{\phi}_p = \bigcap_{t\in\mathbb{R}} \mathrm{Ker}(id - (\phi_t)_{*p}) = T_p\mathrm{Fix}_{G_X}(M).$$

Also, for every smooth vector field Y on M we have

$$\hat{\phi}_p(Y(p)) = \frac{d}{dt}\bigg|_{t=0} (\phi_t)_{*p}(Y(p)) = [Y, \phi_*(X)](p).$$

Now for every pair of smooth vector fields Y_1, Y_2 on M the value of the Hessian $\mathrm{Hess}H(p)$ of the Hamiltonian H at p is

$$\mathrm{Hess}H(p)(Y_1(p), Y_2(p)) = Y_1(p)(Y_2H) = Y_1(p)(dH(Y_2))$$

$$= Y_1(p)(\omega(\phi_*(X), Y_2)) = L_{Y_1}\omega(\phi_*(X), Y_2)(p)$$

$$= (L_{Y_1}\omega)(p)(\phi_*(X)(p), Y_2(p)) + \omega([Y_1, \phi_*(X)](p), Y_2(p)) + \omega(\phi_*(X)(p), [Y_1, Y_2](p))$$

$$= \omega([Y_1, \phi_*(X)](p), Y_2(p)) = \omega(\hat{\phi}_p(Y_1(p)), Y_2(p)) = \langle J(p)(\hat{\phi}_p(Y_1(p))), Y_2(p)\rangle.$$

Hence $J(p) \circ \hat{\phi}_p$ is a self-adjoint operator with respect to the Riemannian metric which represents $\mathrm{Hess}H(p)$. Note that it also commutes with $J(p)$, because

$$(J(p) \circ \hat{\phi}_p) \circ J(p) = -J(p) \circ \hat{\phi}_p \circ J(p)^* = J(p) \circ \hat{\phi}_p^* \circ J(p)^*$$

$$= J(p) \circ (J(p) \circ \hat{\phi}_p)^* = J(p) \circ (J(p) \circ \hat{\phi}_p),$$

and

$$\mathrm{Ker}(J(p) \circ \hat{\phi}_p) = \mathrm{Ker}\hat{\phi}_p = T_p\mathrm{Fix}_{G_X}(M).$$

This proves that H is a Morse-Bott function. Moreover, since $J(p) \circ \hat{\phi}_p$ commutes with $J(p)$, its eigenspaces are $J(p)$-invariant and therefore have even dimensions. □

In case M is compact and connected, each fiber of the Hamiltonian H in Proposition 5.2 is connected. This is a consequence of the connectivity lemma for Morse-Bott functions presented in the Appendix to this section as Proposition 5.9.

Let now (M, ω) be a compact, connected, symplectic $2n$-manifold with a Hamiltonian action $\phi : T^m \times M \to M$. From Lemma 5.1, the fixed point set $\mathrm{Fix}_{T^m}(M)$ of the action is a compact submanifold of M and so has a finite number of connected components. By Theorem 1.7, the action is Poisson, which means that there exists a lift $\rho : \mathfrak{t}_m \to C^\infty(M)$ and a dual momentum map $\mu : M \to \mathfrak{t}_m^*$. We shall analyze the image of such a momentum map.

Lemma 5.3. *The momentum map is constant on each connected component of* $\mathrm{Fix}_{T^m}(M)$.

Proof. The isotropy group of a point $p \in M$ is a subtorus of T^m whose Lie algebra consists of the elements of $\mathfrak{t}_m$ for which the corresponding fundamental vector field of the action vanishes at p. On the other hand, the transpose of $\mu_{*p} : T_pM \to \mathfrak{t}_m^*$ is the linear map $(\mu_{*p})^* : \mathfrak{t}_m \to T_p^*M$ given by the formula

$$(\mu_{*p})^*(X) = d\rho(X)(p).$$

So, the Lie algebra of the isotropy group of p is precisely $\mathrm{Ker}(\mu_{*p})^*$.

If now $p \in \mathrm{Fix}_{T^m}(M)$, then its isotropy group is T^m and the above shows that $(\mu_{*p})^* = 0$, hence also $\mu_{*p} = 0$. $\square$

We pick a basis $\{X_1, X_2, ..., X_m\}$ of commuting elements of $\mathfrak{t}_m$. For example, we can take $X_j = \dfrac{\partial}{\partial t_j}$, $1 \leq j \leq m$, at the identity element of T^m. If $X = \sum_{j=1}^{m} a_j X_j$, then

$$\mu(p) = \sum_{j=1}^{m} a_j \rho(X_j)(p)$$

for every $p \in M$. So, if we identify $\mathfrak{t}_m^*$ with $\mathbb{R}^m$ with respect to this basis of $\mathfrak{t}_m$, the momentum map becomes the smooth function

$$\mu = (\mu_1, \mu_2, ..., \mu_m) : M \to \mathbb{R}^m,$$

where we have put $\mu_j = \rho(X_j)$, $1 \leq j \leq m$. Also, μ_1, μ_2,..., μ_m are Morse-Bott functions whose critical sumanifolds are symplectic and have even Morse indices.

The main subject of this section is the following famous convexity theorem for the image of the momentum map, which was proved independently by M. Atiyah and V. Guillemin-S.Sternberg.

Theorem 5.4. *Let* (M, ω) *be a compact, connected, symplectic manifold and* $\phi : T^m \times M \to M$ *a Hamiltonian action of the* m*-torus with momentum map* $\mu : M \to \mathbb{R}^m$. *Then the following hold.*
(a) The fibers of μ *are connected.*
(b) The image $\mu(M)$ *of* μ *is a convex subset of* $\mathbb{R}^m$.
(c) $\mu(M)$ *is the convex hull of the finite set of the values of* μ *on* $\mathrm{Fix}_{T^m}(M)$.

In order to facilitate the inductive proof of Theorem 5.4, we shall actually prove the following slightly more general version.

Theorem 5.5. *Let* (M, ω) *be a compact, connected, symplectic manifold and let* $\phi : T^m \times M \to M$ *be a Hamiltonian action of the* m*-torus with lift* $\rho : \mathfrak{t}_m \to C^\infty(M)$ *and a dual momentum map* $\mu : M \to \mathfrak{t}_m^*$. *If* f_1, f_2,..., $f_k \in \mathrm{Im}\rho$ *and they commute with respect to the Poisson bracket, then for the smooth function* $f = (f_1, f_2, ..., f_k) : M \to \mathbb{R}^k$ *the following hold.*

(a) The fibers of f are connected.
(b) The image $f(M)$ of f is a convex subset of $\mathbb{R}^k$.
(c) The set of common critical points of f_1, f_2,..., f_k is a disjoint finite union of compact, connected submanifolds of M on each of which f takes a constant value and $f(M)$ is the convex hull of these values.

In the proof of Theorem 5.5(a) we shall use the following.

Lemma 5.6. *Let $\phi : T^m \times M \to M$ be a smooth action on a smooth n-manifold M. Then every point $p \in M$ has an open neighbourhood U such that only a finite number of subgroups of T^m occur as isotropy groups of the points of U.*

Proof. The proof will be carried out by induction on the dimension n of M. For $n = 0$, the conclusion is trivial. Suppose that it is true for manifolds of dimension $n-1$. Observe that the identity element of T^m has an open neighbourhood which contains only the trivial subgroup of T^m and the same is true for the quotient group T^m/T^m_p, which is also a torus, where T^m_p is the isotropy group of p.

First we shall prove that $T^m_x \subset T^m_p$ for points $x \in M$ sufficiently close to p. Indeed, otherwise there exists a sequence of points $(x_k)_{k\in\mathbb{N}}$ converging to p such that $T^m_{x_k}$ is not contained in T^m_p. Let U be an open neighbourhood of the identity in T^m/T^m_p which contains only the trivial subgroup and let $V \subset T^m$ be its inverse image under the quotient map. Then all subgroups of T^m which are contained in V are subgroups of T^m_p. So there are elements $t_k \in T^m_{x_k} \setminus V$, $k \in \mathbb{N}$. Since $T^m \setminus V$ is compact, we may assume that the sequence $(t_k)_{k\in\mathbb{N}}$ converges to some element $t \in T^m \setminus V$. But then the sequence $(x_k)_{k\in\mathbb{N}}$ converges to $\phi(t,p) \neq x$. This contradiction proves that there exists an open neighbourhood W of p such that $T^m_x \subset T^m_p$ for every $x \in W$. So, it suffices to consider only the restricted action of T^m_p on M. Since T^m_p is compact, there is a T^m_p-invariant Riemannian metric on M. Now T^m_p acts on T_pM by linear isometries, because p is a fixed point of the action of T^m_p, and there is some $r > 0$ such that $\exp_p$ maps the open ball in T_pM with center 0 and radius r T^m_p-equivariantly and diffeomorphically onto the geodesic ball in M with center p and radius r. So it suffices to prove the conclusion for the action of T^m_p on T_pM. The isotropy group of a non-zero $v \in T_pM$ coincides with that of $v/\|v\|$. Thus, it suffices to consider only the isotropy groups of the points on the $(n-1)$-sphere with center 0 in T_pM. Since this sphere is a $(n-1)$-dimensional manifold, the conclusion follows by the inductive hypothesis. □

Proof of Theorem 5.5(a). For $k = 1$ the conclusion is an immediate consequence of Proposition 5.2 and Proposition 5.9 in the Appendix. We assume that (a) holds for $k-1$. Let $F = (f_2, ..., f_k) : M \to \mathbb{R}^{k-1}$ and let U be its set of regular values. If $c = (c_2, ...c_k) \in U$, then $N = F^{-1}(c)$ is a connected, compact submanifold of M, by the inductive hypothesis. A point $p \in N$ is critical for $f_1|N$ if and only if $df_1(p)|T_pN = 0$ or equivalently $df_1(p)$ is a linear combination of $df_2(p)$,..., $df_k(p)$, because

$$T_pN = \bigcap_{j=2}^{k} \mathrm{Ker} df_j(p).$$

Let $a = (a_2, ..., a_k) \in \mathbb{R}^{k-1}$ be such that

$$df_1(p) = a_2 df_2(p) + \cdots a_k df_k(p).$$

Since M is compact, only a finite number of subgroups of T^m occur as isotropy groups of points of M, by Lemma 5.6. Thus, only a finite number of linear subspaces of $\mathfrak{t}_m$ occur as Lie algebras of isotropy groups, that is as sets of elements for which the corresponding fundamental vector field of the action vanishes at a specific point.

Since $f_j \in \mathrm{Im}\rho$, there exists $Y_j \in \mathfrak{t}_m$ such that f_j is the Hamiltonian of $\phi_*(Y_j)$, $1 \leq j \leq k$. If b_1, b_2,..., $b_k \in \mathbb{R}$ and $\sum_{j=1}^{k} b_j Y_j$ belongs to the Lie algebra of the isotropy group of the critical point $p \in N$ of $f_1|N$, that is $(\mu_{*p})^*$, then

$$\sum_{j=1}^{k} b_j df_j(p) = 0$$

and substituting $df_1(p)$ we get

$$\sum_{j=}^{k} (a_j b_1 + b_j) df_j(p) = 0.$$

Since $\{df_2(p), ..., df_k(p)\}$ is linearly independent, $b_j = -a_j b_1$ for $2 \leq j \leq k$ and

$$\sum_{j=1}^{k} b_j Y_j = b_1(Y_1 - a_2 Y_2 - \cdots - a_k Y_k).$$

This shows that $\mathrm{Ker}(\mu_{*p})^* \cap \mathrm{span}\{Y_1, ..., Y_k\}$ is generated by

$$Y_a = Y_1 - a_2 Y_2 - \cdots - a_k Y_k.$$

The set $\mathcal{A}$ of all such $a = (a_2, ..., a_k) \in \mathbb{R}^{k-1}$ is finite and the set of critical points of $f_1|N$ is a disjoint union

$$\bigcup_{a \in \mathcal{A}} N \cap Z_a$$

where $Z_a = \{p \in M : Y_a \in \mathrm{Ker}(\mu_{*p})^*\} = \mathrm{Fix}_{G_{Y_a}}(M)$. By Lemma 5.1 and its proof, each Z_a is a disjoint finite union of connected, compact submanifolds of M and $T_p Z_a = \mathrm{Fix}_{G_{Y_a}}(T_p M)$.

Since $\phi_*(Y_a)$ commutes with $\phi_*(Y_j)$, the proof of Proposition 5.2 shows that $\phi_*(Y_j)(p) \in \mathrm{Fix}_{G_{Y_a}}(T_p M)$, $1 \leq j \leq k$. If now J is a T^m-invariant almost complex structure which is compatible with ω and g is the corresponding T^m-invariant Riemannian metric on M, we have

$$df_j(p) = \omega(\phi_*(Y_j)(p), .) = g(J(\phi_*(Y_j)(p)), .)$$

from which follows that

$$\mathrm{grad} f_j(p) = J(\phi_*(Y_j)(p)) \in T_p Z_a$$

where the gradient is considered with respect to g. This shows that $p \in Z_a$ is a regular point of $F|Z_a$, because grad$f_j(p)$, $2 \leq j \leq k$, are linearly independent. Hence $N \cap Z_a$ is submanifold of codimension $k-1$ in Z_a.

Recall now that T_pN is the g-orthogonal complement to

$$\text{span}\{\text{grad}f_j(p) : 2 \leq j \leq k\}$$

and so it contains all the eigenspaces of the Hessian of the Hamiltonian

$$\psi = f_1 - \sum_{j=2}^{k} a_j f_j$$

of Y_a which correspond to non-zero eigenvalues. So the restriction of the Hessisn to T_pN has even index and degeneracy

$$\text{dim}Z_a - (k-1) = \text{dim}(N \cap Z_a).$$

Since

$$\psi(p) = f_1(p) - \sum_{j=2}^{k} a_j c_j$$

for $p \in N$, the Hessian of ψ restricted on T_pN coincides with the Hessian of $f_1|N$ at p. This implies that $f_1|N$ is a Morse-Bott function and its critical submanifolds have even Morse indices. From Proposition 5.9, the fibers of $f_1|N$ are connected. But these are the fibers of f above the points of $pr^{-1}(c)$, where $pr : \mathbb{R}^k \to \mathbb{R}^{k-1}$ is the projection onto the last $k-1$ coordinates. Hence the fibers of f over the points of the open and dense set $pr^{-1}(U)$ are connected.

Finally, if $L \subset \mathbb{R}^k$ is a line, there exist a vector $v \in \mathbb{R}^k$ and a linear map $\sigma : \mathbb{R}^k \to \mathbb{R}^{k-1}$ such that $L = v + \text{Ker}\sigma$. Applying the inductive hypothesis to $\sigma \circ f$, we have that $(\sigma \circ f)^{-1}(\sigma(v)) = f^{-1}(L)$ is connected. From Proposition 5.9 we get conclusion (a).

Proof of Theorem 5.5(b). To prove that $f(M)$ is convex, let q, $u \in f(M)$ be two different points. They define a unique line L which contains them and $f^{-1}(L)$ is connected, by (a). Therefore, $L \cap f(M) = f(f^{-1}(L))$ is connected and since it contains q and u, it must contain the line segment with these as endpoints. This proves that $f(M)$ is convex.

In order to prove the third assertion of Theorem 5.5, we need the following.

Lemma 5.7. *Let X_1, X_2,..., $X_k \in \mathbb{R}^m$ be vectors which generate a linear subspace S of $\mathbb{R}^m$. If $q(S)$ is dense in T^m, where $q : \mathbb{R}^m \to T^m$ is the quotient map, then there exists a dense set $D \subset \mathbb{R}^k$ such that the set*

$$q(\{t \sum_{j=1}^{k} a_j X_j : t \in \mathbb{R}\})$$

is dense in T^m for every $(a_1, a_2, ..., a_k) \in D$.

Proof. Let $X_j = (X_{j1}, X_{j2}, ..., X_{jm})$, $1 \leq j \leq k$. If there are b_1, b_2,..., $b_m \in \mathbb{Z}$ such that

$$\sum_{i=1}^{m} b_i X_{ji} = 0$$

for every $1 \leq j \leq k$, then $q(S)$ is contained in the subtorus

$$\{(x_1, x_2, ..., x_m) + \mathbb{Z}^m \in T^m : \sum_{i=1}^{m} b_i x_i \in \mathbb{Z}\}$$

of T^m and our assumption implies that it must be all of T^m. In this case we conclude that $b_1 = b_2 = \cdots = b_m = 0$. This shows that if $(b_1, b_2, ..., b_m) \in Z^m \setminus \{(0, 0, ..., 0)\}$, there exists some $1 \leq j \leq k$ such that

$$\sum_{i=1}^{m} b_i X_{ji} \neq 0.$$

and since the vector

$$\Big(\sum_{i=1}^{m} b_i X_{1i}, \sum_{i=1}^{m} b_i X_{2i}, ..., \sum_{i=1}^{m} b_i X_{ki}\Big)$$

is non-zero, the set

$$D(b_1, b_2, ..., b_m) = \{(a_1, a_2, ..., a_k) \in \mathbb{R}^k : \sum_{j=1}^{k} \sum_{i=1}^{m} a_j b_i X_{ji} \neq 0\}$$

is open and dense in $\mathbb{R}^k$. By Baire's theorem,

$$D = \bigcap_{(b_1, b_2, ..., b_m) \in \mathbb{Z}^m} D(b_1, b_2, ..., b_m)$$

is dense in $\mathbb{R}^k$. Obviously, for every $(a_1, a_2, ..., a_k) \in D$, the coordinates of the vector $\sum_{j=1}^{k} a_j X_j$ are linearly independent over $\mathbb{Z}$ and the conclusion follows now from Kronecker's theorem. □

Proof of Theorem 5.5(c). Using the notations of the proof of assertion (a), let

$$K = \overline{\exp(\operatorname{span}\{Y_1, Y_2, ..., Y_k\})}.$$

The set of common critical points of f_1, f_2,...,f_k is $\operatorname{Fix}_K(M)$ and is a disjoint finite union of connected compact submanifolds Z_1, Z_2,..., Z_l of M. Moreover, f_j takes a constant value $c_{ji} \in \mathbb{R}$ on Z_i.

By Lemma 5.7, there exists a dense set $D \subset \mathbb{R}^k$ such that

$$\exp(\{t \sum_{j=1}^{k} a_j Y_j : t \in \mathbb{R}\})$$

is dense in the torus K for every $(a_1, a_2, ..., a_k) \in D$. Now Z is the set of critical points of the smooth function $\sum_{j=1}^{k} a_j f_j$ for every $(a_1, a_2, ..., a_k) \in D$ and so this function must take its extreme values somewhere on Z. In particular, this implies that

$$\sum_{j=1}^{k} a_j f_j(p) \leq \max\{\sum_{j=1}^{k} a_j c_{ji} : 1 \leq i \leq l\}$$

for every $p \in M$ and $(a_1, a_2, ..., a_k) \in D$. By continuity, this holds actually for all $(a_1, a_2, ..., a_k) \in \mathbb{R}^k$.

Suppose now that $f(M)$ is not contained in the convex hull H of the points $c_i = (c_{1i}, c_{2i}, ..., c_{ki})$, $1 \leq i \leq l$ and let $p \in M$ be such that $f(p) \notin H$. Since H is compact, there exists an element $h \in H$ closest to $f(p)$. Let $\langle , \rangle$ denote the euclidean inner product in $\mathbb{R}^k$ and $\| \cdot \|$ the corresponding norm. Then $\|f(p) - h\| > 0$ and

$$\langle a, f(p) \rangle \leq \max\{\langle a, c_i \rangle : 1 \leq i \leq l\}$$

or in another form

$$\langle a, f(p) - h \rangle \leq \max\{\langle a, c_i - h \rangle : 1 \leq i \leq l\}$$

for every $a \in \mathbb{R}^k$. Applying this for the special value $a = f(p) - h$ we get

$$0 < \|f(p) - h\|^2 \leq \langle f(p) - h, c_i - h \rangle$$

for some $1 \leq i \leq l$ and therefore

$$\|f(p) - h\|^2 \langle f(p) - h, c_i - h \rangle \leq \langle f(p) - h, c_i - h \rangle^2 \leq \|f(p) - h\|^2 \cdot \|c_i - h\|^2.$$

Hence

$$0 < \frac{\langle f(p) - h, c_i - h \rangle}{\|c_i - h\|^2} \leq 1.$$

Since H is convex, it must contain the point

$$\left(1 - \frac{\langle f(p) - h, c_i - h \rangle}{\|c_i - h\|^2}\right) h + \frac{\langle f(p) - h, c_i - h \rangle}{\|c_i - h\|^2} c_i$$

and this point cannot be closer to $f(p)$ than h. Thus,

$$\|f(p) - h\|^2 \leq \left\| f(p) - h - \frac{\langle f(p) - h, c_i - h \rangle}{\|c_i - h\|^2} (c_i - h) \right\|^2 = \|f(p) - h\|^2 - \frac{\langle f(p) - h, c_i - h \rangle^2}{\|c_i - h\|^2}.$$

This contradiction shows that $f(M) \subset H$ and therefore $f(M) = H$, because $f(M)$ is convex and contains c_1, c_2,..., c_k.

The proof of Theorem 5.5 is now complete. □

Example 5.8. The action $\phi : T^{n+1} \times \mathbb{C}^{n+1} \to \mathbb{C}^{n+1}$ defined by

$$\phi((t_0, t_1, ..., t_n), (z_0, z_1, ..., z_n)) = (t_0 z_0, t_1 z_1, ..., t_n z_n)$$

is the $(n+1)$-fold cartesian product of the action of the unit circle S^1 in Example 3.3(c). So it is Poisson with momentum map $\mu : \mathbb{C}^{n+1} \to \mathbb{R}^{n+1}$ given by

$$\mu(z_0, z_1, ..., z_n) = \Big(-\frac{1}{2}|z_0|^2, -\frac{1}{2}|z_1|^2, ..., -\frac{1}{2}|z_n|^2\Big).$$

Obviously, the origin is a fixed point and the restricted action on $\mathbb{C}^{n+1}\setminus\{(0,0,...,0)\}$ commutes with scalar multiplication with non-zero complex numbers. Therefore, it induces a smooth action $\tilde{\phi} : T^{n+1} \times \mathbb{C}P^n \to \mathbb{C}P^n$ given by

$$\tilde{\phi}((t_0, t_1, ..., t_n), [z_0, z_1, ..., z_n]) = [t_0 z_0, t_1 z_1, ..., t_n z_n]$$

which is Hamiltonian, hence Poisson, with momentum map $\mu : \mathbb{C}P^n \to \mathbb{R}^{n+1}$ given by

$$\tilde{\mu}[z_0, z_1, ..., z_n] = \Bigg(-\frac{|z_0|^2}{2\sum\limits_{j=0}^{n}|z_j|^2}, -\frac{|z_1|^2}{2\sum\limits_{j=0}^{n}|z_j|^2}, ..., -\frac{|z_n|^2}{2\sum\limits_{j=0}^{n}|z_j|^2}\Bigg).$$

The image of $\tilde{\mu}$ lies on the hyperplane

$$x_0 + x_1 + \cdots + x_n = -\frac{1}{2}$$

of $\mathbb{R}^{n+1}$ and is the n-simplex

$$\Delta = \Big\{(x_0, x_1, ..., x_n) \in \mathbb{R}^{n+1} : x_0 + x_1 + \cdots + x_n = -\frac{1}{2} \quad \text{and} \quad x_j \le 0, \quad 0 \le j \le n\Big\}.$$

The fixed points of the action are the $n+1$ points

$$[1, 0, 0, ..., 0], [0, 1, 0, ..., 0], ..., [0, 0, ...0, 1]$$

and their values under $\tilde{\mu}$ are the vertices of Δ.

Appendix

Let M be a connected, smooth n-manifold. We recall that a smooth function $f : M \to \mathbb{R}$ is called a Morse-Bott function if its set of critical points is a disjoint union of a finite number of compact connected submanifolds C_1, C_2,..., C_k of M such that the Hessian Hess$f(p)$ of f at $p \in C_j$ is non-degenerate on a complement of T_pC_j in T_pM. Each C_j is called a critical submanifold of f. The pullback of TM to C_j decomposes as a Whitney direct sum of vector bundles $TC_j \oplus E_j^+ \oplus E_j^-$, where the fiber $E_j^+(p)$ is spanned by the eigenspaces of the positive eigenvalues of Hess$f(p)$ and $E_j^-(p)$ is spanned by the eigenspaces of its negative eigenvalues. The Morse index of f at $p \in C_j$ is dim$E_j^-(p)$ and is constant along C_j, by continuity. So, we talk of the Morse index of C_j.

Let now M be compact and for $c \in \mathbb{R}$ let $M_c^- = f^{-1}((-\infty, c])$. If $a < b$ are two regular values of f with at most one critical value between them, then

$$H_q(M_b^-, M_a^-; \mathbb{Z}_2) \cong \bigoplus_{i=1}^{m} H_q(E^-(C_i), E_0^-(C_i); \mathbb{Z}_2) \cong \bigoplus_{i=1}^{m} H_{q-l_i}(C_i; \mathbb{Z}_2),$$

where C_1, C_2,..., C_m are the connected components of the critical value between a and b and l_1, l_2,..., l_m are their Morse indices, respectively. Recall that the second isomorphism is just the Thom isomorphism in homology with coefficients in $\mathbb{Z}_2$. The main aim of this Appendix is to prove the next proposition which is usually called the connectivity lemma for Morse-Bott functions.

Proposition 5.9. *Let M be a compact, connected, smooth n-manifold and let $f : M \to \mathbb{R}$ be a Morse-Bott function. If no critical submanifold of f has Morse index 1 or $n-1$, then $f^{-1}(c)$ is connected for every $c \in \mathbb{R}$.*

Proof. We put $M_c^+ = f^{-1}([c, +\infty))$ and $M_c = f^{-1}(c)$. In order to avoid problems of orientability, we use singular homology with coefficients in $\mathbb{Z}_2$ throughout the proof without further mention. Let $a < b$ be two regular values of f with at most one critical value between them. The last part of the long exact sequence of the pair (M_b^-, M_a^-) becomes

$$\cdots \to \bigoplus_{i=1}^{m} H_{1-l_i}(C_i) \to H_0(M_a^-) \to H_0(M_b^-) \to \bigoplus_{i=1}^{m} H_{-l_i}(C_i) \to 0.$$

If M_b^- is connected and M_a^- is not empty, we get $H_{-l_i}(C_i) = 0$ for all $1 \leq i \leq m$. This means that $l_i \geq 1$ and by our assumption $l_i \geq 2$, $1 \leq i \leq m$. Hence also $H_{1-l_i}(C_i) = 0$, $1 \leq i \leq m$ and $H_0(M_a^-) \cong H_0(M_b^-)$, which means that M_a^- is connected as well. If M_a^- is empty, then the critical value between a and b is the global minimum of f and $l_i = 0$ for all $1 \leq i \leq m$. Since f is a Morse-Bott function, it has a finite number of critical values and inductively we conclude that if $a < b$ are any two regular values of f with M_b^- is connected, then M_a^- is also connected. If now we pick b larger than the maximum value of f on M, then $M_b^- = M$, which is assumed to be connected. It follows that M_c^- is connected for every regular value $c \in \mathbb{R}$ of f.

The assumption that f has no critical submanifold of Morse index $n-1$ is equivalent to saying that $-f$ has no critical submanifold of Morse index 1. The Morse index of C_i with respect to $-f$ is the dimension of the fiber of $E^+(C_i)$, that is $n - l_i - \mathrm{dim}C_i$. Arguing as above for $-f$ and the interval $[-b, -a]$, we get the exact sequence

$$\cdots \to H_{n-1}(M_b^+) \to H_{n-1}(M_a^+) \to \bigoplus_{i=1}^{m} H_{l_i + \mathrm{dim}C_i - 1}(C_i) \to \cdots$$

Since $l_i \geq 2$, we have $H_{l_i+\mathrm{dim}C_i-1}(C_i) = 0$. Thus, if $H_{n-1}(M_b^+)$ is trivial, then so is $H_{n-1}(M_a^+)$. Picking again b larger than the maximum value of f on M, we have $M_b^+ = \varnothing$, and inductively as before we conclude that $H_{n-1}(M_c^+) = 0$ for every regular value $c \in \mathbb{R}$.

Let now $c \in \mathbb{R}$ be a regular value of f such that $M_c \neq \varnothing$. Applying the above to $-f$ we have $H_{n-1}(M_c^+) = H_{n-1}(M_c^-) = 0$. Since c is a regular value, we have a Mayer-Vietoris long exact sequence

$$0 \to H_n(M) \to H_{n-1}(M_c) \to H_{n-1}(M_c^+) \oplus H_{n-1}(M_c^-) \to H_{n-1}(M) \to \cdots$$

from which follows that $H_{n-1}(M_c)$ is either trivial or $\mathbb{Z}_2$. But since c is a regular value, M_c is a compact $(n-1)$-manifold and the former possibility is excluded. Hence $H_{n-1}(M_c) \cong \mathbb{Z}_2$, which shows that M_c is connected for every regular value $c \in \mathbb{R}$ of f. Since the set of regular values is open and dense in $\mathbb{R}$, by Sard's theorem, the conclusion follows from the following general topological lemma. □

Lemma 5.10. *Let X be a compact, connected, Hausdorff space and $f : X \to \mathbb{R}^k$ a continuous map with the following properties.*
(i) There exists an open, dense set U in $\mathbb{R}^k$ such that for every $c \in U$ the fiber $f^{-1}(c)$ is connected.
(ii) $f^{-1}(L)$ is connected for every line L in $\mathbb{R}^k$.
Then for for every $c \in \mathbb{R}^k$ the fiber $f^{-1}(c)$ is connected.

Proof. Suppose that $c \in \mathbb{R}^k$ is such that $f^{-1}(c)$ is not connected. Then $c \in \mathbb{R}\setminus U$ and there are disjoint closed sets A, $B \subset X$ with $f^{-1}(c) = A \cup B$. Since X is compact, so are A and B and they have disjoint open neighbourhoods G and H, respectively. Then, $f(X \setminus (G \cup H))$ is a compact subset of $\mathbb{R}^k$, which does not contain c. Let V be an open ball in $\mathbb{R}^k$ with center c which is disjoint from $f(X \setminus (G \cup H))$. Obviously, $f^{-1}(V) \subset G \cup H$ and furthermore

$$f(G) \cap f(H) \cap V \subset \mathbb{R}^k \setminus U,$$

because for every $a \in f(G) \cap f(H) \cap V$ we have $f^{-1}(a) \subset f^{-1}(V) \subset G \cup H$ and $f^{-1}(a) \cap G \neq \varnothing$, $f^{-1}(a) \cap H \neq \varnothing$.

Let now L be any line through c. Then $L \setminus \{0\}$ has two connected components, L^- and L^+. We identify L with $\mathbb{R}$ so that c corresponds to 0, and L^- and L^+ correspond to $(-\infty, 0)$ and $(0, +\infty)$, respectively.

We claim that $f(G) \cap f(H)$ contains $V \cap L^-$ or $V \cap L^+$. If not, there are $a \in V \cap L^-$ and $b \in V \cap L^+$ such that none of them belongs to $f(G) \cap f(H)$. We have to consider two cases. If $a \notin f(G)$ and $b \notin f(G)$, then $f^{-1}(a) \cup f^{-1}(b) \subset H$ and

$$f^{-1}(L) = [f^{-1}(a, b) \cap G] \cup [f^{-1}(-\infty, a) \cup f^{-1}(b, +\infty) \cup (H \cap f^{-1}(L))].$$

If $a \notin f(G)$ and $b \notin f(H)$, then $f^{-1}(a) \subset H$, $f^{-1}(b) \subset G$ and

$$f^{-1}(L) = [f^{-1}(b, +\infty) \cup (f^{-1}(a, +\infty) \cap G)] \cup [f^{-1}(-\infty, a) \cup (f^{-1}(-\infty, b) \cap H)].$$

In both cases we we get that $f^{-1}(L)$ is not connected, contrary to our assumption (ii).

If now $g : \mathbb{R}^k \to \mathbb{R}^k$ is the affine homeomorphism with $g(x) = -x + 2c$, then g fixes c and it is such that $g(L^-) = L^+$ and $g(L^+) = L^-$ for every line L through c. Since $V \cap L \subset (\mathbb{R}^k \setminus U) \cap g(\mathbb{R}^k \setminus U)$ for every line L through c and V is a ball with center c, we must have $V \subset (\mathbb{R}^k \setminus U) \cap g(\mathbb{R}^k \setminus U)$. This is impossible, because U is dense in $\mathbb{R}^k$. □

Remark 5.11. The proof of Proposition 5.9 shows that there exist at most one local minimum and at most one local maximum value of f, which are its extreme values on M.

4.6 The topology of the harmonic oscillator

The harmonic oscillator in $\mathbb{R}^2$ is the physical system which consists of a particle, say of unit mass, acted upon by two springs in orthogonal directions, which we assume to have unit spring constants for simplicity. The phase space of the system is $T^*\mathbb{R}^2$, which is identified with $\mathbb{R}^4$ endowed with the canonical symplectic 2-form

$$\omega = dq^1 \wedge dp_1 + dq^2 \wedge dp_2.$$

The Hamiltonian function of the harmonic oscillator is

$$H(q,p) = \frac{1}{2}\|p\|^2 + \frac{1}{2}\|q\|^2,$$

where $q = (q^1, q^2)$ and $p = (p_1, p_2)$. In other words, it is the sum of the kinetic and the potential energy. The equations of motion are

$$\dot{q} = p, \quad \dot{p} = -q$$

or equivalently the corresponding Hamiltonian vector field is $X_H(q,p) = (p, -q)$.

Since the equations of motion form a system of linear differential equations with constant coefficients, it can be solved explicitly. The flow of X_H is given by the formula

$$\phi_t^H(q,p) = ((\cos t)q + (\sin t)p, (-\sin t)q + (\cos t)p)$$

for $t \in \mathbb{R}$ and $(q,p) \in T^*\mathbb{R}^2$. Despite of this explicit formula we want to have a qualitative description of the flow. Note that the flow induces a free smooth action of S^1 on $T^*\mathbb{R}^2$.

As in Example 2.4(a), the action of the special orthogonal group $SO(2,\mathbb{R})$ on $\mathbb{R}^2$ lifts to a Poisson action on $T^*\mathbb{R}^2$ which leaves the Liouville 1-form invariant and gives rise to the angular momentum $\mu : T^*\mathbb{R}^2 \to \mathbb{R}$ defined by the formula

$$\mu(q,p) = q^1 p_2 - q^2 p_1$$

for $q = (q^1, q^2)$ and $p = (p_1, p_2)$. Since the Hamiltonian H is $SO(2,\mathbb{R})$-invariant, μ is a integral of motion, that is $\{H, \mu\} = 0$. In order to verify that X_H is completely integrable in the sense of section 4.4, it suffices to show that dH and $d\mu$ are linearly independent on a dense open subset $T^*\mathbb{R}^2$. Thus, we need to find the critical points, the critical values and the range of the energy-momentum map

$$E = (H, \mu) : T^*\mathbb{R}^2 \to \mathbb{R}^2.$$

The Jacobian matrix of E at $(q,p) \in T^*\mathbb{R}^2$ is

$$DE(q,p) = \begin{pmatrix} dH(q,p) \\ d\mu(q,p) \end{pmatrix} = \begin{pmatrix} q^1 & q^2 & p_1 & p_2 \\ p_2 & -p_1 & -q^2 & q^1 \end{pmatrix}$$

for $q = (q^1, q^2)$ and $p = (p_1, p_2)$, and $dH(q,p)$, $d\mu(q,p)$ are linearly independent if and only if $DE(q,p)$ has rank 2. Obviously, $DE(q,p)$ has rank 0 if and only

$(q,p)=(0,0)$. Suppose that $DE(q,p)$ has rank 1. Then $dH(q,p)$ and $d\mu(q,p)$ are both non-zero and there exists $a \neq 0$ such that

$$(p_2, -p_1, -q^2, q^1) = (aq^1, aq^2, ap_1, ap_2).$$

This implies that $a^2 = 1$. Consequently, the set of points (q,p) in $T^*\mathbb{R}^2$ at which $DE(q,p)$ has rank 1 consists of the two punctured 2-dimensional linear subspaces

$$P_+ = \{(q^1, q^2, -q^2, q^1) \in T^*\mathbb{R}^2 : (q^1, q^2) \neq (0,0)\},$$
$$P_- = \{(q^1, q^2, q^2, -q^1) \in T^*\mathbb{R}^2 : (q^1, q^2) \neq (0,0)\}.$$

Thus, dH and $d\mu$ are linearly independent on an open dense subset of $T^*\mathbb{R}^2$.

Each energy level set $H^{-1}(c)$, $c \in \mathbb{R}$, is either a 3-sphere of radius $\sqrt{2c}$ for $c > 0$, a singleton for $c = 0$ or the empty set for $c < 0$. Let $c > 0$ and $(q,p) \in H^{-1}(c)$. Then, $(q,p) \neq (0,0)$ and the orbit of (q,p) lies in the intersection of $H^{-1}(c)$ with the 2-dimensional linear subspace of $T^*\mathbb{R}^2 = \mathbb{R}^4$ generated by $\{(q,p),(p,-q)\}$. Thus, each orbit of X_H in $H^{-1}(c)$ is a great circle.

We are going to analyse how the level sets of the restriction of the angular momentum $\mu|_{H^{-1}(c)}$ fit in each energy level set $H^{-1}(c)$ for $c > 0$. The critical values of the energy-momentum map E is the set

$$E(P_+) \cup E(P_-) = \{(c,c) : c > 0\} \cup \{(c,-c) : c > 0\}.$$

Observe that

$$H^{-1}(c) \cap P_+ = \{(q^1, q^2, -q^2, q^1) \in T^*\mathbb{R}^2 : (q^1)^2 + (q^2)^2 = c\},$$
$$H^{-1}(c) \cap P_- = \{(q^1, q^2, q^2, -q^1) \in T^*\mathbb{R}^2 : (q^1)^2 + (q^2)^2 = c\}$$

are great circles which are orbits of X_H and $H^{-1}(c) \cap (P_+ \cup P_-)$ is the set of critical values of $\mu|_{H^{-1}(c)}$. So, $\mu|_{H^{-1}(c)}$ has two critical submanifolds, namely $H^{-1}(c) \cap P_+$ and $H^{-1}(c) \cap P_-$.

Proposition 6.1. *The restriction of the angular momentum $\mu|_{H^{-1}(c)}$ on an energy level set $H^{-1}(c)$ for $c > 0$ is a Morse-Bott function and has two critical great circles on which the Morse indices are 2 and 0, respectively.*

Proof. We shall show that the critical level $H^{-1}(c) \cap P_+$ is non-degenerate of Morse index 2. Let $(q,p) = (q^1, q^2, p_1, p_2) \in H^{-1}(c) \cap P_+$, that is $p_1 = -q^2$ and $p_2 = q^1$. From the above, (q,p) is a critical point of $\mu|_{H^{-1}(c)}$ with Lagrange multiplier 1 and the Hessian of $\mu|_{H^{-1}(c)}$ at (q,p) is

$$D^2(\mu - H)(q,p)|_{\mathrm{Ker}DH(q,p)} = \begin{pmatrix} -1 & 0 & 0 & 1 \\ 0 & -1 & -1 & 0 \\ 0 & -1 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix} |_{\mathrm{Ker}DH(q,p)}.$$

Since $H^{-1}(c) \cap P_+$ is an orbit of X_H, its tangent space at (q,p) is generated by

$$X_H(q,p) = (-q^2, q^1, -q^1, -q^2).$$

Also, $\mathrm{Ker}DH(q,p)$ is generated by the basis consisting of

$$X_H(q,p), \quad v_1 = (q^2, -q^1, -q^1, -q^2), \quad v_2 = (q^1, q^2, q^2, -q^1)$$

and $\{v_1, v_2\}$ is an orthogonal basis of the orthogonal complement of the tangent space of $H^{-1}(c) \cap P_+$ at (q,p) in $\mathrm{Ker}DH(q,p)$. Now a simple computation gives

$$D^2(\mu - H)(q,p)v_1 = -2v_1, \quad D^2(\mu - H)(q,p)v_2 = -2v_2.$$

Hence $H^{-1}(c) \cap P_+$ is a non-degenerate critical level of Morse index 2. A similar computation shows that the critical level $H^{-1}(c) \cap P_-$ is non-degenerate of Morse index 0. □

Corollary 6.2. *For every $c > 0$ the orbit space $H^{-1}(c)/S^1$ of $X_H|_{H^{-1}(c)}$ is diffeomorphic to the 2-sphere S^2.*

Proof. Since the flow of $X_H|_{H^{-1}(c)}$ induces a free smooth action of S^1 on $H^{-1}(c)$, the orbit space $H^{-1}(c)/S^1$ is a smooth 2-dimensional manifold and the quotient map is a fibre bundle projection. According to Proposition 6.1 the angular momentum induces a Morse function $\tilde{\mu} : H^{-1}(c)/S^1 \to \mathbb{R}$ with two critical points, namely $H^{-1}(c) \cap P_+$ of Morse index 2 and $H^{-1}(c) \cap P_-$ of Morse index 0. It follows from Reeb's Theorem that $H^{-1}(c)/S^1$ is homeomorphic to S^2. Since it is 2-dimensional and smooth, it is diffeomorphic to S^2. □

From Proposition 6.1 we conclude thet the angular momentum $\mu|_{H^{-1}(c)}$ takes its maximum value c on $H^{-1}(c) \cap P_+$ and its minimum value $-c$ on $H^{-1}(c) \cap P_-$. Therefore,

$$E(T^*\mathbb{R}^2) = \{(c,d) \in \mathbb{R}^2 : |d| \leq c\}.$$

We shall now determine the topological structure of the level sets

$$(\mu|_{H^{-1}(c)})^{-1}(d) = H^{-1}(c) \cap \mu^{-1}(d)$$

for $|d| \leq c$. We consider the linear automorphism Q of $\mathbb{R}^4$ such that

$$(q^1, q^2, p_1, p_2) = Q(\xi_1, \xi_2, \eta_1, \eta_2),$$

where

$$q^1 = -\frac{1}{\sqrt{2}}\eta_1 - \frac{1}{\sqrt{2}}\eta_2, \quad q^2 = \frac{1}{\sqrt{2}}\xi_1 - \frac{1}{\sqrt{2}}\xi_2,$$

$$p_1 = \frac{1}{\sqrt{2}}\xi_1 + \frac{1}{\sqrt{2}}\xi_2, \quad q^2 = \frac{1}{\sqrt{2}}\eta_1 - \frac{1}{\sqrt{2}}\eta_2$$

or in matrix form

$$\begin{pmatrix} q \\ p \end{pmatrix} = Q \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix},$$

where

$$A = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix}, \quad B = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

Note that $A^tB = AB^t = 0$ and $A^tA + B^tB = I_2$. Therefore,

$$Q^tJQ = \begin{pmatrix} A^t & B^t \\ -B^t & A^t \end{pmatrix} \begin{pmatrix} 0 & -I_2 \\ I_2 & 0 \end{pmatrix} \begin{pmatrix} A & -B \\ B & A \end{pmatrix} = J$$

which means that $Q \in \mathrm{Sp}(2, \mathbb{R})$. Moreover, $Q^tQ = I_4$ and so $Q \in SO(4, \mathbb{R})$. These imply that $X_{H\circ Q} = Q^{-1} \circ X_H \circ Q$ and the Hamiltonian of $X_{H\circ Q}$ is given by the formula

$$(H \circ Q)(\xi, \eta) = \frac{1}{2}\|\xi\|^2 + \frac{1}{2}\|\eta\|^2.$$

The angular momentum $\mu \circ Q$ is also a first integral of $X_{H\circ Q}$ and is given by the formula

$$(\mu \circ Q)(\xi_1, \xi_2, \eta_1, \eta_2) = \frac{1}{2}(\eta_2^2 - \eta_1^2 + \xi_2^2 - \xi_1^2).$$

Thus, the symplectic orthogonal transformation Q diagonalizes simultaneously the quadratic forms of the energy and the angular momentum. Now

$$(E\circ Q)^{-1}(c, d) = \{(\xi_1, \xi_2, \eta_1, \eta_2) \in \mathbb{R}^4 : \xi_1^2+\xi_2^2+\eta_1^2+\eta_2^2 = 2c, \quad \eta_2^2-\eta_1^2+\xi_2^2-\xi_1^2 = 2d\}$$

$$= \{(\xi_1, \xi_2, \eta_1, \eta_2) \in \mathbb{R}^4 : \eta_1^2 + \xi_1^2 = c - d, \quad \eta_2^2 + \xi_2^2 = c + d\}.$$

If $|d| < c$, then (c, d) is a regular value of $E \circ Q$ and we conclude that $E^{-1}(c, d) = Q((E \circ Q)^{-1}(c, d))$ is diffeomorphic to the 2-torus. The rotation number of each orbit of X_H on $E^{-1}(c, d)$ is 1.

Finally, we want to examine the way the level sets $H^{-1}(c) \cap \mu^{-1}(d)$, $|d| \leq c$, of the energy-momentum map fill out the energy level set $H^{-1}(c)$. Summarizing the above, the energy level set $H^{-1}(c)$ for $c > 0$ is diffeomorphic to the 2-sphere S^2 and $H^{-1}(c) \cap \mu^{-1}(d)$ for $|d| < c$ is diffeomorphic to the 2-torus, while $H^{-1}(c) \cap \mu^{-1}(\pm c)$ are great circles.

The action of $SO(2, \mathbb{R})$ on $T^*\mathbb{R}^2$ can be naturally extended to an action of $U(2)$. If $A + iB \in U(2)$, where A, $B \in \mathbb{R}^{2\times 2}$, then

$$I_2 = (A + iB)\overline{(A + iB)}^t = (A + iB)(A^t - iB^t) = AA^t + BB^t + i(BA^t - AB^t)$$

which means that $AA^t + BB^t = I_2$ and $BA^t = AB^t$. Moreover, $(A + iB)(ix) = -Bx + iAx$ for every $x \in \mathbb{R}^2$, so that in particular

$$(A + iB)e_1 = Ae_1 + B(ie_1),$$

$$(A + iB)e_2 = Ae_2 + B(ie_2),$$

$$(A + iB)(ie_1) = -Be_1 + A(ie_1),$$

$$(A + iB)(ie_2) = -Be_2 + A(ie_2).$$

Hence $A + iB \in U(2)$ can be represented a a real 4×4 matrix by

$$\begin{pmatrix} A & -B \\ B & A \end{pmatrix},$$

where $A,\ B \in GL(2,\mathbb{R})$ are such that $AA^t + BB^t = I_2$ and $BA^t = AB^t$. Since $U(2) = \mathrm{Sp}(2,\mathbb{R}) \cap GL(2,\mathbb{C})$, the natural action of $U(2)$ on $\mathbb{R}^4$ is linear symplectic, hence Hamiltonian. For $u \in \mathfrak{u}(2)$ we have

$$u = \begin{pmatrix} A & -B \\ B & A \end{pmatrix}$$

for some $A,\ B \in \mathbb{R}^{2\times 2}$ such that $A = -A^t$, $B = B^t$ and the integral curves of the corresponding fundamental vector field of the action are the solutions of the Hamiltonian system of linear differential equations

$$\begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} = \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix}.$$

The Hamiltonian function is given by the formula

$$\rho(u)(q,p) = \frac{1}{2}\omega(u\begin{pmatrix} q \\ p \end{pmatrix}, \begin{pmatrix} q \\ p \end{pmatrix}) = \langle Aq, p\rangle - \frac{1}{2}\langle q, Bq\rangle - \frac{1}{2}\langle p, Bp\rangle,$$

for $(q,p) \in T^*\mathbb{R}^*$, where $\langle .,. \rangle$ denotes the euclidean inner product in $\mathbb{R}^4$. If u, $v \in \mathfrak{u}(2)$, by the linearity of $X_{\rho(u)}$ and $X_{\rho(v)}$ we have $\{\rho(u), \rho(v)\}(0,0) = 0 = \rho([u,v])(0,0)$. So, the above formula defines a lift $\rho : \mathfrak{u}(2) \to C^\infty(T^*\mathbb{R}^2)$ which is a Lie algebra homomorphism. Therefore, the natural action of $U(2)$ on $T^*\mathbb{R}^2 = \mathbb{R}^4$ is Poisson and gives rise to a momentum map $\ell : T^*\mathbb{R}^2 \to \mathfrak{u}(2)^*$, which is equivariant with respect to the coadjoint action of $U(2)$ on $\mathfrak{u}(2)^*$, according to Theorem 2.2. Since each element of $U(2)$ is also an element of $SO(4,\mathbb{R})$ and the Hamiltonian H is $U(2)$-invariant, it follows that ℓ is an integral of motion of the harmonic oscillator. So, for every $u \in \mathfrak{u}(2)$ the smooth function $\rho(u)$ is a (homogeneous, quadratic) first integral of X_H.

A basis of $\mathfrak{u}(2)$ consists of the elements (considered as 4×4 matrices)

$$u_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

$$u_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad u_4 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix},$$

with corresponding first integrals

$$W_1(q,p) = \rho(u_1)(q,p) = q^1q^2 + p_1p_2,$$

$$W_2(q,p) = \rho(u_2)(q,p) = q^1p_2 - qq^2p_1 = \mu(q,p),$$

$$W_3(q,p) = \rho(u_3)(q,p) = \frac{1}{2}((q^1)^2 - (q^2)^2 + p_1^2 - p_2^2),$$

$$W_4(q,p) = \rho(u_4)(q,p) = \frac{1}{2}((q^1)^2 + (q^2)^2 + p_1^2 + p_2^2) = H(q,p)$$

which satisfy the quadratic relation

$$W_1^2 + W_2^2 + W_3^2 = W_4^2, \quad W_4 \geq 0.$$

We observe that for $c > 0$ the map $W = (W_1, W_2, W_3) : H^{-1}(c) \to S_c^2$, where S_c^2 is the 2-sphere in $\mathbb{R}^3$ of radius c, is essentially the Hopf fibration. Recall that the Hopf fibration is the map $f : S^3 \to S^2$ defined by the formula

$$f(x_1, x_2, y_1, y_2) = (2x_1x_2 + 2y_1y_2, 2x_2y_1 - 2x_1y_2, x_1^2 + y_1^2 - x_2^2 - y_2^2).$$

If now $h : S^3 \to H^{-1}(c)$ and $g : S^2 \to S_c^2$ are the restrictions of the linear diffeomorphisms $h = \sqrt{2c}I_4$ and

$$g = \begin{pmatrix} c & 0 & 0 \\ 0 & -c & 0 \\ 0 & 0 & c \end{pmatrix},$$

then the following diagram commutes.

$$\begin{array}{ccc} S^3 & \xrightarrow{\quad f \quad} & S^2 \\ \Big\downarrow{\scriptstyle h} & & \Big\downarrow{\scriptstyle g} \\ H^{-1}(c) & \xrightarrow{\quad W \quad} & S_c \end{array}$$

Proposition 6.3. *The fibres of W are the great circles in $H^{-1}(c)$.*

Proof. Let $a = (a_1, a_2, a_3) \in S_c^2$. If $(q, p) = (q^1, q^2, p_1, p_2) \in W^{-1}(a)$, then

$$\begin{aligned} q^1q^2 + p_1p_2 &= a_1, \\ q^1p_2 - q^2p_1 &= a_2, \\ (q^1)^2 - (q^2)^2 + p_1^2 - p_2^2 &= 2a_3, \\ (q^1)^2 + (q^2)^2 + p_1^2 + p_2^2 &= 2c. \end{aligned}$$

The first two equations can be written in matrix form

$$\begin{pmatrix} q^1 & p_1 \\ -p_1 & q^1 \end{pmatrix} \begin{pmatrix} q^2 \\ p_2 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix},$$

and from the last two we obtain $(q^1)^2 + p_1^2 = a_3 + c$.

If $(a_1, a_2, a_3) \neq (0, 0, -c)$, then $a_3 + c > 0$ and we can invert to get

$$\begin{pmatrix} q^2 \\ p_2 \end{pmatrix} = \frac{1}{a_3 + c} \begin{pmatrix} q^1 & -p_1 \\ p_1 & q^1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

or equivalently

$$\begin{pmatrix} a_1 & -a_2 - c & -a_2 & 0 \\ a_2 & 0 & a_1 & -a_3 - c \end{pmatrix} \begin{pmatrix} q^1 \\ q^2 \\ p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Therefore $W^{-1}(a)$ is contained in the kernel of

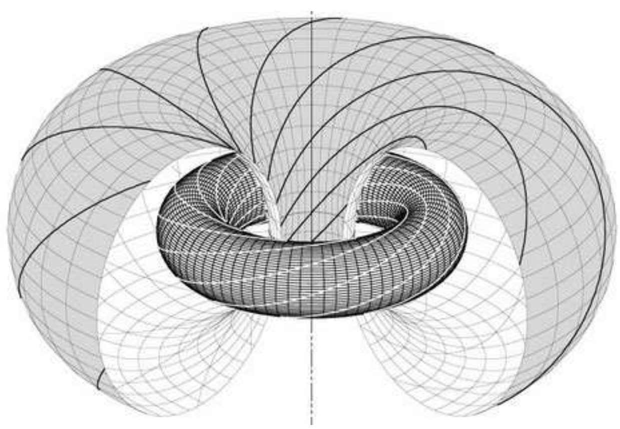
$$\begin{pmatrix} a_1 & -a_2-c & -a_2 & 0 \\ a_2 & 0 & a_1 & -a_3-c \end{pmatrix}$$

which is 2-dimensional, because this matrix has rank 2, since $(a_1, a_2, a_3) \neq (0, 0, -c)$. In case $(a_1, a_2, a_3) = (0, 0, -c)$ we have

$$(q^1)^2 - (q^2)^2 + p_1^2 - p_2^2 = -2c,$$

$$(q^1)^2 + (q^2)^2 + p_1^2 + p_2^2 = 2c$$

and hence $q^1 = p_1 = 0$. Thus, $W^{-1}(0, 0, -c)$ is contained in the 2-dimensional linear subspace $\{(0, q^2, 0, p_2) : q^2, p_2 \in \mathbb{R}\}$ of $\mathbb{R}^4$. Since W is a submersion, $W^{-1}(a)$ is a compact 1-dimensional submanifold of $H^{-1}(c)$ and from the above it is contained in a great circle. It follows that $W^{-1}(a)$ is a great circle in $H^{-1}(c)$. □



Combining the above, each fibre of $W : H^{-1}(c) \to S_c^2$, $c > 0$, coincides with a single (unoriented) orbit of X_H. So, there is a well defined bijective continuous map $\phi : S_c^2 \to H^{-1}(c)/S^1$ which makes the following diagram commutative.

$$\begin{array}{ccc} H^{-1}(c) & \xrightarrow{\text{quotient map}} & H^{-1}(c)/S^1 \approx S^2 \\ {\scriptstyle W}\downarrow & \nearrow_{\phi} & \\ S_c^2 & & \end{array}$$

By compactness, it is a homeomorphism. Moreover, since W is a smooth locally trivial fibre bundle projection and the flow of X_H on $H^{-1}(c)$ has smooth local sections by the Flow Box Theorem (actually the induced free smooth action of S^1 has smooth tubes), the map ϕ is a diffeomorphism. Summarizing, on each energy level set $H^{-1}(c)$, $c > 0$, the flow of X_H is smoothly equivalent to the Hopf flow on S^3.

Bibliography

These notes benefited from some of the books listed below. The list also contains items which are recommended for further reading of material not covered here that is more advanced or specialized. Of course the list is by no means exhaustive.

[1] B. Aebischer, M. Borer, M. Kälin, Ch. Leuenberger, H.M. Reimann, Symplectic Geometry: An introduction based on the seminar in Bern 1992, Progress in Mathematics, vol. 124, Birkhäuser, 1994.

[2] M. Audin, Torus actions on symplectic manifolds, Progress in Mathematics, vol. 93, Second revised edition, Birkhäuser, 2004.

[3] M. Audin, J. Lafontaine, Holomorphic curves in symplectic geometry, Progress in Mathematics, vol. 117, Birkhäuser, 1994.

[4] R. Berndt, An introduction to symplectic geometry, Graduate Studies in Mathematics, vol. 26, American Mathematical Society, 2001.

[5] R.H. Cushman, L.M. Bates, Global aspects of classical integrable systems, Birkhäser, 2015.

[6] H. Hofer, E. Zehnder, Symplectic invariants and Hamiltonian dynamics, Birkhäuser, 1994.

[7] D. McDuff, D. Salamon, Introduction to symplectic topology, Oxford University Press, 1998.

[8] D. McDuff, D. Salamon, J-holomorphic curves and symplectic topology, Colloquium Publications, vol. 52, American Mathematical Society, 2004.

[9] Ana Cannas da Silva, Lectures in Symplectic Geometry, Lecture Notes in Mathematics, vol. 1764, Springer-Verlag, 2001.