The Moment Map

J.W.R

March 21, 2007

Throughout M is a manifold, G is a Lie group and LG is the Lie algebra of G. Denote by $\mathrm{Diff}(M)$ the group of smooth self diffeomorphisms of M, and by $\mathcal{X}(M)$ the Lie algebra of smooth vector fields on M. We denote by $\mathrm{Flow}(X)$ the time one map of the flow of the vector field X. Note that

$$\mathrm{Flow} : \mathcal{X}(M) \to \mathrm{Diff}(M)$$

is formally analogous to the exponential map

$$\exp : LG \to G$$

More precisely if

$$f^t = \mathrm{Flow}(tX)$$

then

$$f^0 = id_M$$

the identity map of M and

$$\left.\frac{d}{dt} f^t\right|_{t=t_0} = X \circ f^{t_0}$$

Given some structure ω on M we denote by $\mathrm{Diff}(M, \omega)$ the subgroup of $\mathrm{Diff}(M)$ consisting of those diffeomorphisms which preserve ω. Similarly, $\mathcal{X}(M, \omega)$ denotes the Lie subalgebra of $\mathcal{X}(M)$ consisting of those vector fields which preserve the structure. For example, if ω is a differential form then

$$\mathrm{Diff}(M, \omega) = \{f \in \mathrm{Diff}(M) : f^*\omega = \omega\}$$

and

$$\mathcal{X}(M, \omega) = \{X \in \mathcal{X}(M) : \ell(X)\omega = 0\}$$

where $\ell(X)\omega$ denotes the Lie derivative of ω in the direction X:

$$\ell(X)\omega = \frac{d}{dt}\mathrm{Flow}(tX)^*\omega\bigg|_{t=0}$$

A Lie group **action** of G on M is a group homomorphism

$$G \to \mathrm{Diff}(M) : a \mapsto a_M$$

for which the evaluation map is smooth. A Lie group action determines (and in case G is connected and simply connected is determined by) a Lie algebra action

$$LG \to \mathcal{X}(M) : A \mapsto A_M$$

via the formula

$$\exp(tA)_M = \mathrm{Flow}(tA_M)$$

for $A \in LG$ and $t \in \mathbb{R}$. We can calculate A_M by the formula

$$A_M = \frac{d}{dt}\bigg|_{t=0} \exp(tA)_M.$$

1 Some formulas

The notations used in the exterior calculus vary slightly from author to author depending on the choice of multiplicative constants. In the following three definitions one can take any value of e and the various laws (viz. $\wedge$ is associative, skew-commutative; $\iota()$ and d are $\wedge$ skew-derivations, etc.) will hold. Most authors take $e = 1$; some take $e = 0$.

For $\alpha \in \mathcal{D}^p(M)$, $\beta \in \mathcal{D}^q(M)$, $X_0, X_1, \dots, X_p \in \mathcal{X}(M)$ we define (with $e = 1$)

$$\alpha \wedge \beta = \binom{p+q}{q}^e \mathrm{ALT}(\alpha \otimes \beta)$$

$$\iota(X_1)\alpha X_2, \dots, X_p = p^{1-e}\alpha(X_1, X_2, \dots, X_p)$$

$$(d\alpha)(X_0, X_1, \dots, X_p) = (p+1)^{e-1}\sum_i (-)^i (D_{X_i}\alpha)(X_0, \dots, \hat{X}_i \dots, X_p)$$

In the last formula D is any covariant derivative (the result is independent of the choice)

Wedge product of one-forms

$$\alpha \wedge \beta(X, Y) = \alpha(X)\beta(Y) - \alpha(Y)\beta(X)$$

Wedge product of a one-form with a two-form

$$\alpha \wedge \beta(X, Y, Z) = \alpha(X)\beta(Y, Z) + \alpha(Y)\beta(Z, X) + \alpha(Z)\beta(X, Y)$$

Cartan's infinitesimal homotopy formula

$$\ell(X)\omega = d\iota(X)\omega + \iota(X)d\omega.$$

Palais's Formula

$$(d\alpha)(X_0, X_1, \ldots, X_p) = \sum_j (-)^j \ell(X_j) \left(\alpha(X_0, \ldots, \hat{X}_j, \ldots, X_p\right)$$

$$+ \sum_{j<k} (-)^{j+k} \alpha(\ell(X_j)X_k, X_0 \ldots, X_j \ldots, X_k \ldots, X_p).$$

Palais's formula for zero-forms

$$d\phi(X) = \ell(X)\phi$$

Palais's formula for one-forms

$$d\theta(X, Y) = \ell(X)(\theta(Y)) - \ell(Y)(\theta(X)) + \theta([X, Y])$$

Palais's formula for two-forms

$$\begin{aligned} d\omega(X, Y, Z) &= \ell(X)(\omega(Y, Z)) + \ell(Y)(\omega(Z, X)) + \ell(Z)(\omega(X, Y)) \\ &\quad -\omega(X, [Y, Z]) - \omega(Y, [Z, X]) - \omega(Z, [X, Y]) \end{aligned}$$

2 Symplectic Mechanics

Let (M, ω) be a **symplectic manifold**: i.e. ω is a symplectic form (=closed, non-degenerate two-form) on M. We call the diffeomorphisms f and vector fields X which preserve ω **symplectic**. As usual $\mathrm{Diff}(M, \omega)$ denotes the group of symplectic diffeomorphisms, and $\mathcal{X}(M, \omega)$ denotes the Lie algebra of symplectic vector fields.

Since $d\omega = 0$ Cartan's formula gives

$$\ell(X)\omega = d\iota(X)\omega$$

so $X \in \mathcal{X}(M, \omega)$ iff $\iota(X)\omega$ is closed. Thus any function H on M determines a vector field $X \in \mathcal{X}(M, \omega)$ via the formula

$$\iota(X)\omega = dH$$

(X is the "symplectic gradient" of H) and when $H^1(M) = 0$ every vector field $X \in \mathcal{X}(M, \omega)$ has this form. We call X the **Hamiltonian vectorfield** and H a **Hamiltonian** for X. We also write

$$X = H_M$$

for the Hamiltonian vector field with Hamiltonian H. Note that X determines H only up to a locally constant function (an additive constant when M is connected.)

The Poisson brackets

$$\mathcal{F}(M) \times \mathcal{F}(M) \to \mathcal{F}(M) : (H, K) \mapsto \{H, K\}$$

defined by

$$\{H, K\} = \omega(H_M, K_M)$$

give $\mathcal{F}(M)$ the structure of a Lie algebra which renders the map $H \mapsto H_M$ a homomorphism of Lie algebras:

$$\{H, K\}_M = [H_M, K_M].$$

In other words, the Poisson brackets $\{H, K\}$ is a Hamiltonian for the Lie brackets $[H_M, K_M]$. This Lie algebra is called the **Poisson algebra** and denoted by $\mathcal{F}(M, \omega)$ so

$$\mathcal{F}(M, \omega) = \mathcal{F}(M)$$

as vector spaces.

The kernel of the homomorphism $H \mapsto H_M$ consists of the locally constant functions and the homomorphism is onto if the first De Rham cohomology

$H^1(M)$ of M vanishes. Hence in case M is connected and simply connected we have an exact sequence

$$0 \to \mathbb{R} \to \mathcal{F}(M,\omega) \to \mathcal{X}(M,\omega) \to 0$$

of Lie Algebras.

According to Darboux's theorem there exist, at any point of M, co-ordinates $q_1, \ldots, q_n, p_1, \ldots, p_n$ such that

$$\omega = \sum_{i=1}^{n} dq_i dp_i.$$

Such co-ordinates are called **symplectic co-ordinates**. In symplectic co-ordinates the Poisson brackets are given by

$$\{H, K\} = \sum_{i=1}^{n} \left(\frac{\partial H}{\partial p_i} \frac{\partial K}{\partial q_i} - \frac{\partial H}{\partial q_i} \frac{\partial K}{\partial p_i} \right)$$

and the trajectories of the vector field H_M are the solutions of **Hamiltonian's equations**:

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}.$$

3 Darboux

Darboux's theorem is easily seen to be equivalent to the following

Theorem 3.1 *Given two symplectic forms ω_0 and ω_1 defined in a neighborhood of 0 in $\mathbb{R}^{2n}$ there is a diffeomorphism f, defined in a possibly smaller neighborhood of 0 with $f(0) = 0$ and*

$$f^* \omega_1 = \omega_0.$$

For the proof first make a linear change of co-ordinates so that ω_1 and ω_0 agree at 0. Then define

$$\omega_t = \omega_0 + t(\omega_1 - \omega_0)$$

for $0 \le t \le 1$. Since that value of ω_t at the origin $0 \in \mathbb{R}^{2n}$ is independent of t (and hence non-degenerate) it follows that all the ω_t are symplectic in a neighborhood of $0 \in \mathbb{R}^{2n}$. We shall find a time-dependent vector field X_t defined near $0 \in \mathbb{R}^{2n}$ so that the curve of diffeomorphism germs $t \mapsto t$ determined by solving the differential equation

$$\frac{d}{dt} f_t = X_t \circ f_t \tag{1}$$

with initial condition

2) $$f_0(x) = x$$

satisfies

$$f_t(0) = 0 \tag{3}$$

and

$$f_t^* \omega_t = \omega_0. \tag{4}$$

Assume for the moment that (1)-(4) hold. Differentiate with respect to t and evaluate at t; we obtain

$$\frac{d}{dt} f_t^* \omega_t = f_t^* \Big(\ell(X_t)\omega_t + \omega_1 - \omega_0 \Big). \tag{5}$$

Hence we require X_t to satisfy

$$\ell(X_t)\omega_t + \omega_1 - \omega_0 = 0. \tag{6}$$

We will also imose the condition

$$X_t(0) = 0. \tag{7}$$

Suppose we have found X_t satisfying (6) and (7). By the existence theorem for ordinbary differential equations we may find f_t satisfying (1) and (2). Equation (6) implies equation (3) so that the diffeomorphisms f_t fix the origin. (This will assure that the change of variables is valid in a neighborhood of the origin.) Equations (5) and (6) give that $f_t^* \omega_t$ is independent of t. Hence, taking $f = f^1$ gives $f^* \omega_1 = \omega_0$ as required. Thus it is enough to construct X_t solving (6) and (7).

By Cartan's formula it suffices to choose X_t to be the unique solution of

$$\iota(X_t)\omega_t = \theta$$

where $d\theta = \omega_1 - \omega_0$. To achieve $X_t(0) = 0$ we require that θ vanishes at 0. This is can be achieved by replacing θ by $\theta + d\phi$ for some function ϕ.

4 Invariance

The formula

$$\ell(H_M)K = \{K, H\}$$

holds on any symplectic manifold (M, ω). Here's the proof:

$$\begin{aligned} \ell(H_M)K &= \iota(H_M)dK \\ &= \iota(H_M)\iota(K_M)\omega \\ &= \omega(K_M, H_M) \\ &= \{K_M, H_M\}. \end{aligned}$$

Thus if $\{K, H\} = 0$, the function K is constant along the trajectories of the vector field H_M.

5 Volume

A symplectic manifold (M, ω) admits a volume

$$\Omega = \omega^n$$

where the dimension of M is $2n$. Obviously, any symplectic diffeomorphism (or vector field) is volume-preserving (that is, it preserves Ω). This is called *Liouville's theorem.* (It is easy to see that a Hamiltonian vector field is volume-preserving, since its divergence is zero in symplectic co-ordinates.)

When M is compact we can use Ω to choose a canonical Hamiltonian H for any Hamiltonian vector field X: we simply impose the condition

$$\int_M H\Omega = 0$$

in addition to the condition $X = H_M$. When M is connected, this condition determines a Hamiltonian H for X uniquely.[1]

For any functions H and K we have

$$\int_M \{H, K\}\Omega = 0$$

[1] If M is not connected, impose this condition on each component.

The proof is by Stokes theorem and

$$\begin{aligned}\{H, K\}\Omega &= (\ell(K_M)H)\Omega \\ &= \ell(K_M)(H\Omega) \\ &= d\iota(K_M)(H\Omega).\end{aligned}$$

Hence the map $X \mapsto H$ determines a Lie algebra splitting of the exact sequence

$$0 \to \mathbb{R} \to \mathcal{F}(M, \omega) \to \mathcal{X}(M, \omega) \to 0.$$

6 Classical mechanics

When $M = \mathbb{R}^n \times \mathbb{R}^n$ and (q, p) are the position and momentum of a particle of mass m subjected to a potential $V = V(q)$, we take

$$H = \frac{1}{2m}\|p\|^2 + V(q)$$

to be the energy of the particle and Hamilton's equations reduce to Newton's equations:

$$\begin{aligned} m\dot{q} &= p \\ \dot{p} &= -grad\, V. \end{aligned}$$

where $grad\, V$ is the gradient of V.

7 Symplectic Group Actions

A **symplectic** Lie group action is a Lie group action where each element of the group G is represented by a symplectic diffeomorphism:

$$G \to \mathrm{Diff}(M, \omega) : a \mapsto a_M.$$

In case G is connected, the action of G is symplectic if and only if corresponding action of its Lie algebra LG is symplectic:

$$A_M \in \mathcal{X}(M, \omega)$$

for $A \in LG$.

We call the action **Hamiltonian** iff each vector field A_M (for $A \in LG$) is Hamiltonian. (This is always the case when M is simply connected.) The action is Hamiltonian iff the Lie algebra homomorphism

$$LG \to \mathcal{X}(M, \omega) : A \mapsto A_M$$

lifts to a linear map

$$LG \to \mathcal{F}(M, \omega) : A \mapsto \mu_A.$$

For each $A \in LG$ the function μ_A is a Hamiltonian for the vector field A_M:

$$(\heartsuit) \qquad d\mu_A = \iota(A_M)\omega.$$

The map

$$\mu : M \to LG^*$$

given by

$$\langle \mu(z), A \rangle = \mu_A(z)$$

is called a **moment map** for the action. Equation $(\heartsuit)$ determines the moment map μ only up to an additive constant which depends on A; i.e. for $\alpha \in LG^*$

$$\tilde{\mu}_A = \mu_A + \alpha(A)$$

is also a Hamiltonian for A_M.

8 Equivariance

Let $\mu : M \to LG^*$ be a moment for a Hamiltonian group action. Recall the co-adjoint action

$$G \to Aut(LG^*) : a \mapsto ad(a^{-1})^*.$$

of the Lie group on the dual of its Lie algebra.

Proposition 8.1 *Assume that G is connected. Then the map*

$$\mu : M \to LG^*$$

is equivariant iff the map

$$LG \to \mathcal{F}(M, \omega) : A \mapsto \mu_A$$

is a homomorphism of Lie algebras:

$$\{\mu_A, \mu_B\} = \mu_{[A,B]}$$

for $A, B \in LG$.

Proof: Since G is connected, equivariance with respect to the Lie group actions is the same as equivariance with respect to the corresponding Lie algebra action. In other words, the map μ is equivariant iff

$$\ell(A_M)\mu = -Ad(A)^*\mu$$

for $A \in LG$. Evaluate both sides at $B \in LG$. The left side becomes

$$\ell(A_M)\mu_B = -\{\mu_A, \mu_B\}$$

(since μ_A and μ_B are Hamiltonians for A_M and B_M) and the right hand side becomes $-\mu_{[A,B]}$.

Now in case M is connected, any two Hamiltonians for the same vector field differ by an additive constant so that each moment map μ determines a skew-symmetric bilinear form $\beta \in \Lambda^2(LG)$ by

$$\{\mu_A, \mu_B\} = \mu_{[A,B]} + \beta(A, B).$$

Now β solves a "co-cycle" identity:

$$d\beta = 0$$

and replacing μ by

$$\tilde{\mu} = \mu + \alpha$$

(where $\alpha \in LG^*$) replaces β by $\tilde{\beta}$ defined by

$$\tilde{\beta} = \beta + d\alpha$$

where $d : \Lambda^2(LG) \to \Lambda^3(LG)$ is defined by

$$d\beta(A, B, C) = \beta([A, B], C) + \beta([B, C], A) + \beta([C, A], B)$$

and $d : \Lambda^1(LG) \to \Lambda^2(LG)$ is defined by

$$d\alpha(A, B) = \alpha([A, B]).$$

If we can solve $d\alpha = \beta$ (for example, if the Lie algebra cohomology $H^2(LG)$ vanishes which is the case when LG is semi-simple) then we can choose μ so that the map $A \mapsto \mu_A$ is a homomorphism of Lie algebras; i.e. so that μ is equivariant.

Of course, when M is compact it is easy to choose an equivariant moment: we impose the condition that

$$\int \mu_A \Omega = 0.$$

(See section 5.)

9 Angular Momentum

Take $M = \mathbb{R}^n \times \mathbb{R}^n$ as before, and $G = SO(n)$ the special orthogonal group:

$$SO(n) = \{a \in \mathbb{R}^{n\times n} : \det(a) = 1,\ a^* = a^{-1}\}.$$

Take the action to be given by:

$$a_M(q,p) = (aq, ap).$$

This action is symplectic. To prove this note that $\omega = d\theta$ where

$$\theta = \sum_{i=1}^{n} p_i dq_i = \langle p, dq\rangle$$

so

$$\begin{aligned} a_M^*\theta &= \langle ap, d(aq)\rangle \\ &= \langle ap, a(dq)\rangle \\ &= \langle p, dq\rangle \\ &= \theta \end{aligned}$$

so $a_M^*\omega = a_M^* d\theta = da_M^*\theta = d\theta = \omega$. The Lie algebra of $SO(n)$ is

$$so(n) = \{A \in \mathbb{R}^{n\times n} : A^* = -A\}$$

A (equivariant) moment map is given by

$$\mu_A(q,p) = \langle p, Aq\rangle.$$

When $n = 3$ there is a bijective correspondence

$$so(3) \to \mathbb{R}^3 : A \mapsto \alpha$$

determined by

$$Av = \alpha \times v$$

for $v \in \mathbb{R}^3$ ($\times$ is the cross product) which identifies μ_A with the **angular momentum** about the axis α.

10 Heisenberg

The action of $\mathbb{R}^2$ on itself by translations is Hamiltonian. To see this let $\omega = dp \wedge dq$ and note that the infinitessimal translations $\partial/\partial p$ and $\partial/\partial q$ are Hamiltonian with respective Hamiltonian's

$$H(q,p) = q + h, \quad K(q,p) = -p + k.$$

For any choice of the additive constants h and k the map $\mu = (H, K)$ is a moment for this action, but

$$\{H, K\} = 1$$

whereas the group is abelian so there can be no equivariant moment.

The Heisenberg group is the set of all real 3×3 matrices of form

$$a = \begin{bmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{bmatrix}.$$

It acts on $M = \mathbb{R}^2$ via the formula

$$a_M(p,q) = (p + x, q + y)$$

and this action is also Hamiltonian. This time however there is a equivariant moment

$$\mu(p,q) = (q, -p, 1)$$

where the triple $\alpha = (\xi, \eta, \zeta) \in \mathbb{R}^3$ is identified with a point in the dual of the Heisenberg algebra via the formula

$$\langle \alpha, A \rangle = \xi\hat{x} + \eta\hat{y} + \zeta\hat{z}$$

for

$$A = \begin{bmatrix} 0 & \hat{x} & \hat{z} \\ 0 & 0 & \hat{y} \\ 0 & 0 & 0 \end{bmatrix}.$$

11 Co-adjoint

A Lie group G acts linearly on its Lie algebra LG via the adjoint action:

$$ad : G \to Aut(LG),$$

$$ad(a)A = aAa^{-1}$$

for $a \in G$ and $A \in LG$. Hence G acts linearly on the dual LG^* of the Lie algebra LG via the co-adjoint action:

$$G \to Aut(LG^*) : a \mapsto ad(a^{-1})^*.$$

Theorem 11.1 *Let $N \subset LG^*$ be an orbit of the co-adjoint action of G. Denote by $a \mapsto a_N$ the restriction of the co-adjoint action to N:*

$$a_N = ad(a^{-1})^*|N$$

for $a \in G$. Then there is a unique G-invariant symplectic form ω on N such that the inclusion

$$\mu : N \to LG^*$$

is an equivariant moment. It is defined by the formula

$$(\diamondsuit) \qquad \omega(A_N(\alpha), B_N(\alpha)) = \langle \alpha, [A, B] \rangle$$

for $\alpha \in N \subset LG^*$ and $A, B \in LG$.

Proof: Before we start the proof, note that the tangent space $T_\alpha N$ to N at $\alpha \in N$ is the vector subspace of LG^* given by

$$T_\alpha N = \{A_N(\alpha) : A \in LG\}.$$

This is because G acts transitively on N.

Now assume ω is G-invariant and that the inclusion μ is a moment for the (symplectic) action of G on the orbit N. We prove uniqueness of ω by proving the formula ($\diamondsuit$). Evidently for $A \in LG$, a Hamiltonian μ_A for A_N is given by

$$\mu_A(\alpha) = \langle \alpha, A \rangle$$

for $\alpha \in N$. Hence

$$\begin{aligned}
\omega(A_N(\alpha), B_N(\alpha)) &= (\iota(A_N)\omega)(B_N(\alpha)) \\
&= d\mu_A(\alpha)B_N(\alpha) \\
&= \quad \langle B_N(\alpha), A\rangle \\
&= -\langle Ad(B)^*\alpha, A\rangle \\
&= -\langle \alpha, Ad(B)A\rangle \\
&= -\langle \alpha, [B, A]\rangle \\
&= \quad \langle \alpha, [A, B]\rangle
\end{aligned}$$

Now we prove existence. Define ω by ($\diamondsuit$). If $A_N(\alpha) = 0$ then the right side of ($\diamondsuit$) vanishes (by part of the last calculation) and similarly if $B_N(\alpha) = 0$. Hence, ω is well-defined. The above computation verifies the formula

$$d\mu_A = \iota(A_N)\omega.$$

The formula

$$Ta_N \circ A_N \circ a_N^{-1} = (ad(a)A)_N$$

(which holds for any action) easily implies the invariance of ω. To see that ω is closed we use Palais's formula:

$$\begin{aligned}
d\omega(A_N, B_N, C_N) &= \ell(A_N)\omega(B_N, C_N) + \ell(B_N)\omega(C_N, A_N) + \ell(C_N)\omega(A_N, B_N) \\
&\quad +\omega(A_N, [B_N, C_N]) + \omega(B_N, [C_N, A_N]) + \omega(C_N, [A_N, B_N]).
\end{aligned}$$

Both the first three and the last three terms on the right vanish by the Jacobi identity.

12 Reduction

Let P be a manifold and ω_P be a closed two-form on P. (Typically P will be a submanifold of a symplectic manifold M and ω_P will be the restriction to P of the symplectic form ω_M on M.) For each point $p \in P$ define a subspace

$$K_p = \{\hat{p} \in T_pP : \iota(\hat{p})\omega_P = 0\}$$

Proposition 12.1 *Assume that $p \mapsto K_p$ is a sub-bundle of TW, i.e. that* $\dim(K_p)$ *is constant. Then the subundle K is integrable.*

Proof: If X is a vector field on P then X is a section of K iff $\iota(X)\omega_P = 0$. When X is a section of K, we have $\ell(X)\omega_P = d\iota(X)\omega_P + \iota(X)d\omega_P = 0$ since ω_P is closed. Hence, any section of K leaves ω_P invariant. Hence, if X and Y are sections of K, then

$$\begin{aligned}\iota([X,Y])\omega_P &= \iota(\ell(Y)X)\omega_P \\ &= \ell(Y)\left(\iota(X)\omega_P\right) - \iota(X)\ell(Y)\omega_P \\ &= 0\end{aligned}$$

as required.

Now assume that the foliation admits a quotient manifold. This means that there is a surjective submersion

$$\pi : P \to B$$

whose fibers are the leaves of the foliation. Thus

$$K_p = T_p\pi^{-1}(\pi(p))$$

for $p \in P$.

Proposition 12.2 *There is a unique two-form ω_B on B such that*

$$\omega_P = \pi^*\omega_B.$$

Moreover, ω_B is symplectic. The symplectic manifold (B, ω_B) is called the **reduction** *of (P, ω_P).*

Proof: To define ω_B at a point $b \in B$ choose a point $p \in P$ with $\pi(p) = b$ and define

$$\omega_B(v, w) = \omega_P(\tilde{v}, \tilde{w}) \tag{$\clubsuit$}$$

for $v, w \in T_bN$ where $\tilde{v}, \tilde{w} \in TzW$ are lifts of v, w:

$$(T_p\pi)\tilde{v} = v, \quad (T_p\pi)\tilde{w} = w.$$

Since the right hand side of ($\clubsuit$) vanishes if either $v = 0$ or $w = 0$ the definition is independent of the choice of the lifts $\tilde{v}, \tilde{w}$. Since the vector fields tangent to the fiber leave ω_P invariant and act transitively on the fiber, the definition is independent of the choice of $p \in \pi^{-1}(b)$. Since $\pi^*d\omega_B = d\omega_P = 0$ and π is a surjective submersion, it follows that $d\omega_B = 0$, i.e. ω_B is closed. Finally, ω_B is non-degenerate, for if $v \in T_bB$ is such that $\omega_B(v, w) = 0$ for all $w \in T_bB$, then $\omega_P(\tilde{v}, \tilde{w}) = 0$ for all $\tilde{w} \in T_pP$, so $\tilde{v} \in K_p$, so $v = 0$.

13 Reduction and Moment

Now assume that $\mu : M \to LG^*$ is an equivariant moment for a sympletic G-manifold and put

$$P = \mu^{-1}(\alpha)$$

where $\alpha \in LG^*$ is a regular value for μ. The tangent space to P at a point $p \in P$ is given by

$$T_p P = \{\hat{p} \in T_p M : d\mu_A(p)\hat{p} = 0 \ \ \forall \ A \in LG\}.$$

This can be rewritten in the form

$$T_p P = \{\hat{p} \in T_p M : \omega(A_M(p), \hat{p}) = 0 \ \ \forall \ A \in LG\}$$

i.e.

$$T_p P = (T_p \mathcal{O})^\perp$$

where $\mathcal{O}$ is the orbit of G through P and $\perp$ denotes “sympletic orthogonal complement”. Thus

$$\begin{aligned} K_p &= (T_p P) \cap (T_p P)^\perp \\ &= (T_p P) \cap (T_p \mathcal{O}) \\ &= \{B_M(p) : d\mu_A(p) B_M(p) = 0 \ \forall A \in LG\} \\ &= \{B_M(p) : Ad(B)^* \alpha = 0\} \\ &= \{B_M(p) : B \in LStab(\alpha, G)\} \end{aligned}$$

where

$$Stab(\alpha, G) = \{a \in G : ad(a)^* \alpha = \alpha\}$$

is the stabilizer (isotropy) group of $\alpha = \mu(p)$. This stabilizer group $Stab(\alpha, G)$ leaves invariant the manifold $P = \mu^{-1}(\alpha)$ since the moment μ is equivariant. If the stabilizer group is connected its orbits are the leaves of the foliation tangent to the sub-bundle K so the reduction $\pi : P \to B$ (if it exists) is the projection of the $Stab(\alpha, G)$-manifold $P = \mu^{-1}(\alpha)$ onto its orbit space.

14 Reduction of the order

Classically, the reduction construction was used to study a Hamiltonian system which admitted some symmetry. By restricting to a level surface

$P = \mu^{-1}(\alpha)$ of the moment and passing to the reduction B one obtains a system on a lower dimensional manifold B. One says the new system is obtained by *reduction of the order*.

Here's how it works. Take a symplectic manifold (M, ω) with a symplectic group action $G \to \mathrm{Diff}(M, \omega) : a \mapsto a_M$ and equivariant moment $\mu : M \to LG^*$. Assume that the Hamiltonian H is invariant by G:

$$a_M^* H = H$$

for $a \in G$. Let $P = \mu^{-1}(\alpha)$ be a regular level of the moment μ, $\omega_P = \omega|P$ the restriction, $B = P/\mathit{Stab}(\alpha, G)$ the orbit space of P by the stabilizer subgroup, $\pi : P \to B$ the projection onto the orbit space, and ω_B the unique symplectic form on B satisfying $\pi^*\omega_B = \omega_P$.

If we differentiate the equation $a_M^* H = H$ we obtain that $\ell(A_M)H = 0$ for $A \in LG$. Hence

$$\ell(H_M)\mu_A = -\{H, \mu_A\} = \ell(A_M)H = 0$$

for all A so $\ell(H_M)\mu = 0$ so the vector field H_M is tangent to the level surfaces of μ. In particular, H_M is tangent to $P = \mu^{-1}(\alpha)$ so we define the vector field $H_P \in \mathcal{X}(P, \omega_P)$ to be the restriction

$$H_P = H_M|P.$$

On the other hand the condition $a_M^* H = H$ for $a \in G$ shows that $H|P$ is invariant by the stabilizer subgroup $\mathit{Stab}(\alpha, G)$. Hence there is a function $K \in \mathcal{F}(B, \omega_B)$ which lifts to $H|P$:

$$\pi^* K = H|P.$$

This K determnes a Hamiltonian vector field K_B on B.

It is easy to see that the projection $\pi : P \to B$ intertwines the vector field H_P on P and the vector field K_B on B:

$$T\pi \circ H_P = K_B.$$

Here's the proof. Choose $b \in B$ and $p \in \pi^{-1}(b)$. We must show that $T\pi H_P(p) = K_B(b)$. For this it is enough to show that $\omega_B(T\pi H_P(p), \hat{b}) =$

$\omega_B(K_B(b), \hat{b})$ for all $\hat{b} \in T_bB$. Choose $\hat{b} \in T_bB$ and $\hat{p} \in T_pP$ with $T\pi\hat{p} = \hat{b}$. Then

$$\begin{aligned}
\omega_B(T\pi H_P(p), \hat{b}) &= \pi^*\omega_B(H_P(p), \hat{p}) \\
&= \omega_P(H_P(p), \hat{p}) \\
&= \omega(H_M(p), \hat{p}) \\
&= dH(p)\hat{p} \\
&= d(\pi^*K(b))\hat{p} \\
&= (\pi^*dK)(b)\hat{p} \\
&= dK(b)\hat{b} \\
&= \omega_B(K_B(b), \hat{b})
\end{aligned}$$

as required.

15 Projective space

For example, take $M = \mathbb{C}^{n+1} \setminus \{0\}$, $G = S^1$ acting diagonally

$$a_M(z) = \exp(i\theta)z$$

for $z \in \mathbb{C}^{n+1}$, $a = \exp(i\theta) \in S^1$, and

$$\omega = -i\sum_{j=0}^{n} dz_j d\bar{z}_j = 2\sum_{j=0}^{n} dy_j dx_j$$

where $z_j = x_j + iy_j$ are the standard co-ordinates on $\mathbb{C}^{n+1}$. Also ω is the imaginary part of the (flat) Hermitean inner product on $\mathbb{C}^{n+1}$:

$$\omega(v, w) = \langle v, iw\rangle - \langle w, iv\rangle$$

for $v, w \in T_zM = \mathbb{C}^{n+1}$ where

$$\langle v, w\rangle = \sum_{j=0}^{n} v_j\bar{w}_j$$

for $v = (v_0, v_1, \ldots, v_n)$, $w = (w_0, w_1, \ldots, w_n)$.

A moment for the action is given by

$$\mu(z) = \|z\|^2.$$

The level surface $P = \mu^{-1}(1)$ is the unit sphere S^{2n+1} and the stabalizer subgroup $Stab(1, G)$ is the whole group S^1 (because the group is abelian) so the orbit space $N = \mathbb{C}P^n$ is complex projective space and the projection $\pi : S^{2n+1} \to \mathbb{C}P^n$ is the Hopf map.

A skew-Hermitean matrix $A \in u(n+1)$ determines a real valued function $H : \mathbb{C}^{n+1} \to \mathbb{R}$ by

$$H(z) = \langle Az, iz \rangle.$$

The corresponding Hamiltonian vector field on $M = \mathbb{C}^{n+1} \setminus \{0\}$ is given by

$$H_M(z) = Az.$$

Since H is invariant by the circle action, we obtain a symplectic vector field A_N on $N = \mathbb{C}P^n$. Thus the usual action of the unitary group $U(n+1)$ on $\mathbb{C}P^n$ is symplectic.

16 Linear actions

If $G \to GL(V)$ is a linear representation of a compact group G on a real vector space V there is a direct sum decomposition

$$V = V_0 \oplus W_1 \oplus \cdots \oplus W_q$$

where V_0 is the fixed point set of the linear representation and $W_1, \ldots, W_q$ are the irreducible components of the representation. By averaging we may find an invariant inner product $\langle \cdot, \cdot \rangle$ on V and take the decomposition to be orthogonal.

Now suppose ω is a non-degenerate skew symmetric form on V invariant by G. Then there is a skew-symmetric automorphism Ω of V with

$$\omega(v, w) = \langle \Omega v, w \rangle$$

for $v, w \in V$. Write $\Omega = PJ$ where P is positive definite and J is orthogonal. Since $\Omega^* = -\Omega$, Ω and Ω^* commute, hence $P = \sqrt{\Omega\Omega^*}$ and $J = P^{-1}\Omega$ commute. By replacing $\langle \cdot, \cdot \rangle$ by $(v, w) \mapsto \langle Pv, w \rangle$ we may assume $\Omega = J$.

Now $J^{-1} = J^* = -J$ so $J^2 = -I$: J is a complex structure on V. The Hermitean inner product

$$(v, w) \mapsto \langle v, w \rangle + i\omega(v, w)$$

is invariant by G. We may decompose V into irreducible complex representations (this might be different from the real decomposition)[2] and obtain a J-invariant orthogonal decomposition

$$V = V_0 \oplus V_1 \oplus \cdots V_p.$$

Now the summands are orthogonal with respect to ω as well as with repect to the inner product. Any infinitessimally unitary endomorphism A of V is Hamiltonian:

$$A = J\, grad\, Q$$

where

$$Q(v) = \frac{1}{2}\langle JAv, v \rangle.$$

Now specialize to the case where G is the torus $T = T^n = \mathbb{R}^n/\mathbb{Z}^n$. In this case the irreducible components $V_1, \ldots, V_p$ are all two dimensional (over $\mathbb{R}$) and the decomposition over $\mathbb{R}$ agrees with the decomposition over $\mathbb{C}$. Any A in the image of LT has form $A = J\, grad\, Q$ where

$$Q = a_1(x_1^2 + y_1^2) + \cdots a_p(x_p^2 + y_p^2)$$

where (x_i, y_i) are suitable symplectic linear co-ordinates on V_i for $i = 1, \ldots, p$. Note in particular that

(1) The index of the quadratic form Q is even.

(2) The restriction of the skew-symmetric form ω to the space V_i is non-degenerate. In particular, $\omega|V_0$ is non-degenerate.

[2] An example where the decomposition into real irreducibles is different from the decomposition into complex irreducibles is given by $O(n)$ acting on $\mathbb{C}^n$. (The skew form ω is the imaginary part of the standard Hermitean inner product on $\mathbb{C}^n$.)

17 Symplectic linearization

The action of a compact group G on a manifold M may be "linearized" near a fixed point $z \in M$: use an equivariant exponential map $T_zM \to M$, say the one arising from the geodesic spray of an invariant Riemannian metric. The equivariant exponential map intertwines the linear action of G on T_zM with the action of G on M. Since the fixed point set of the linearized action of G on T_zM is a vector subspace of T_zM, this proves that the fixed point set of the action of G on M is a submanifold of M.

Now suppose that the action is symplectic. The equivariant exponential map need not be symplectic, but we shall prove that there is a symplectic linearization of the action. What is required is an equivariant Darboux's theorem: if ω_0 and ω_1 are G-invariant symplectic forms which agree at a fixed point z of the G-action, then there is an equivariant diffeomorphism f, defined in an invariant neighborhood of z, satisfying $f(z) = z$ and $f^*\omega_1 = \omega_0$.

To prove this we simply imitate the proof of Darboux's theorem given in section 3 above. First write $\omega_1 - \omega_0 = d\theta$ and choose θ to be invariant by averaging over the group. Then the vector field X_t defined by $\iota(X_t)\omega_t = \theta$ is equivariant with respect to G since ω_t and θ are. Hence the solution f_t commutes with the G-action. In particular, the diffeomorphism $f = f_1$ commutes with the G-action. As it solves $f^*\omega_1 = \omega_0$ it provides the required linearization.

18 Morse inequalities

In their most general form, the Morse inequalities state that

$$\dim H^k(M) \leq \sum_{i=1}^{r} \dim H^k(M_i, M_{i-1})$$

where the M_i form a filtration of the space M:

$$\emptyset = M_0 \subset M_1 \subset M_2 \subset \cdots \subset M_r = M.$$

These inequalities are usually applied in the following situation:

$$M_i = \{x \in M : f(x) \leq a_i\}$$

where $f : M \to \mathbb{R}$ is a smooth function and the a_i are regular values of f and there is exactly one critical value c_i with

$$a_{i-1} < c_i < a_i.$$

We denote by $\Sigma = \Sigma(f)$ the set of critical points of f:

$$\Sigma = \{x \in M : df(x) = 0\}$$

and by Σ_i those at level c_i:

$$\Sigma_i = \{x \in \Sigma : f(x) = c_i\}$$

so that

$$\Sigma = \Sigma_1 \cup \Sigma_2 \cup \cdots \cup \Sigma_r$$

and

$$\Sigma_i \subset \operatorname{int}\left(M_i \setminus M_{i-1}\right).$$

The **critical elements** of f we mean the connected components of the set $\Sigma(f)$ of critical points of f. To ease the exposition we shall assume that each Σ_i connected, so that the critical elements are $\Sigma_1, \Sigma_2, \ldots, \Sigma_r$. This assumption can easily be dropped.

Let x be a critical point f. The Hessian $D^2 f(x)$ is the quadratic form on the tangent space defined by

$$D^2 f(x)\hat{x}^2 = \left.\frac{d^2}{dt^2}\right|_{t=0} f(\gamma(t))$$

where $\hat{x} \in T_x M$ and γ is any curve in M with $\gamma(0) = x$ and $\dot{\gamma}(0) = \hat{x}$. (Since $df(x) = 0$ the definition is independent of the choice of γ.) The dimension of a maximal subspace of $T_x M$ on which the Hessian $D^2 f(x)$ is negative-definite is called the **index** of the critical point x. Thus at any critical point x we have a (non-unique) direct sum decomposition

$$T_x M = N_x^0 \oplus N_x^s \oplus N_x^u$$

orthogonal with respect to the Hessian $D^2 f(x)$ such that $D^2 f(x)$ vanishes on N_x^0, is positive-definite on N_x^s, and negative-definite on N_x^u. The dimension u of N_x^u is the index of the critical point x.

The function f is called **non-degenerate in the sense of Morse** iff each critical point $x \in \Sigma$ is non-degenerate meaning that the Hessian $D^2 f(x)$ is a non-degenerate quadratic form on the tangent space T_xM. It follows in this case that the critical points are isolated so that Σ is a discrete set (finite if M is compact). In this case the critical elements are points. By our assumption that each Σ_i is connected it follows that Σ_i is a single point and M_i/M_{i-1} has the homotopy type of S^u where u is the index. Hence

$$\dim H^k(M_i, M_{i-1}) = \begin{cases} 1 & \text{if } k = u, \\ 0 & \text{otherwise.} \end{cases}$$

We obtain

Proposition 18.1 (Morse inequalities) *For a function f which is non-degenerate in the sense of Morse we have that*

$$\dim H^k(M) \le n_k(f)$$

where $n_k(f)$ is the number of critical points of f with index k.

The function f is called **non-degenerate in the sense of Bott** iff the critical set Σ is a submanifold of M and at each $x \in \Sigma$ the restriction of the Hessian $D^2f(x)$ to the normal space $T_xM^\perp$ is non-degenerate. (This condition is independent of the choice of Riemannian metric used to compute the normal space.) In this case the index of a critical point x is constant for $x \in \Sigma_i$. In this case for each $x \in \Sigma$ the Hessian vanishes on $T_x\Sigma$. In fact the restriction of TM to each critical element Σ_i of Σ admits a (non-unique) vector bundle splitting

$$TM|\Sigma_i = T\Sigma_i \oplus N_i^s \oplus N_i^u$$

orthogonal with respect to the Hessian D^2f such that the Hessian vanishes on $T_x\Sigma_i$, is positive definite on N_i^s and negative definite on N_i^u. The fiber dimension u of N_i^u is the the common index of the critical points $x \in \Sigma_i$. The space M_i/M_{i-1} has the homotopy type of the Thom space $D/\partial D$ where D is the unit disk bundle of N_i^u. Hence, by the Thom isomorphism, we have that

$$\dim H^k(M_i, M_{i-1}) = \dim H^{k-u}(\Sigma_i).$$

In particular, $\dim H^k(M_i, M_{i-1}) = 0$ for $k < u$ and $\dim H^u(M_i, M_{i-1}) = 1$ (by our assumption that Σ_i is connected). We obtain

Proposition 18.2 (Morse-Bott inequalities) *For a function f which is non-degenerate in the sense of Bott we have that*

$$\dim H^k(M) \leq n_k(f)$$

where $n_k(f)$ is the number of critical elements of f with index k.

19 Conected levels

Lemma 19.1 (Atiyah) *Assume that M is compact and connected, that the function $f : M \to \mathbb{R}$ is non-degenerate in the sense of Bott, and that neither f nor $-f$ has a critical element of index 1. The each level $f^{-1}(c)$ for $c \in \mathbb{R}$ is connected (or empty).*

Proof: By continuity we may assume w.l.o.g that c is a regular value. Let

$$M_c^+ = f^{-1}([c, \infty)), \quad M_c^- = f^{-1}((-\infty, c]),$$

so

$$f^{-1}(c) = M_c^+ \cap M_c^-.$$

By Mayer-Vietoris we have an exact sequence

$$\cdots \to H^1(M) \to H^0(f^{-1}(c)) \to H^0(M_c^+) \oplus H^0(M_c^-) \to H^0(M).$$

By the Morse-Bott inequalities we have

$$H^1(M) = 0$$

so it is enough to prove that M_c^- is connected. (To see that M_c^+ is also connected replace f with $-f$.)

Now M_c^- is diffeomorphic to M_k for some $k = 1, 2, \ldots, r$. Consider the exact sequence

$$\cdots \to H^1(M_i, M_{i-1}) \to H^0(M_{i-1}) \to H^0(M_i) \to H^0(M_i, M_{i-1}).$$

Now $H^1(M_i, M_{i-1}) \neq 0$ only if the index of the critical element Σ_i is 1 or 0. The former case is excluded by hypothesis and in the latter case Σ_i is a local minimum. In any case (no matter what the index is) the map $H^1(M_i, M_{i-1}) \to H^0(M_{i-1})$ is zero so $\dim H^0(M_{i-1}) \leq \dim H^0(M_i)$. We have equality for each i since by hypothesis $\dim H^0(M_r) = \dim H^0(M) = 1$. Thus $\dim H^0(M_i) = 1$ for each $i = 1, 2, \ldots, r$ so each M_i is connected as required.

20 Symplectic toral actions

Let (M, ω) be a compact symplectic manifold and $T^n \to \mathrm{Diff}(M, \omega)$ be a Hamiltonian action of the torus $T^n = \mathbb{R}^n/\mathbb{Z}^n$. Using an invariant inner product we identify

$$LT^n = LT^{n*} = \mathbb{R}^n.$$

Let

$$\mu = (f_1, f_2, \ldots, f_n) : M \to \mathbb{R}^n.$$

be an equivariant moment map for the action and

$$f = \mu_A$$

be a linear combination of the components of μ Thus the corresponding Hamiltonian vector field $f_M = A_M$ is an infinitessimal generator of the toral action. Its trajectories have tori as closures: sub-tori of the orbits of the action. (Atiyah calls f_M "almost periodic".)

Lemma 20.1 *Under these hypotheses the function $f : M \to \mathbb{R}$ is non-degenerate in the sense of Bott and has only crtical elements of even index. (Hence, every level set $f^{-1}(c)$ is connected.)*

Proof: The critical points of f are the zeros of the corresponding vector field A_M. With out loss of generality (replace T by a sub-torus) we may assume that the one-parameter group subgroup of T generated by A is dense in T. Now critical points of $f = \mu_A$ are the fixed points of the action. At any fixed point z of the action we may choose symplectic co-ordinates which linearize the action. In such co-ordinates f is a quadratic form

$$f = a_1(x_1^2 + y_1^2) + \cdots a_p(x_p^2 + y_p^2)$$

as in section 16 above which shows the critical point has even index.

21 Symplectic and convex

Theorem 21.1 (Atiyah) *Let M be compact and symplectic and*

$$\mu : M \to LT^* = \mathbb{R}^n$$

be an equivariant moment for a symplectic action of the n dimensional torus $T = T^n$ on M. Let $Z_1, \ldots, Z_r$ be the connected components of the fixed point set of the action. Then

(1) *Each (non-empty) fiber $\mu^{-1}(\alpha)$ ($\alpha \in LT^*$) is connected.*

(2) *The moment μ is constant on each Z_i ($i = 1, \dots, r$).*

(3) *The image $\mu(M)$ is a convex hull of the points $f(Z_1), \dots, f(Z_r)$.*

Proof: We prove part (1) by induction on n. The case $n = 1$ is lemma 20.1. Now assume the theorem for $n - 1$ and write

$$\mu(x) = (\nu(x), f(x))$$

where $\nu : M \to \mathbb{R}^{n-1}$ and $f : M \to \mathbb{R}$. Also put

$$\alpha = (\beta, c)$$

where $\beta \in \mathbb{R}^{n-1}$ and $c \in \mathbb{R}$. It is enough to prove (1) for a dense set of α so by Sard's theorem we may assume that β is a regular value of ν. Form the manifold

$$N = \nu^{-1}(\beta).$$

By the induction hypothesis N is connected. Since the torus is abelian, the stabalizer is everything: $Stab(\beta, T^{n-1}) = T^{n-1}$.

As a warm-up we make the (unwarranted) assumption that the quotient $Q = N/T^{n-1}$ is a manifold. We'll see how to eliminate this assumption at the end of this section. Form the reduction $\pi : N \to Q$ where $Q = N/T^{n-1}$. The moment μ is invariant by the action of T^n (since the torus is abelian) so the function f is invariant by the action of T^{n-1} so there is a function $g : Q \to \mathbb{R}$ with $f|N = g \circ \pi$. Evidentally $\mu^{-1}(\alpha) = \pi^{-1}(g^{-1}(c))$. The fiber of π is T^{n-1} which is connected and the level set $g^{-1}(c) \subset N$ is connected by lemma 20.1. Hence $\mu^{-1}(\alpha)$ is connected as required.

For part (2) note that the fixed point set of the action is precisely the set of points $z \in M$ where $d\mu_A(z) = 0$ for every $A \in LT$: in other words it is the set of points z where $d\mu(z) = 0$. As this set is a manifold it is clear that μ is constant on each of its components.

For part (3) we first show that the image $\mu(M)$ is convex, i.e. that it intersects every line L in $\mathbb{R}^n = LT^*$ in a segment. Choose a linear projection $\pi : \mathbb{R}^n \to \mathbb{R}^{n-1}$ and a point $\beta \in \mathbb{R}^{n-1}$ so that $L = \pi^{-1}(\beta)$. Apply part (1) to $\pi \circ \mu$. Then $(\pi \circ \mu)^{-1}(\beta)$ is connected so its image

$$\mu\Big((\pi \circ \mu)^{-1}(\beta)\Big) = \mu(M) \cap \pi^{-1}(\beta)$$

is also connected as required.

For almost every $A \in \mathbb{R}^n = LT$ the one-parameter group $t \mapsto \exp(tA)$ is dense in T and so the trajectories of the corresponding vector field A_M are dense in the orbits of T. This means that the zeros of A_M are precisly the fixed points of the T-action. But the zeros of A_M are the critical points of its Hamiltonian μ_A. Hence μ_A achieves its maximum on one of the components Z_i of the fixed point set. This means that the image $\mu(M)$ lies to one side of the hyperplane of all $\alpha \in LT^*$ such that

$$\langle \alpha, A\rangle = \langle \mu(Z_i), A\rangle.$$

Since this holds for almost every A it follows that the image $\mu(M)$ is a subset of the convex hull of the points $\mu(Z_i) \in \mathbb{R}^n = LT^*$. But we have already seen that the image is convex so it must equal this convex hull.

Here's how to get rid of the unwaranted assumption that Q is a manifold. We work directly with $N = \nu^{-1}(\beta)$. Since

$$\mu^{-1}(\alpha) = (f|N)^{-1}(c)$$

it is enough to show that $f|N$ satisfies the hypothesis of Lemma 19.1, namely that $f|N$ is non-degenerate in the sense of Bott and the critical elements of $\pm f|N$ have even index. For this we choose a critical point $x \in N$ of $\pm f|N$. By the method of Lagrange multipliers x is a critical point of the function

$$\phi = f + \lambda \circ \nu : M \to \mathbb{R}$$

for some choice of the linear functional $\lambda \in \mathbb{R}^{(n-1)*}$. Now ϕ is a linear combination of the components of the moment map μ so by Lemma 20.1 its critical set Σ (that is, the fixed point set of its Hamiltonian flow) is a submanifold of M and $D^2 f(x) : T_xM \times T_xM \to \mathbb{R}$ has even index. We will show

22 Eigenvalue inequalities

Theorem 22.1 (Schur-Horn) *A vector $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ is the diagonal of a Hermitean matrix whose eigenvalues are the components of the vector $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ if and only if x is a convex combination of the vectors $\sigma_* \lambda$ for $\sigma \in S_n$. Here S_n is the permutation group on $\{1, 2, \ldots n\}$ and $\sigma_* \lambda = (\lambda_{\sigma(1)}, \lambda_{\sigma(2)}, \ldots, \lambda_{\sigma(n)})$.*

The theorem may be reformulated as follows. Let D_n denote the set of all real diagonal matrices. Let $\Lambda \in D_n$ and let

$$\mathcal{O}(\Lambda) = \{a\Lambda a^{-1} : a \in U(n)\}$$

denote the orbit of Λ under unitary similarity. For any square $n \times n$ matrix A, let $\delta(A)$ denote the $n \times n$ diagonal matrix having the same diagonal as A. Then the Schur-Horn theorem takes the form

$$\delta(\mathcal{O}(\Lambda)) = \text{ convex hull of } \mathcal{O}(\Lambda) \cap D_n.$$

Let $G = U(n)$ be the unitary group and $LG = u(n)$ be its Lie algebra. Thus $u(n)$ is the set of all matrices iA where A is Hermitean. Notice that $LT = iD_n$ is the Lie algebra of the maximal torus $T = \exp(iD_n)$ in $U(n)$ consisting of all diagonal unitary matrices. The Lie algebra $u(n)$ admits an invariant inner product

$$\langle A, B \rangle = \text{ real part of } tr(AB^*)$$

and with respect to this inner product the map

$$\delta : u(n) \to iD_n$$

which assigns to each matrix $A \in u(n)$ the diagonal matrix having the same diagonal entries as A is orthogonal projection onto iD_n. Thus the Schur-Horn theorem is a special case of the following

Theorem 22.2 (Kostant) *Let G be a compact Lie group, LG be its Lie algebra, T be a maximal torus in G, and LT be its Lie Algebra. Let LG be endowed with a G-invariant inner product. Then the orthogonal projection of any orbit of G in LG onto LT is the convex hull of the intersection of that orbit with LT.*

Remark 22.3 The orthogonal projection $LG \to LT$ is independent of the choice of the invariant inner product. To see this write

$$G = G_1 \times G_2 \times \cdots \times G_k \times T_0$$

where each G_i is compact with simple Lie algebra and T_0 is a torus. Then

$$T = T_1 \times T_2 \times \cdots \times T_k \times T_0$$

where T_i is a maximal torus in G_i for $i = 1, 2, \ldots, k$. Now consider the corresponding decompostion of the Lie algebra:

$$LG = LG_1 \oplus LG_2 \oplus \cdots \oplus LG_k \oplus LT_0.$$

Each G_i acts irreducibly on LG_i (else LG_i would not be simple) so by Schur's lemma this direct sum decomposition must be orthogonal with respect to any invariant inner product $\langle \cdot, \cdot \rangle$ on LG. Thus the invariant inner product is given by

$$\langle A, B \rangle = \langle A_1, B_1 \rangle_1 + \langle A_2, B_2 \rangle_2 + \cdots \langle A_k, B_k \rangle_k + \langle A_0, B_0 \rangle_0$$

where $\langle \cdot, \cdot \rangle_i$ is a G_i-invariant inner product on LG_i for $i = 1, 2, \ldots, k$ and $\langle \cdot, \cdot \rangle_0$ is a T_0-invariant inner product on LT_0. But the only invariant inner product on a simple Lie algebra is a multiple of the Killing form, so the orthogonal complement is given by

$$LT^\perp = LT_1^\perp \oplus LT_2^\perp \oplus \cdots \oplus LT_k^\perp$$

where $LT_i^\perp$ is the sum of the root spaces of LG_i. (This argument also shows that to prove Kostant's theorem it is enough to prove it when LG is simple.)

Proof of Kostant's theorem: Using the G-invariant inner product we identify LG and LG^*. The orbit N is a symplectic G-manifold and the inclusion $N \subset LG$ is an equivariant moment. Restrict the action to T: then the inclusion followed by the orthogonal projection onto LT is an equivariant moment. The fixed points of the action are the points $\alpha \in N$ at which

$$\langle \alpha, [A, B] \rangle = 0$$

for all $A \in LT$ and all $B \in LG$. In other words α is a fixed point for the T-action exactly when $[A, \alpha] = 0$ for all $A \in LT$ and this the case exactly when $\alpha \in LT$.

23 References

[AM] R. Abraham, J. Marsden: *Foundations of Mechanics, (Second edition)*, Benjamin-Cummings, 1978.

[A] M.F. Atiyah: Convexity and commuting Hamiltonians, *Bull. London Math. Soc.* **14** (1982) 1-15.

[**B**] R. Bott: Non-degenerate critical manifolds, *Ann. of Math.* **60** (1954) 248-261.

[**H**] A. Horn: Doubly stochastic matrices and the diagonal of a rotation matrix, *Amer. J. Math.*, **76** (1954), 620-630.

[**Ki**] F.C. Kirwan: *Cohomology of quotients in symplectic and algebraic geometry*, Princeton Mathematical notes, **31** (1984).

[**Ko**] B. Kostant: On convexity, the Weyl group, and the Iwasawa decompositionm, *Ann. Sci. Ec. Norm. Sup,*, **6** (1973), 413-455.

[**M**] J. Milnor: *Morse theory*, Annals of Math Studies **51**, Princeton University Press (1963).

[**MW**] J. Marsden, A. Weinstein: Reduction of symplectic manifolds with symmetry, *Rep. Math. Phys.*, **5** (1974) 121-130.

[**S**] I. Schur: Uber eine klasse von mittelbildungen mit anwendungen auf der determinantentheorie, Sitzungsberichte der Berliner Mathematischen Gesellschaft **22** (1923) 9-20.