THE MOMENT MAP IN SYMPLECTIC GEOMETRY

PHILIP BOALCH

1. Introduction

Symplectic geometry has its origins in the study of phase spaces in Hamiltonian mechanics, where the points of such spaces are described by position and momentum coordinates (q, p). If $f(q, p)$ is a smooth function on the phase space we obtain an associated vector field:

$$X_f = \sum_i \frac{\partial f}{\partial p_i}\frac{\partial}{\partial q_i} - \frac{\partial f}{\partial q_i}\frac{\partial}{\partial p_i}$$

Thus a curve $(q(t), p(t))$ is an integral of the flow of X_f iff

$$\frac{dq_i}{dt} = \frac{\partial f}{\partial p_i} \qquad \text{and} \qquad \frac{dp_i}{dt} = -\frac{\partial f}{\partial q_i} \text{ for each } i$$

If $f = H$ is the Hamiltonian of some mechanical system then these are Hamilton's equation of motion, determining the evolution of the system. By differentiating, it can be seen that $H(q(t), p(t))$ is constant on any trajectory; i.e. H is a constant of the motion.

The primary example of a moment (or momentum) map is the angular momentum of a particle of unit mass moving freely in 3 dimensions. Here $H = \frac{1}{2}\sum_1^3 p_i^2$ so $X_H = \sum_1^3 p_i \frac{\partial}{\partial q_i}$ and the trajectories are straight lines $(q(t), p(t)) = (a + tb, b)$. The angular momentum

$$\mu(q, p) = q \times p$$

is constant on the trajectories. If $f(q, p) = (q \times p) \cdot v$ is the component of the angular momentum in some fixed direction v then the associated vector field X_f generates a rotation about v and these rotations make up a group of symmetries of the system.

A second example, whose importance will come to light in later chapters, is the system of n independent harmonic oscillators. Here $H = \frac{1}{2}\sum_i (q_i^2 + p_i^2)$ and the trajectories are the products of circles about the origin in each (q_i, p_i)-plane. The moment map turns out to be given by the parts of the Hamiltonian:

$$\mu(q, p) = \frac{1}{2}(q_1^2 + p_1^2, \ldots, q_n^2 + p_n^2)$$

In this case the flow generated by a component of the moment map actually is a trajectory.

Most of Chapter 2 will be devoted to elucidating what we mean by 'symplectic geometry' and how the concept of a moment map is abstracted, by describing the key properties we want a general moment map to enjoy.

The point of this is that there has been much recent interest in symplectic geometry (where the symplectic manifolds studied arise in many guises; we are no longer restricted to phase spaces) and the moment map is frequently a useful tool. I've included a section of references at the end of this essay pointing out many of the directions that we are leading towards, although most of these will not be used here.

The contents of this essay are as follows:

As mentioned above Chapter 2 provides a brisk general introduction to symplectic geometry and incorporates the general results we'll subsequently use. In fact most of the results in this chapter will be seen in use in the rest of the essay.

Chapter 3 gives a proof of the Atiyah,Guillemin-Sternberg Convexity Theorem (see [A],[GS] and [GS1], although we follow Atiyah most closely). This result basically says that the image of a moment map (obtained when the symmetry group is a torus and the space is compact) is in fact a convex polyhedron. Apart from being attractive in its own right, this Theorem provides foundation for much else and the methods used in the proof are quite widely applicable.

The vertices of such a polyhedron turn out to be the image under the moment map of the fixed points of the symmetry group. In Chapter 4 we show how, in a neighbourhood of a fixed point, the group acts similarly to the harmonic oscillator example above, and that this action is in some sense determined by the geometry of the polyhedron at the corresponding vertex. We then go on to show that if the polyhedron is of a particular type then the 'harmonic oscillator actions' obtained from each vertex can be pieced together in a natural way to construct a manifold on which a torus acts and admits a moment map whose image is the convex polyhedron we started with.

In [D], Delzant proves that the space obtained in this way is in fact uniquely determined by the polyhedron up to a diffeomorphism that preserves the symplectic structure. Once we have this construction there are many interesting questions to ask, for instance how does the space change if we smoothly vary the faces of the polyhedron? It is closely related to the theory of Toric Varieties (see e.g. [Da] or [F]) and changing the number of vertices corresponds in some sense to the algebro-geometric notion of 'blowing-up' at a singular point. The polyhedron also contains information about the 'quantisation' of the corresponding symplectic manifold (see [G], where Guillemin proceeds to use it to illustrate many exciting results). Unfortunately there isn't space or time to delve into any of these questions here and I can only point you towards the references.

As for the intended level, some effort has been made to make this essay accessible to anyone who went to the Part III Differential Geometry course.

2. Definitions and Preliminary Results

2.1. Symplectic and Complex Vector Spaces.

A *symplectic vector space* is a real vector space V endowed with a non-degenerate anti-symmetric bilinear form, i.e. it has a bilinear form ω such that for all $x, y \in V$ $\omega(x,y) = -\omega(x,y)$ and, if $\omega(x,y) = 0$ for all $y \in V$, then $x = 0$. From elementary linear algebra we can check that V is then necessarily of even dimension. The non-degeneracy gives an isomorphism

$$x \mapsto \omega(\cdot, x) = -i_x\omega \tag{1}$$

between V and its dual $V^* = Hom(V, \mathbb{R})$, where i_x is interior multiplication by x. The reason for the minus sign will seen in Proposition 2.3

A *complex structure* on a real vector space V is a linear map $\Im : V \to V$ such that $\Im^2 = -id_V$. Again V is necessarily even dimensional. A complex structure allows us to regard V as a complex vector space with scalar multiplication defined by the rule: $(a + ib)x = ax + b\Im(x)$. Conversely a complex vector space can be regarded as a real vector space with a complex structure defined by $\Im(x) = ix$.

The following result will be useful later:

Proposition 2.1. *Given a symplectic vector space (V, ω) and any inner product $\langle\ ,\ \rangle$ on V we can canonically associate a complex structure $\Im$ and inner product $g(\ ,\ )$ with V such that $\Im$ preserves ω and g and they are 'calibrated', i.e for all x,y in V:*

$$\omega(\Im x, \Im y) = \omega(x,y),\ \ g(\Im x, \Im y) = g(x,y) \text{ and } \omega(x,y) = g(\Im x, y)$$

Proof. Since ω and $\langle\ ,\ \rangle$ are non-degenerate we can define an invertible linear map $A : V \to V$ by $\langle Ax, y\rangle = \omega(x,y)$. Let * denote the adjoint (transpose) operation with respect to $\langle\ ,\ \rangle$ so that A^*A is self-adjoint and positive definite (as $\langle A^*Ax, x\rangle = \langle Ax, Ax\rangle > 0$ if $x \neq 0$). This implies V decomposes canonically into eigenspaces of A^*A with positive (real) eigenvalues and we can obtain a well defined self-adjoint square root P of A^*A by taking the positive (scalar) square roots in each eigenspace.

Claim that $\Im = AP^{-1}$ does the trick: Note that we necessarily have $g(x,y) = \omega(x, \Im y) = \langle Ax, \Im y\rangle$. Two observations make the checking simple: 1) the anti-symmetry of ω implies that $A^* = -A$ and 2) A commutes with $\Im$ (it is straight forward to check that A preserves the eigenspaces of A^*A so commutes with P^{-1} and therefore also with $\Im = AP^{-1}$). It's now easy to verify what's required:

Firstly

$$\Im^2 = AP^{-1}AP^{-1} = A^2P^{-2} = -A^*A(A^*A)^{-1} = -id_V$$

so we do have a complex structure and note also that $\mathfrak{I}^* = (AP^{-1})^* = P^{-1}(-A) = -\mathfrak{I}$ and so $\mathfrak{I}^*\mathfrak{I} = id_V$.

$\mathfrak{I}$ preserves ω and g since:

$$\omega(\mathfrak{I}x, \mathfrak{I}y) = \langle A\mathfrak{I}x, \mathfrak{I}y\rangle = \langle \mathfrak{I}Ax, \mathfrak{I}y\rangle = \langle \mathfrak{I}^*\mathfrak{I}Ax, y\rangle = \langle Ax, y\rangle = \omega(x, y), \text{ and}$$

$$g(\mathfrak{I}x, \mathfrak{I}y) = \langle A\mathfrak{I}x, \mathfrak{I}^2 y\rangle = \langle \mathfrak{I}Ax, -y\rangle = \langle Ax, -\mathfrak{I}^* y\rangle = g(x, y).$$

g is an inner product because:

$$g(y, x) = \langle Ay, \mathfrak{I}x\rangle = \langle \mathfrak{I}x, Ay\rangle = \langle Ax, \mathfrak{I}y\rangle = g(x, y), \text{ and}$$

$$g(x, x) = \langle Ax, \mathfrak{I}x\rangle = \langle x, A^*AP^{-1}x\rangle = \langle x, Px\rangle > 0 \text{ if } x \neq 0$$

since P is self-adjoint and positive definite. Finally they are calibrated as:

$$\omega(x, y) = \langle Ax, y\rangle = \langle Ax, -\mathfrak{I}^2 y\rangle = -g(x, \mathfrak{I}y) = g(\mathfrak{I}x, y) \quad \square$$

2.2. Symplectic Manifolds.

A *symplectic manifold* is a differentiable manifold, M endowed with a closed, non-degenerate 2-form, i.e. a 2-form ω such that $\mathrm{d}\omega = 0$ and at each point m of M the tangent space $T_m M$ is symplectic with form ω_m. The simplest example of a symplectic manifold is $\mathbb{C}^n$ with the constant symplectic form

$$\omega = \sum_{i=1}^{n} dy_i \wedge dx_i \tag{2}$$

where $(x_1 + iy_1, \ldots, x_n + ix_n)$ are coordinates on $\mathbb{C}^n$. We'll see later (Proposition 2.18) that every point in any symplectic manifold has a coordinate neighbourhood in which ω has this form.

A diffeomorphism $\theta : M \to M$ is said to be a *symplectomorphism* if the induced map $\theta^* = \delta\theta$ on forms preserves the symplectic form ω. This notion extends to maps between symplectic manifolds in the obvious way.

The non-degeneracy of ω at each point gives an isomorphism between vector fields and 1-forms on M using the isomorphism (1) between $T_m M$ and $T_m^* M$ at each point. In particular given a smooth real valued function $f \in C^\infty(M)$ we obtain a vector field X_f corresponding to the 1-form df, so that

$$df = -i_{X_f}\omega$$

X_f is called the *Hamiltonian vector field* or *symplectic gradient* of f. Thus X is a Hamiltonian vector field iff $i_X\omega$ is an exact 1-form and then X determines f up to the addition a function that's constant on each connected component of M. Observe that the points of M where f is critical (i.e. where $df = 0$) are precisely where the vector field X_f vanishes.

From any vector field we can obtain a (local) flow $\Phi(t,m)$ defined on a neighbourhood of $0\times M\subset\mathbb{R}\times M$.

Proposition 2.2. *The flow of a Hamiltonian vector field is a symplectomorphism wherever it is defined.*

Proof. Suppose the function $f\in C^\infty(M)$ generates the flow $\Phi(t,m)$ defined on a neighbourhood of $0\times M\subset\mathbb{R}\times M$. Using Cartan's formula that as maps between forms

$$\mathcal{L}_X = i_X\circ d + d\circ i_X \qquad \text{(where } \mathcal{L} \text{ is the Lie derivative)} \tag{3}$$

we see

$$\mathcal{L}_{X_f}\omega = i_{X_f}(d\omega) + d(i_{X_f}\omega) = 0 - d(df) = 0 \tag{4}$$

The result now follows from fairly standard properties of the Lie derivative. We'll give the argument for completeness.

We want to show that $\Phi_t^*\omega=\omega$, i.e. for each $m\in M$ that $\Phi_t^*\omega_{\Phi_t(m)}=\omega_m$ (where $\Phi_t(m)=\Phi(t,m)$). Fix $m\in M$ then $\Phi_t^*\omega_{\Phi_t(m)}$ is an element of the vector space $\Lambda^2(T_m^*M)$ which we wish to show is independent of t (clearly its value at $t=0$ is ω_m). Observe:

$$\begin{aligned}
\frac{d}{dt}\Big(\Phi_t^*\omega_{\Phi_t(m)}\Big)\Big|_{t=s} &= \lim_{h\to 0}\Big(\frac{\Phi_{s+h}^*\omega_{\Phi_{s+h}(m)} - \Phi_s^*\omega_{\Phi_s(m)}}{h}\Big) \\
&= \Phi_s^*\Big(\lim_{h\to 0}\Big(\frac{\Phi_h^*\omega_{\Phi_{s+h}(m)} - \omega_{\Phi_s(m)}}{h}\Big)\Big) && \text{(as } \Phi_s^* \text{ continuous)} \\
&= \Phi_s^*((\mathcal{L}_{X_f}\omega)_{\Phi_s(m)}) && \text{(by definition of } \mathcal{L}) \\
&= 0 && \text{(by (4))} \quad \square
\end{aligned}$$

Thus if the flow of a Hamiltonian vector field is complete we obtain an action of $\mathbb{R}$ on M by symplectomorphisms.

The *Poisson bracket* on a symplectic manifold is the bilinear map

$$\{\ ,\ \}: C^\infty(M)\times C^\infty(M)\to C^\infty(M);\quad (f,g)\mapsto \omega(X_f,X_g)$$

Note that $\{f,g\} = (-i_{X_g}\omega)(X_f) = dg\cdot X_f = X_f(g) = \mathcal{L}_{X_f}g$.

The sign in the isomorphism (1) was chosen so that:

Proposition 2.3. $[X_f,X_g]$ *is the Hamiltonian vector field of* $\{f,g\}$

Proof. We use the identity that, as maps between forms

$$i_{[X,Y]} = \mathcal{L}_X\circ i_Y - i_Y\circ\mathcal{L}_X \tag{5}$$

(this and Cartan's formula (3) can be proved by checking them on functions and 1-forms and then seeing that both sides of each act the same on products). Combining them we get:

$$i_{[X,Y]} = di_X i_Y + i_X di_Y - i_Y di_X - i_Y i_X d \tag{6}$$

so that since df, dg and ω are closed we have

$$i_{[X_f,X_g]}\omega = di_{X_f} i_{X_g}\omega = d(\omega(X_g, X_f)) = -d\{f,g\} \quad \square$$

The Poisson bracket in fact endows $C^\infty(M)$ with a Lie algebra structure. To establish this we only need check the Jacobi identity:

$$\begin{aligned}
\{f,\{g,h\}\} &= \mathcal{L}_{X_f}(\omega(X_g, X_h)) \\
&= \omega([X_f, X_g], X_h) + \omega(X_g, [X_f, X_h]) \qquad \text{(by Leibniz and (4))} \\
&= \{\{f,g\},h\} + \{g,\{f,h\}\} \qquad \text{(by Proposition 2.3)}
\end{aligned}$$

Thus Proposition 2.3 says precisely that the map $f \mapsto X_f$ is a Lie algebra homomorphism.

Given a symplectomorphism $\theta : M \to M$ we can compare the induced maps on $C^\infty(M)$ and $Vect(M)$ to obtain:

Proposition 2.4. *If θ is a symplectomorphism of (M,ω) then*

$$d\theta(X_f) = X_{f\circ\theta^{-1}}$$

Proof. We manipulate the definitions as follows:

$$\begin{aligned}
d(f\circ\theta) = \theta^*(df) &= \theta^*(\omega(\cdot, X_f)) \\
&= \omega(\theta_*(\cdot), X_f) \\
&= (\theta^*\omega)(\cdot, \theta_*^{-1}X_f) \\
&= \omega(\cdot, \theta_*^{-1}X_f) \qquad (\theta \text{ preserves } \omega)
\end{aligned}$$

the result follows by replacing θ by θ^{-1} $\square$

2.3. Lie group actions.

Suppose a Lie group G acts on a manifold M, that is we have a smooth map

$$\Phi : G \times M \to M$$

such that $\Phi(gh, m) = \Phi(g, \Phi(h, m))$ and $\Phi(e, m) = m$ for all $g, h \in G$ and $m \in M$ where $e = id_G$. The notation $\Phi_g(m) = \Phi_m(g) = g \cdot m = \Phi(g, m)$ will be used if there is no risk of ambiguity.

Given $x \in \mathfrak{g} = T_eG$ (the Lie algebra of G), let X denote the corresponding left invariant vector field on G. Thus $X_e = x$ and at $g \in G$

$$X_g = (dL_g)_e(x)$$

where $L_g : G \to G$ is left multiplication by g. Left invariance just means that for all $g \in G$, $(dL_g)(X) = X$, and follows directly from the chain rule.

An action Φ of G on M allows us to define the *fundamental vector field* $\widetilde{X}$ associated to $x \in \mathfrak{g}$:

$$\widetilde{X}_m = (d\varphi_m)_e(x)$$

where $\varphi_m : G \to M$, associated to Φ, maps $g \mapsto g^{-1} \cdot m$. This definition ensures that the map $x \mapsto \widetilde{X}$ is a Lie algebra homomorphism; which we'll deduce in Corollary 2.7. Note that we have

Proposition 2.5. *The flow of $\widetilde{X}$ is given by $\exp(-tx) \cdot m$, where $\exp(tx)$ is the 1-parameter subgroup of G generated by $x \in \mathfrak{g}$*

Proof. By definition $g \exp(tx)$ is the flow of the vector field X on G, through the point $g \in G$. This implies that $\frac{d}{dt}\exp(tx)|_{t=0} = x$ and so it follows from the chain rule that $\varphi_m \circ \exp(\cdot x)$ has derivative $(d\varphi_m)_e(x) = \widetilde{X}_m$ at $t = 0$. The result follows as $\varphi_m \circ \exp(tx) = \exp(-tx) \cdot m$. $\square$

G can act on itself by conjugation: $c_g(h) = ghg^{-1}$. Clearly $c_g(e) = e$ so its derivative at the identity, $(dc_g)_e$, is an endomorphism of $T_eG = \mathfrak{g}$. This is the adjoint action of G on $\mathfrak{g}$ which we'll denote by Ad_g. Dualising we get the coadjoint action Ad^*_g of G on the vector space dual $\mathfrak{g}^*$ of $\mathfrak{g}$ determined by:

$$\langle Ad^*_g \xi, x\rangle = \langle \xi, Ad_{g^{-1}} x\rangle$$

for all $\xi \in \mathfrak{g}^*$ and $x \in \mathfrak{g}$ where $\langle\ ,\ \rangle$ is the canonical pairing between $\mathfrak{g}$ and $\mathfrak{g}^*$. Inverting g ensures we do get an action. Part of the significance of this action lies in the following

Proposition 2.6. *If Φ is an action of G on a manifold M then the map*

$$\mathfrak{g} \to Vect(M); \quad x \mapsto \widetilde{X}$$

intertwines the adjoint action of G on $\mathfrak{g}$ and the induced action of G on $Vect(M)$ given by $d\Phi_g$, i.e:

$$\widetilde{Ad_g(x)} = (d\Phi_g)(\widetilde{X})$$

Proof. We just calculate:

$$\begin{aligned}
(\widetilde{Ad_g(x)})_m &= (d\varphi_m)_e(Ad_g(x)) && \text{(definition of } \widetilde{\ }) \\
&= (d(\varphi_m \circ c_g))_e(x) && \text{(chain rule)} \\
&= (d(\Phi_g \circ \varphi_{g^{-1}m}))_e(x) && (\text{as } \varphi_m \circ c_g = \Phi_g \circ \varphi_{g^{-1}m}) \\
&= (d\Phi_g)_{g^{-1}m}\widetilde{X}_{g^{-1}m} && \square
\end{aligned}$$

Corollary 2.7. *In the situation of Proposition 2.6, for any $x, y \in \mathfrak{g}$:*

$$[x, y] = \frac{d}{dt}\Big(Ad_{\exp(tx)}(y)\Big)\Big|_{t=0} \tag{7}$$

$$[\widetilde{X}, \widetilde{Y}] = \widetilde{[x, y]} \tag{8}$$

where, as usual, $\exp(tx)$ *denotes the 1-parameter subgroup of G generated by $x \in \mathfrak{g}$.*

Proof. At any point $m \in M$ we have:

$$\begin{aligned}
[\widetilde{X}, \widetilde{Y}]_m &= (\mathcal{L}_{\widetilde{X}}\widetilde{Y})_m \\
&= \frac{d}{dt}\Big(d\Phi_{\exp(tx)}\widetilde{Y}\Big)_m\Big|_{t=0} && (2.5 \text{ gives the flow of } \widetilde{X}) \\
&= \frac{d}{dt}\Big((d\varphi_m)_e Ad_{\exp(tx)}(y)\Big)\Big|_{t=0} && (\text{by } 2.6) \\
&= \left(\left(\frac{d}{dt}\Big(Ad_{\exp(tx)}(y)\Big)\Big|_{t=0}\right)^{\sim}\right)_m
\end{aligned} \tag{9}$$

Now if we firstly take Φ to be the action of G on itself given by

$$\Phi(g, h) = hg^{-1} = R_{g^{-1}}(h)$$

where R denotes right multiplication then $\widetilde{X}_h = X_h$ at all points $h \in G$ since here we have $\varphi_h = L_h$. In this case (9) at the point $e \in G$ is precisely (7).

For a general group action substituting (7) into the last part of (9) yields (8). $\square$

Note that by definition $[x, y] = [X, Y]_e$ and that the linear map $ad_x : \mathfrak{g} \to \mathfrak{g}$ defined by $ad_x(y) = [x, y]$ gives the so-called adjoint representation of $\mathfrak{g}$ on $\mathfrak{g}$.

2.4. Lie group actions on symplectic manifolds.

If Φ is an action of a Lie group G on a symplectic manifold (M, ω) then Φ is said to be a *symplectic group action* if each of the induced diffeomorphisms of M are actually symplectomorphisms, that is, if $\Phi_g^*\omega = \omega$ for all $g \in G$.

Proposition 2.8. *If G acts symplectically on (M, ω) then for any $x \in \mathfrak{g}$ the 1-form $i_{\widetilde{X}}\omega$ is closed.*

Proof. We just calculate:

$$\begin{aligned} d(i_{\widetilde{X}}\omega) &= i_{\widetilde{X}}d\omega + di_{\widetilde{X}}\omega && \text{(since } d\omega = 0) \\ &= \mathcal{L}_{\widetilde{X}}\omega && \text{(Cartan's formula)} \\ &= \lim_{t\to 0}\frac{1}{t}(\Phi^*_{\exp(tx)}\omega - \omega) \\ &= 0 && \text{(as } G \text{ acts symplectically)} \quad \square \end{aligned}$$

Under certain mild extra conditions we can ensure $\widetilde{X}$ is actually Hamiltonian. In particular we have

Proposition 2.9. *If G acts symplectically on (M,ω) and if also either*

(i) The first de Rham cohomology group of M, $H^1_{dR}(M)$, is zero, or

(ii) Each element of $\mathfrak{g}$ can be written as a finite linear combination of Lie brackets of other elements (i.e. $\mathfrak{g} = [\mathfrak{g},\mathfrak{g}]$)
then $\widetilde{X}$ is Hamiltonian for each $x \in \mathfrak{g}$.

Note that (i) is a topological condition on M since we have the de Rham theorem (that the de Rham cohomology is isomorphic to the real Cech cohomology) and also note that (ii) holds if G is semi-simple.

Proof. If (i) holds then every closed 1-form is exact, so Proposition 2.8 implies $\widetilde{X}$ is Hamiltonian.

Suppose now that (ii) holds. Proposition 2.8 enables us to emulate the proof of Proposition 2.3 to establish that, for all $x, y \in \mathfrak{g}$:

$$i_{[\widetilde{X},\widetilde{Y}]}\omega = d(i_{\widetilde{X}}i_{\widetilde{Y}}\omega)$$

By linearity and the fact that the map $x \mapsto \widetilde{X}$ is a Lie algebra homomorphism (equation (8)) we see that any element of $[\mathfrak{g},\mathfrak{g}]$ maps to a Hamiltonian vector field. $\square$

The group action is said to be Hamiltonian if $\widetilde{X}$ is Hamiltonian for each $x \in \mathfrak{g}$. In this case, for example by choosing a basis of $\mathfrak{g}$, we can get a map

$$\mu : M \to \mathfrak{g}^* \text{ such that } \langle \mu, x\rangle \in C^\infty(M) \text{ has Hamiltonian vector field } \widetilde{X} \tag{10}$$

where $\langle\ ,\ \rangle$ is the canonical pairing between $\mathfrak{g}$ and $\mathfrak{g}^*$ so that $\langle \mu, x\rangle$ is the component of μ along $x \in \mathfrak{g}$. In 1890 Sophus Lie observed that in this situation, using our notation we have

Proposition 2.10. *For each connected component of M there is a map $\theta : G \to \mathfrak{g}^*$ such that on that connected component, for each $g \in G$*

$$Ad^*_g \circ \mu = \mu \circ \Phi_g + \theta(g)$$

Proof. Pick any $x \in \mathfrak{g}$. We can obtain two functions that have Hamiltonian vector field $d\Phi_{g^{-1}}\widetilde{X}$. Firstly by Proposition 2.6

$$d\Phi_{g^{-1}}\widetilde{X} = \widetilde{Ad_{g^{-1}}(x)}$$

which, by (10), is the Hamiltonian vector field of $\langle \mu, Ad_{g^{-1}}(x)\rangle$. Also

$$\begin{aligned} d\Phi_{g^{-1}}\widetilde{X} &= d\Phi_{g^{-1}}(X_{\langle \mu, x\rangle}) \quad &\text{(by (10))} \\ &= X_{\langle \mu\circ\Phi_g, x\rangle} \quad &\text{(by Proposition 2.4)} \end{aligned}$$

Thus $\langle Ad_g^* \circ \mu, x\rangle$ and $\langle \mu \circ \Phi_g, x\rangle$ differ by a locally constant function on M and the result follows. □

Note that, by a straight-forward calculation, such θ will satisfy

$$\theta(gh) = \theta(g) + Ad_g^*\theta(h) \tag{11}$$

in which case θ is said to be a cocycle for the coadjoint representation in terms of Lie algebra cohomology. At last we come to

Definition 2.11. *A moment map for a symplectic action of a Lie group G on a symplectic manifold M is a smooth map*

$$\mu : M \to \mathfrak{g}^*$$

such that

(i) $i_{\widetilde{X}}\omega = -d\langle \mu, x\rangle$ for all $x \in \mathfrak{g}$, and

(ii) μ intertwines the action of G on M and the coadjoint action of G on $\mathfrak{g}^*$

Remarks:

- (i) is precisely (10)
- (ii) just means that each θ in Proposition 2.10 is zero. If it holds we'll say that μ is *equivariant.* It follows that orbits of G in M are mapped by μ onto coadjoint orbits in $\mathfrak{g}^*$.
- The associated map $\widetilde{\mu} : \mathfrak{g} \to C^\infty(M)$; $x \mapsto \langle \mu, x\rangle$ is known as the comomentum map. Often condition (ii) is replaced by the condition

(ii)′ $\widetilde{\mu} : \mathfrak{g} \to C^\infty(M)$ is a Lie algebra homomorphism.

which is really just the infinitesimal version of (ii), since the brackets are derivatives of the corresponding actions. More precisely we can prove

Proposition 2.12. *If μ is equivariant then $\widetilde{\mu}$ is a Lie algebra homomorphism and the converse holds if G is connected.*

Proof. For the first part we just unwind the definitions. For all $y, z \in \mathfrak{g}$ and $m \in M$:

$$\{\widetilde{\mu}(y), \widetilde{\mu}(z)\}(m) = (X_{\widetilde{\mu}(y)})_m(\widetilde{\mu}(z)) = \widetilde{Y}_m(\widetilde{\mu}(z))$$

$$= \frac{d}{dt}\Big(\widetilde{\mu}(z)(\exp(-ty)\cdot m)\Big)\Big|_{t=0} \qquad \text{(it's the Lie derivative)}$$

$$= \frac{d}{dt}\langle \mu(\exp(-ty)\cdot m), z\rangle\Big|_{t=0} \tag{12}$$

$$= \frac{d}{dt}\langle Ad^*_{\exp(-ty)}(\mu(m)), z\rangle\Big|_{t=0} \qquad \text{(by condition (ii))} \tag{13}$$

$$= \frac{d}{dt}\langle \mu(m), Ad_{\exp(ty)}z\rangle\Big|_{t=0}$$

$$= \langle \mu(m), [y,z]\rangle \qquad \text{(by (7) of Corollary 2.7)}$$

$$= \widetilde{\mu}([y,z])(m)$$

Conversely if $\widetilde{\mu}$ is a Lie algebra homomorphism we have that lines (12) and (13) are equal, which is equivalent to $\frac{d}{dt}\theta(\exp(tx))\big|_{t=0} = 0$ for all $x \in \mathfrak{g}$ and each cocycle θ as defined in Proposition 2.10. By using the cocycle condition (11) we see this implies

$$\frac{d}{dt}\theta(\exp(tx))\Big|_{t=s} = ad^*_x(\theta(\exp(sx))).$$

for any $s \in \mathbb{R}$ where ad^*_x is the dual linear map to ad_x. This is a linear first order ordinary differential equation for $\theta(\exp(\cdot x)) : \mathbb{R} \to \mathfrak{g}^*$ with initial condition $\theta(e) = 0$ and so the solution is identically zero. Thus θ is zero on the image of exp which is dense in G if G is connected. □

Much of the usefulness of moment maps stems from the observation that for all $v \in T_mM$ and $x \in \mathfrak{g}$

$$\langle (d\mu)_m(v), x\rangle = (d\langle \mu, x\rangle)_m(v) \qquad \text{(as } \langle\cdot, x\rangle \text{ is linear)} \tag{14}$$

$$= -(i_{\widetilde{X}}\omega)_m \cdot v$$

$$= \omega_m(v, \widetilde{X}_m)$$

In particular this is used in the proof of

Theorem 2.13. *Suppose a compact Lie group G acts symplectically on (M, ω) and admits a moment map μ. Let N be the subset of M that μ maps to 0 (i.e. $N = \mu^{-1}(0)$). If G acts freely on N (i.e. only the identity fixes any point of N) then*

(i) N is a submanifold of M

(ii) The space of orbits in N, N/G, has the structure of a symplectic manifold with induced symplectic form $\overline{\omega}$ determined by

$$i^*\omega = p^*\overline{\omega}$$

where $i : N \to M$ is the inclusion and $p : N \to N/G$ maps each point onto its orbit.

In this situation the orbit space N/G is called the *symplectic quotient* or *Marsden-Weinstein reduced space.* We won't go through the proof here; it was given in the Part III Differential Geometry Course.

We will also need the following result later

Proposition 2.14. *Suppose in the situation of the previous Theorem, another group H acts symplectically on M, admitting a moment map ν. If the action of H commutes with that of G then H acts symplectically on N/G and admits a moment map $\nu_{N/G}$ such that*

$$\nu_{N/G} \circ p = \nu \circ i$$

2.5. **Incorporating a Riemannian Metric.**

It is a well known result that any differentiable manifold can be given a Riemannian metric (i.e. an inner product on each tangent space that varies smoothly over the manifold). In this section we'll see how a metric can be made to fit into the picture.

If (M, ω) is a symplectic manifold that also has a Riemannian metric $\langle , \rangle$ then we can apply Proposition 2.1 to each tangent space $T_m M$ to obtain a calibrated complex structure $\Im_m$ and inner product $g_m(,)$ at each point. The whole point about the rather long-winded proof of Proposition 2.1 was that in the current situation it becomes clear that $\Im_m$ and $g_m(,)$ are chosen in such a way that they vary smoothly with m. Such a $\Im$ (that varies smoothly over M and gives a complex structure on each tangent space) is called an *almost-complex structure* on the manifold M. Thus we have:

Proposition 2.15. *If $(M, \omega, \langle , \rangle)$ is a symplectic manifold with Riemannian metric then we can find a Riemannian metric g and an almost-complex structure $\Im$ such that $\Im$ preserves g and ω and they are calibrated (i.e. the conditions in 2.1 hold at each point)*

We also have the following

Proposition 2.16. *If a Lie group G acts symplectically on (M, ω) and M admits a G-invariant Riemannian metric $\langle , \rangle$, then after applying Proposition 2.15, g is G-invariant and $\Im$ commutes with the action of G.*

Proof. Pick any $h \in G$. The invariance of ω and $\langle , \rangle$ implies that the induced action of Φ_h between tangent spaces intertwines the maps A_m and $A_{h\cdot m}$ (where we are using the notation of 2.1). The result now follows from the definitions of $\mathfrak{I}$ and g in 2.1. □

Note that if G is compact then it is possible, by integration, to 'smooth out' any given metric to make it G-invariant.

On a Riemannian manifold the usual gradient vector field , ∇f, of a function $f \in C^\infty(M)$ is defined by the isomorphism between 1-forms and vector fields induced from the metric, i.e.

$$df = g(\nabla f, \cdot)$$

Thus in the current situation, a simple calculation shows that ∇f is related to the symplectic gradient X_f as:

$$X_f = \mathfrak{I}\nabla f$$

In particular X_f and ∇f are orthogonal ($\mathfrak{I}$ rotates vectors into their orthogonal space generally, since $g(x, \mathfrak{I}x) = g(\mathfrak{I}x, -x) = -g(x, \mathfrak{I}x)$). This ties in with the fact that the flow of ∇f is in the direction of maximum change of f and X_f flows along the levels of f.

A Riemannian metric also brings to M a notion of geodesics. Given a tangent vector $x \in T_mM$ there is a unique geodesic curve, $a : (-\epsilon, \epsilon) \to M$ through m, defined for some positive ϵ, whose derivative at m is x. Thus if x is short enough we can map it to the point of M that is distance $|x|$ along the geodesic it defines (the 'distance' on M and the norm $|\ |$ on T_mM are obtained from the metric). This mapping in fact defines a diffeomorphism between a neighbourhood U of $m \in M$ and W of $0 \in T_mM$. So-called 'normal coordinates' about m are obtained by using an orthonormal basis to provide coordinates on T_mM and using this diffeomorphism. See, for example, [KN] p162 for more precise details.

A very nice property of normal coordinates is the following:

Proposition 2.17. *In the situation of the preceding paragraph, if a diffeomorphism θ preserves the metric and fixes the point $m \in M$ then, where defined, the action of θ on U corresponds via normal coordinates to the linear action induced by θ on $W \subset T_mM$.*

Proof. The result follows directly from the fact that θ maps geodesics to geodesics (which is true since it preserves the metric). If $a : (-\epsilon, \epsilon) \to M$ is the geodesic through m and some point $n \in U$ (corresponding to the tangent vector $x = \frac{da}{dt}|_{t=0} = (da)_0(\frac{d}{dt}|_{t=0}) \in T_mM$) then $\theta \circ a$ is the geodesic through m and $\theta(n)$.

By the chain rule it corresponds to

$$((d\theta)_m \circ (da)_0)\Big(\frac{d}{dt}\Big|_{t=0}\Big) = (d\theta)_m(x) \quad \square$$

We will also need the following basic result of symplectic geometry:

Theorem 2.18. *(The Darboux Theorem) Suppose a manifold M is endowed with two symplectic forms ω_1 and ω_2. If the two forms coincide at some point $m \in M$ then there is an open neighbourhood U containing m and a diffeomorphism*

$$\psi : U \to U$$

such that $\psi^*\omega_1 = \omega_0$ throughout U.

The proof is neither long nor very hard but it is not really very enlightening so has been omitted. See e.g. [Au] p39 or [HZ] p10. If there a Lie group acting that preserves both of the symplectic forms then in fact ψ can be chosen to commute with the group action.

It follows directly from this Theorem that any point of any symplectic manifold has a coordinate neighbourhood with respect to which the symplectic form ω is as in the simple example (2): We just use normal coordinates at the point; by linear algebra the constant symplectic form on the tangent space (defined to agree with the lift of ω at the origin) takes the simple form with respect to some basis. An application of the Darboux Theorem adjusts the coordinates so that the two forms on the tangent space (the constant form and the lift of ω) agree throughout a neighbourhood of the origin. c.f. Section 4.2 as well.

3. The Atiyah,Guillemin-Sternberg Convexity Theorem

In this chapter we'll prove the following result:

Theorem 3.1. *Suppose a torus T acts symplectically on a compact connected symplectic manifold (M, ω) and admits a moment map $\mu : M \to \mathfrak{t}^*$ (where $\mathfrak{t}$ is the Lie algebra of T). Then for $n = 1, 2, 3, \ldots$:*

If $f_1, \ldots, f_n \in C^\infty(M)$ are some components of μ (i.e. each for each $i, f_i = \langle \mu, x_i \rangle$ for some $x_i \in \mathfrak{t}$). Then $F = (f_1, \ldots, f_n) : M \to \mathbb{R}^n$ satisfies:

(A_n) : $F^{-1}(c)$ *is connected or empty for each* $c \in \mathbb{R}^n$

(B_n) : $F(M)$ *is convex*

Moreover if $Z_1, \ldots, Z_m$ are the connected components of the set of common critical points of all the f_i's then $F(Z_j) = c_j$ is a single point and $F(M)$ is the convex hull of the c_j's

Note that the statement in the introduction follows by choosing the components of μ with respect to a basis of $\mathfrak{t}^*$. It is phrased in this way here since we will prove it by induction on n.

Proof. The first thing to see is that B_{n+1} follows from A_n:

$F(M)$ is convex iff its intersection with any line l in $\mathbb{R}^{n+1}$ is connected or empty. If $\pi : \mathbb{R}^{n+1} \to \mathbb{R}^n$ is any linear projection then the components of πF satisfy the conditions for A_n since each of its components are linear combinations of $f_1, \ldots, f_{n+1}$. Choose π such that $l = \pi^{-1}(c)$ for some $c \in \mathbb{R}^n$, then by A_n, $(\pi F)^{-1}(c)$ is connected or empty. Now

$$l \cap F(M) = \pi^{-1}(c) \cap F(M) = \{F(m) | m \in M \text{ and } \pi(F(m)) = c\} = F((\pi F)^{-1}(c))$$

This is connected or empty since F is certainly continuous.

Next observe that B_1 holds as M is connected so it is sufficient to prove A_n by induction.

The induction uses two lemmas, the first of which is quite general and requires some results about homology and Morse theory which are really beyond the scope of this essay, so the proof has been omitted. Firstly some more definitions:

Definition 3.2.

• *The <u>Hessian</u> of a smooth function $f \in C^\infty(M)$ at a critical point $z \in M$ is the symmetric bilinear form on the tangent space T_zM given by*

$$H_z(X_z, Y_z) = X_z(Y(f))$$

where $X_z, Y_z \in T_zM$ and Y is any vector field taking the value Y_z at z.

• *f is said to be <u>clean</u> (or non-degenerate in the sense of Bott) if:*

(i) Its critical points form a submanifold Z of M, and

(ii) At each critical point the Hessian of f is non-degenerate in the directions normal to Z

• *The <u>index</u> of f at a critical point is the number of negative entries in any diagonalisation of the Hessian of f at that point.*

Remarks

It's easy to check the Hessian is well-defined and symmetric using the fact that, for any vector fields X, Y:

$$X(Y(f)) - Y(X(f)) = (\mathcal{L}_X Y)(f) = (\mathcal{L}_X Y) \cdot df$$

which is zero at any critical point of f.

By (i), if $z \in Z$ then $T_z Z$ is a subspace of $T_z M$ and (ii) just means that if $H_z(x, y) = 0$ for all $y \in T_z M$ then $x \in T_z Z$ or equivalently that H_z is non-degenerate when restricted to any complement of $T_z Z$ in $T_z M$.

Lemma 3.3. *If N is a compact connected manifold and $f \in C^\infty(N)$ is clean and both f and $-f$ have no critical points of index 1 then for any $c \in \mathbb{R}$, $f^{-1}(c)$ is connected or empty.*

(Proof omitted) See [Au] p63, [B] or [B1].

We now study the Hamiltonian action generated by a component of the moment map μ and see how it relates to this lemma, by examining the behaviour in the neighbourhood of a critical pont.

Suppose $f = \langle \mu, x \rangle$ where $x \in \mathfrak{t}$. Let T' be the closure of the 1-parameter subgroup $\exp(tx)$ of T generated by x; it's a compact connected subgroup and therefore a subtorus of T. Let Z be the set of critical points of f in M. Observe that $Z = \{\text{zeros of } X_f\} = \{\text{fixed points of the flow of } X_f\}$. It follows that Z is the set of fixed points of the induced action of T' on M since the flow of X_f acts as a dense subgroup of T' and the action is smooth.

Since T' is compact we can give M a T'-invariant metric and use Propositions 2.15 and 2.16 to endow M with a calibrated T'-invariant metric g and almost-complex structure $\mathfrak{J}$ that is invariant under the action of T'.

Choose any critical point $z \in Z$ and let $V = T_z M$. We can set up normal coordinate neighbourhoods: U a neighbourhood of z diffeomorphic to a neighbourhood W of the origin in V, as described earlier. Since every diffeomorphism from the action of T' fixes z we obtain a linear representation of T' on V. By Proposition 2.17 this representation corresponds to the action of T' on M via normal coords, at least for all group elements in some neighbourhood of the identity. This implies that the subspace $V_0 \subset V$ of fixed points of the representation provides a coordinate patch on $Z \cap U$ about z. It follows that Z is a submanifold of M and $T_z Z = V_0$.

The properties of $\mathfrak{J}$ and g allow us to regard V as a complex vector space on which T' is represented by unitary transformations. This allows us to choose a complex basis for V such that T' acts by diagonal matrices. To see how to do this consider the following commutative diagram:

$$\begin{array}{ccc} \mathfrak{t}' & \xrightarrow{\exp} & T' \\ \Big\downarrow{\scriptstyle (d\rho)_e} & & \Big\downarrow{\scriptstyle \rho} \\ \mathfrak{u} & \xrightarrow{\exp} & U \end{array}$$

where $\mathfrak{u}$ is the Lie algebra of the unitary group U and ρ is the representation. Recall that $\mathfrak{u}$ can be identified with the skew-adjoint matrices which are diagonalisable. Since the map on the top row is a surjection, the image of ρ is a set of pairwise commuting diagonalisable transformations. Any such set can be simultaneously diagonalised.

Thus we can decompose V into a direct sum of complex subspaces

$$V = V_0 \oplus V_1 \oplus \cdots \oplus V_k$$

where V_0 is the subspace of fixed points of the representation (as above) and the others are of complex dimension 1. T' acts on each of $V_1, \ldots, V_k$ nontrivially by unitary transformations, i.e. multiplication by complex numbers of modulus one.

In particular we have nonzero real numbers $\lambda_1, \ldots, \lambda_k$ such that $\exp(-tx)$ acts on V_j by multiplication by $e^{it\lambda_j}$. They are nonzero as $\exp(tx)$ generates a dense subgroup of T' and so if it fixes V_j then so does all of T' contradicting the definition of T'. This allows us to describe how X_f behaves near z since X_f induces the flow $\exp(-tx) \cdot m$ which has this simple description in terms these coordinates. To get at df near z we need to know how ω behaves and for this, Darboux's Theorem will serve our purposes.

We claim that the Hessian of f at z is given by

$$H_z(u, v) = \sum_{j=1}^{k} \lambda_j g_z(u_j, v_j)$$

where $u_j, v_j \in V_j$ are the decompositions of $u, v \in V$. It follows directly from this expression that the Hessian is non-degenerate in the normal directions and has even index.

To prove the claim first observe that given $Y_z \in T_z M$ we can find a vector field Y taking the value Y_z at z such that $di_Y\omega$ is zero in a neighbourhood of z. We can obtain such Y by expressing Y_z in terms of the basis obtained from local coordinates provided by Darboux's Theorem, in which the symplectic form ω is constant. Then we have

$$\begin{aligned} X(Y(f)) &= X(Y.df) = X(\omega(Y, X_f)) \\ &= X \cdot d(\omega(Y, X_f)) = X \cdot i_{[X_f, Y]}\omega \quad \text{(c.f proof of Prop. 2.3)} \\ &= \omega([X_f, Y], X) \end{aligned} \tag{15}$$

And also if $Y_z \in V_j$ then

$$\begin{aligned} [X_f, Y]_z &= \frac{d}{dt}\big(d\Phi_{\exp(tx)} Y_z\big)\,\big|_{t=0} \\ &= \frac{d}{dt}\big(e^{-it\lambda_j} Y_z\big)\,\big|_{t=0} \\ &= -i\lambda_j Y_z = -\lambda_j \Im Y_z \end{aligned} \tag{16}$$

Thus the claim follows by evaluating (15) at z, substituting in (16) for each j and recalling the calibration identity $\omega(x, y) = g(\Im x, y)$

Thus we have proved:

Lemma 3.4. *In the situation of Theorem 3.1, if $f \in C^\infty(M)$ is a component of μ then f is clean and only has critical point of even index.*

Combining Lemma 3.3 and Lemma 3.4 we have:

Corollary 3.5. *In the situation of Theorem 3.1, if $f \in C^\infty(M)$ is a component of μ then for any $c \in \mathbb{R}$, $f^{-1}(c)$ is connected or empty.*

This is precisely A_1 of Theorem 3.1. We'll now assume A_n and show that A_{n+1} holds.

Pick any $c \in \mathbb{R}^{n+1}$ and let

$$N = f_1^{-1}(c_1) \cap \ldots \cap f_n^{-1}(c_n)$$

We want to show that $N \cap f_{n+1}^{-1}(c_{n+1}) = (f_{n+1}|_N)^{-1}(c_{n+1})$ is connected or empty (where $f_{n+1}|_N$ denotes the restriction to N), so it is sufficient to prove that $f_{n+1}|_N$ satisfies the conditions of Lemma 3.3. Without loss of generality we may assume that:

• $f_1, \ldots, f_{n+1}$ are linearly independent (since otherwise we could omit one and apply A_n).

• By continuity, c is a regular value of $F = (f_1, \ldots, f_{n+1})$ (this can be seen with just point set topology)

This implies that N is a submanifold of codimension n in M and by A_n it is connected.

Let z be a critical point of $f_{n+1}|_N$, so that on M

$$df_{n+1} + \sum_{i=1}^{n} \nu_i df_i$$

is zero at z for some constants ν_i, since T_zN is the subspace of T_zM annihilated by $\{df_1, \ldots, df_n\}$. In particular z is a critical point of the function on M given by

$$f = f_{n+1} + \sum_{i=1}^{n} \nu_i f_i \tag{17}$$

which is a component of μ by linearity. Note that $f|_N$ and $f_{n+1}|_N$ differ only by the constant $\sum_{i=1}^n \nu_i c_i$ and so have identical critical points. This is slightly subtle: critical points of f that are in N are clearly critical points of $f_{n+1}|_N$ but conversely a critical point of $f_{n+1}|_N$ is not necessarily a critical point of f, although we can always choose some f (by (17)) that is critical at that point.

The idea is that we can now use the previous analysis of the behaviour of such f in the neighbourhood of a critical point z to deduce that $f|_N$, and therefore $f_{n+1}|_N$, satisfy the conditions of Lemma 3.3 as required.

We have $T_zM = V_0 \oplus V_1 \oplus \cdots \oplus V_k$ where, as before, V_0 is tangent to the critical manifold Z (which corresponds via normal coordinates to the fixed points of the flow of X_f) and $V_1 \oplus \cdots \oplus V_k$ represent the non-degenerate normal directions. We claim

$$V_1 \oplus \cdots \oplus V_k \subset T_zN$$

or equivalently that each df_i annihilates $V_1 \oplus \cdots \oplus V_k$ for $i = 1, \ldots, n$. To prove this first observe that each tangent vector $(X_{f_i})_z$ is in V_0 since the flow generated by X_{f_i} preserves Z (Z is the set of fixed points of the flow of X_f and these flows commute since all the flows come from the torus T). If $v \in V_1 \oplus \cdots \oplus V_k$ then

$$df_i(v) = \omega(v, (X_{f_i})_z) = g(\Im v, (X_{f_i})_z)$$

and it is easy to check this is zero since $\Im v \in V_1 \oplus \cdots \oplus V_k$ and $(X_{f_i})_z \in V_0$.

From this claim we deduce $T_zZ + T_zN = V_0 + T_zN = T_zM$ so that Z and N intersect transversely at z. It follows that $Z \cap N$ is a submanifold and $T_z(Z \cap N) = T_zZ \cap T_zN = V_0 \cap T_zN$. Thus the Hessian restricted to N is diagonal with respect to the decomposition $(V_0 \cap T_zN) \oplus V_1 \oplus \ldots \oplus V_k$ of T_zN, so is of even index and non-degenerate in directions normal to $Z \cap N$.

Non-degenerate in the normal directions implies that given any 'normal vector' $v \in T_zN \backslash T_z(N \cap Z)$ we can choose a component of $df|_N$ which has nonzero rate of change in the direction of v. It follows that there is a neighbourhood of z such that any critical point of $f|_N$ is in $N \cap Z$ so is also a critical point of f.

Hence the connected component of $N \cap Z$ containing z is the component of the critical submanifold of $f|_N$ containing z so we can apply Lemma 3.3, since $-f$ can be done by the same argument, to establish A_{n+1}.

We'll now see how the rest of Theorem 3.1 follows. Firstly if Z_j is a connected component of Z (the set of common critical points of $f_1, \ldots, f_n$) then F is constant on Z_j and so $F(Z_j) = c_j$ is a single point. Note that it is a general result that there are only a finite number of connected components of Z. Next observe that if $\lambda \in \mathbb{R}^n$ is such that the flow of $\sum_{i=1}^n \lambda_i f_i$ generates a dense subgroup of the torus action generated by $f_1, \ldots, f_n$ then Z is also the set of critical points of $\sum_{i=1}^n \lambda_i f_i$. Moreover this occurs for a dense set of $\lambda \in \mathbb{R}^n$. Thus the points of the compact set $F(M)$ where a generic linear form $\sum_{i=1}^n \lambda_i x_i$ on $\mathbb{R}^n$ will take its extremal values are in $F(Z) = \{c_1, \ldots, c_m\}$. In particular any convex set containing $\{c_1, \ldots, c_m\}$ will contain $F(M)$ (for example consider 'supporting hyperplanes' of the convex set). Of course, $\{c_1, \ldots, c_m\}$ is a subset of $F(M)$ and by B_n, $F(M)$ is convex, so $F(M)$ is the smallest convex set containing $\{c_1, \ldots, c_m\}$, i.e.

$$F(M) \text{ is the convex hull of } \{c_1, \ldots, c_m\} \quad \square$$

4. Cones and Reconstruction

4.1. Basic example.

As we saw in the proof of Theorem 3.1 any unitary representation of a torus on $\mathbb{C}^d$ can be diagonalised. Now it is clear that any such representation of the circle, regarded as $\mathbb{R}/2\pi\mathbb{Z}$ here, on $\mathbb{C}$ has the form

$$s \cdot z \mapsto e^{isn_1} z$$

for some integer n_1. Thus in the general case by considering the generators of the torus $\mathbb{R}^n/\mathbb{Z}^n$ and using the diagonal basis of $\mathbb{C}^d$ we deduce that the representation has the form

$$s \cdot z \mapsto (e^{i\langle s, n_1\rangle} z_1, \dots, e^{i\langle s, n_d\rangle} z_d)$$

for some $n_1, \dots, n_d \in (\mathbb{Z}^n)^*$. We claim that this admits the moment map

$$\mu(z) = \frac{1}{2}|z_1|^2 n_1 + \dots + \frac{1}{2}|z_d|^2 n_d + \lambda$$

for any constant $\lambda \in (\mathbb{R}^n)^*$. This is not hard to check by generalising the calculation for the one dimensional case. In particular note that the image of μ is the cone with vertex λ and generated by $n_1, \dots, n_d$.

4.2. The AGS vertex cones.

We can use the example above together with the Darboux Theorem to deduce more information about the convex polyhedron Δ produced by the Atiyah, Guillemin-Sternberg Convexity Theorem. Suppose we are in the situation of that Theorem and $z \in M$ is a fixed point of the action. Set up normal coordinates about z and diagonalise the action as we did there. Consider the following diagram:

$$\begin{tikzcd} W \arrow[r, "\psi"] & W \arrow[r, "\exp"] \arrow[d, "\mu_0"] & U \arrow[d, "\mu"] \\ & \mathfrak{t}^* & \mathfrak{t}^* \end{tikzcd}$$

where $U \subset M$ is the normal coordinate neighbourhood of z, diffeomorphic to W, a neighbourhood of the origin in T_zM, and we've called this diffeomorphism exp here. μ is the moment map on M as required by the Theorem. W has two symplectic forms on it, namely the pullback $\exp^*\omega$ and the constant form ω_0 obtained when we identified the tangent space as a complex vector space; it agrees with $\exp^*\omega$ at the origin. The diagonalisation of the action means that the linear action of T^n on T_zM is precisely as in the basic example above. As such it is symplectic with respect to ω_0 and admits a moment map, μ_0. Of course μ_0 is determined up to a constant by the property that

$$\text{For any } x \in \mathfrak{t},\ d\langle\mu_0, x\rangle = \omega_0(\cdot, \widetilde{X}_W)$$

where $\widetilde{X}_W$ is the fundamental vector field corresponding to the action on W. Similarly by pulling μ back to W we see that $\mu \circ \exp$ satisfies

$$d\langle \mu \circ \exp, x\rangle = (\exp^* \omega)(\cdot, \widetilde{X}_W) \tag{18}$$

since exp is equivariant. If we apply the equivariant version of the Darboux Theorem to W we obtain, after maybe shrinking W, a diffeomorphism $\psi : W \to W$ commuting with the action of T^n and such that $\psi^* \exp^* \omega = \omega_0$. Thus by applying ψ^* to (18) we obtain:

$$d\langle \mu \circ \exp \circ \psi, x\rangle = (\psi^* \exp^* \omega)(\cdot, \widetilde{X}_W) = \omega_0(\cdot, \widetilde{X}_W)$$

So that $\mu \circ \exp \circ \psi$ is the same as μ_0 on W up to an additive constant. In particular the image under μ of U (a neighbourhood of $z \in M$) is a neighbourhood of the vertex of the cone $\mu_0(T_z M)$ (translated appropriately).

Choosing a basis of a torus gives an induced basis of its Lie algebra and the dual basis gives a basis of the dual of the Lie algebra. Thus we have established:

Proposition 4.1. *The cone at each vertex $c \in \Delta$ is of the form*

$$\{c + \alpha_1 n_1 + \cdots + \alpha_p n_p \mid \alpha_i \in \mathbb{R}_{\geq 0}\}$$

near c for some subset $\{n_1, \dots, n_p\}$ of the integer lattice in $(\mathbb{R}^n)^*$, where p is half the real dimension of the manifold M.

4.3. **The Delzant construction.**

We've seen that the image of a moment map of a symplectic torus action on a compact manifold is a convex polyhedron which looks like a cone generated by integral vectors near each vertex. We've also seen how simple torus actions on $\mathbb{C}^n$ give rise to moment maps whose image is a full cone generated by integral vectors.

Now suppose we are given a convex polyhedron in $(\mathbb{R}^n)^*$ and wonder whether it arises as the image of some moment map. In this section we'll answer this question positively for a certain class of convex polyhedra by explicitly constructing the required symplectic manifold.

Suppose the polyhedron Δ has vertices

$$\{c_1, \dots, c_m\} \subset (\mathbb{R}^n)^*$$

If it is to be the image of a moment map of a T^n action then necessarily the vertex cones are generated by integral vectors (we'll use the standard basis of $\mathbb{R}^n$ and regard T^n as $\mathbb{R}^n/2\pi\mathbb{Z}^n$)

The polyhedra we are interested in here are those that satisfy the following hypothesis:

Hypothesis 4.2. *The cone at each vertex $c \in \Delta$ is of the form*

$$\{c + \alpha_1 n_1 + \cdots + \alpha_n n_n \mid \alpha_i \in \mathbb{R}_{\geq 0}\}$$

near c for some $\underline{\mathbb{Z}\text{-basis}}$ $\{n_1, \ldots, n_n\}$ of the integer lattice in $(\mathbb{R}^n)^*$

The simplest cases of this construction yield the complex projective spaces $\mathbb{P}^n(\mathbb{C})$ from the standard n-simplex in $(\mathbb{R}^n)^*$, having vertices at the origin and at the end of each standard basis vector. If $n = 1$ we will get the 2-sphere $S^2 = \mathbb{P}^1(\mathbb{C})$ with a circle action rotating it about a fixed axis and the moment map being the 'height function' along this axis (so the image is a closed interval in the real line).

Recall that the cone in the hypothesis is the image of a moment map of the T^n-action on $\mathbb{C}^n$ given by

$$s \cdot z = (e^{i\langle s, n_1\rangle} z_1, \ldots, e^{i\langle s, n_n\rangle} z_n)$$

The hypothesis implies that we can choose a $\mathbb{Z}$-basis $\{u_1, \ldots, u_n\}$ of $\mathbb{Z}^n$ such that $\langle u_i, n_j\rangle = \delta_{ij}$ (this is equivalent to changing basis on the torus; each u_i generates a circle). Observe that if $s = s_1 u_1 + \cdots + s_n u_n$ with respect to this basis then the action takes the simpler form:

$$s \cdot z = (e^{is_1} z_1, \ldots, e^{is_n} z_n) \tag{19}$$

It is clear that u_i is normal to the face of Δ spanned by $\{n_1, \ldots, \widehat{n}_i, \ldots, n_n\}$. The fact that $\langle u_i, n_i\rangle = 1$ implies that u_i is 'inward pointing' and of minimal length. Thus u_i is determined uniquely by the face of Δ it defines (i.e. if we start from two vertices on the same face then we get the same u_i). By considering other vertices too, we obtain vectors $u_1, \ldots, u_d \in \mathbb{Z}^n$ (one for each face) each generating a circle in T^n. Thus there are $\lambda_1, \ldots, \lambda_d \in \mathbb{R}$ such that

$$\Delta = \bigcap_{i=1}^{d} \{c \in (\mathbb{R}^n)^* \mid \langle c, u_i\rangle \geq \lambda_i\}$$

Now let S_k denote the set of indices corresponding to the faces meeting at the vertex $c_k \in \Delta$ (without loss of generality $S_1 = \{1, \ldots, n\}$). Thus $\{u_i \mid i \in S_k\}$ is a $\mathbb{Z}$-basis of $\mathbb{Z}^n$ and is a set of normals to the faces meeting at c_k. As we saw above, in (19), these bases are precisely what we need to write the T^n actions yielding the vertex cones in a simple form. We now use the overlapping of these bases to combine the actions into the action of a big torus T^d on $\mathbb{C}^d$.

To this end, we let T^d act on $\mathbb{C}^d$ via

$$s \cdot z = (e^{is_1} z_1, \ldots, e^{is_d} z_d) \tag{20}$$

where $s = \sum_1^d s_i e_i$ and $e_1, \ldots, e_d$ is the standard basis. We obtain an embedding homomorphism

$$i_k : T^n \to T^d$$

for each vertex of Δ by using the basis $\{u_i \mid i \in S_k\}$ of T^n and mapping u_i to e_i for $i \in S_k$ (so the image of i_k is the subtorus of T^d generated by $\{e_i \mid i \in S_k\}$). Observe that the action of T^n on $\mathbb{C}^d$ via i_k is really the same as that in (19) for T^n on $\mathbb{C}^n$ since the components z_i of $z \in \mathbb{C}^d$ are fixed if $i \notin S_k$. In particular we have an induced moment map $\mu_k : \mathbb{C}^d \to (\mathbb{R}^n)^*$ for each embedding, whose image is the cone at the vertex c_k. Thus μ_k is such that if $i \in S_k$ then the function $\langle \mu_k, u_i \rangle : \mathbb{C}^d \to \mathbb{R}$ generates the same action as that generated by e_i (regarded as an element of the Lie algebra of T^d), and the value of this function at the origin is $\langle c_k, u_i \rangle = \lambda_i$. Then if we take the momentum map for the T^d action (20) to be

$$\mu(z) = (\frac{1}{2}|z_1|^2 + \lambda_1, \ldots, \frac{1}{2}|z_d|^2 + \lambda_d)$$

it follows that

$$\mu_k = i_k^* \circ \mu$$

We also have the surjective map

$$\pi : T^d \to T^n \ ; \ \sum_1^d s_i e_i \mapsto \sum_1^d s_i u_i$$

Observe that we have $\pi \circ i_k = id_{T^n}$ for each i_k.

Let $N = \ker(\pi)$. N is a torus since for instance it has basis

$$\{e_i - \sum_{j=1}^n a_{ij} e_j \mid i = n+1, \ldots, d\}$$

where the integers a_{ij} are defined by expressing u_i in terms of the basis from the vertex c_1, i.e. $u_i = \sum_{j=1}^n a_{ij} u_j$ in $\mathbb{Z}^n$. N measures the extent to which the embeddings of T^n in T^d differ: If $s \in T^n$ then the difference $i_k s - i_l s$ is in N and conversely, e.g. by using the basis, any element of N can be decomposed as the sum of such differences. Thus we are looking to quotient by N in order to obtain a single well defined T^n action.

By differentiating and transposing the exact sequence

$$0 \longrightarrow N \xrightarrow{i} T^d \xrightarrow{\pi} T^d \longrightarrow 0$$

we obtain the following commutative diagram:

$$\begin{array}{ccccccccc} & & & & \mathbb{C}^d & = & \mathbb{C}^d & & \\ & & & & \big\downarrow{\scriptstyle \mu} & & \big\downarrow{\scriptstyle \mu_N} & & \\ 0 & \longrightarrow & (\mathbb{R}^n)^* & \xrightarrow{\pi^*} & (\mathbb{R}^d)^* & \xrightarrow{i^*} & \mathfrak{n}^* & \longrightarrow & 0 \end{array}$$

where μ_N is the induced moment map for the action of N on $\mathbb{C}^d$, and the row is exact. Let $\Delta' = \pi^*\Delta$ be the embedded image of Δ in $(\mathbb{R}^d)^*$.

Claim: $\mu_N^{-1}(0) = \mu^{-1}(\Delta')$

Since the rows are exact $\mu_N^{-1}(0) = \mu^{-1}(\pi^*((\mathbb{R}^n)^*))$ and so it is sufficient to prove that

$$\pi^*((\mathbb{R}^n)^*) \cap \mu(\mathbb{C}^d) = \pi^*(\Delta)$$

This follows from the crucial observation that since $u_i = \pi(e_i)$,

$$\langle u_i, c\rangle \geq \lambda_i \iff \langle e_i, \pi^*(c)\rangle \geq \lambda_i \tag{21}$$

Note that the induced action of π on the Lie algebras is simply the linear map defined in the same fashion as π (that is $e_i \mapsto u_i$) since these maps agree in a neighbourhood of the identity.

We deduce from this claim that N acts freely on $\mu_N^{-1}(0)$: If $z \in \mu_N^{-1}(0)$ then $\mu(z) = \pi^*(c)$ for some $c \in \Delta$. c is on at most n faces of Δ and the intersection of these faces contains some vertex c_k. It follows that the inequalities (21) are strict if $i \notin S_k$ and so (from the definition of μ) $z_i \neq 0$ if $i \notin S_k$. In particular any $s \in T^d$ fixing z has $s_i = 0$ if $i \notin S_k$ (by (20)) and so $s \in i_kT^n$. Thus $i_k\pi s = s$ (since $\pi i_k = id_{T^n}$) so that if $s \in N$ and fixes z then s is the identity.

This enables us to apply Theorem 2.13 to establish that the quotient $X_\Delta = \mu_N^{-1}(0)/N$ is a symplectic manifold. X_Δ is compact since firstly the formula for μ implies that $\mu^{-1}(\Delta')$ is closed and bounded and secondly, by definition, the projection on to the quotient is continuous. Observe $\mu_N^{-1}(0)$ is a union of T^d orbits in $\mathbb{C}^d$ (since μ is equivariant and $\mu_N^{-1}(0) = \mu^{-1}(\Delta')$). Proposition 2.14 implies that the T^d action passes down to a symplectic action on X_Δ and admits a moment map $\overline{\mu}$ such that

$$\overline{\mu} \circ q = \mu \circ j$$

where $q : \mu_N^{-1}(0) \to X_\Delta$ is the projection and $j : \mu_N^{-1}(0) \to \mathbb{C}^d$ is the inclusion. Since we have quotiented by N this action is the same as each induced T^n action. Moreover the image of the moment map obtained is

$$\overline{\mu}(q(\mu_N^{-1}(0))) = \mu(\mu^{-1}(\Delta')) = \Delta'$$

which is identified via π^* with Δ.

5. Bibliography

5.1. Sources used directly.

[**A**] **M.F.Atiyah** Convexity and Commuting Hamiltonians, Bull. LMS 14, 1-15, 1982

[**A1**] **M.F.Atiyah** Angular Momentum, Convex Polyhedra and Algebraic Geometry, Proc.Edinburgh.Math.Soc. 26, 121-138, 1983

[**Au**] **M.Audin** The Topology of Torus Actions on Symplectic Manifolds, Birkhauser, Boston, 1991

[**B**] **R.Bott** Nondegenerate Critical Manifolds, Ann. Math. 60, 2, 248-261,Sept 1954

[**B1**] **R.Bott** Lectures on Morse Theory, Bonn Uni. Lecture Notes, 1958

[**D**] **T.Delzant** Hamiltoniens Periodiques et Image Convex de l'Application Moment, Bull. Soc. Math. France 116, 315-339, 1988

[**G**] **V.Guillemin** Moment Maps and Combinatorial Properties of Hamiltonian T^n-spaces, Birkhauser, 1994

[**GS**] **V.Guillemin & S.Sternberg** Convexity Properties of the Moment Mapping, Invent.Math. 67, 491-513, 1982

[**GS1**] **V.Guillemin & S.Sternberg** Convexity Properties of the Moment Mapping II, Invent.Math. 77, 533,546, 1984

[**GS2**] **V.Guillemin & S.Sternberg** Symplectic Techniques in Physics, CUP, Cambridge, 1984

[**HZ**] **H.Hofer & E.Zehnder** Symplectic Invariants & Hamiltonian Dynamics, Birkhauser, Basil, 1994

[**KN**] **S.Kobayashi & K.Nomizu** Foundations of Differential Geometry vol.1, Wiley Interscience, 1963

[**LM**] **P.Libermann & C-M.Marle** Symplectic Geometry and Analytical Mechanics, D.Reidel Publishing Co., 1987

5.2. Further reading.

[**A2**] **M.F.Atiyah** The Moment Map in Symplectic Geometry, in: Global Riemannian Geometry, eds N.J.Hitchin & T.J.Willmore, Ellis Horwood, 1984

[**A3**] **M.F.Atiyah** The Geometry and Physics of Knots, CUP, 1990

[**AB**] **M.F.Atiyah & R.Bott** The Yang-Mills Equations over Riemann Surfaces, Phil Trans. Roy. Soc. London 308, 523-615, 1982

[**AB1**] **M.F.Atiyah & R.Bott** The Moment Map and Equivariant Cohomology, Topology, 23, 1-28, 1984

[**AP**] **M.F.Atiyah & A.N.Pressley** Convexity and Loop Groups, in: Geom. and Arithmetic, Birkhauser, 1983

[**Da**] **V.I.Danilov** The Geometry of Toric Varieties, Russ.Math.Surv. 33:2, 97-154, 1978

[**F**] **W.Fulton** Introduction to Toric Varieties, Ann. Math. Studies, PUP, Princeton, NJ, 1993

[**G2**] **V.Guillemin** Kahler Structures on Toric Varieties, J.Diff.Geom. 40, 285-309, 1994

[**GS3**] **V.Guillemin & S.Sternberg** Geometric Quantisation and Multiplicities of Group Representations, Invent. Math. 67, 515-538, 1982

[**H**] **N.J.Hitchin** Hyperkahler Manifolds, Seminaire Bourbaki, Asterisque 206, 137-166, 1992

[**JW**] **L.C.Jeffrey & J.Weitsman** Toric Structures on the Moduli Space of Flat Connections on a Riemann Surface: Volumes and the Moment Map, Adv. in Math. 106, 151-168, 1994

[**K**] **F.Kirwan** Cohomology of Quotients in Symplectic and Algebraic Geometry, PUP, 1984

[**K1**] **F.Kirwan** Convexity Properties of the Moment Mapping III, Invent. Math. 77,547-552, 1984

[**KT**] **Y.Karshon & S.Tolman** The Moment Map and Line Bundles over Presymplectic Toric Manifolds, J.Diff.Geom. 38, 465-484, 1993

[**MFK**] **D.Mumford, J.Fogarty & F.Kirwan** Geometric Invariant Theory, 3rd edition, Springer Verlag, 1994

[**W**] **N.M.J.Woodhouse** Geometric Quantisation, OUP, 1991