CONVEXITY AND COMMUTING HAMILTONIANS

M. F. ATIYAH

§1. *Introduction*

A well-known result of Schur [9] asserts that the diagonal elements $(a_1, \ldots, a_n)$ of an $n \times n$ Hermitian matrix A satisfy a system of linear inequalities involving the eigenvalues $(\lambda_1, \ldots, \lambda_n)$. In geometric terms, regarding $\mathbf{a}$ and λ as points in R^n and allowing the symmetric group Σ_n to act by permutation of coordinates, this result takes the form

(1.1) $\mathbf{a}$ *is in the convex hull of the points* $\Sigma_n \lambda$.

The converse was proved by A. Horn [5], so that all points in this convex hull occur as diagonals of some matrix A with the given eigenvalues. Kostant [7] generalized these results to any compact Lie group G in the following manner. We consider the adjoint action of G on its Lie algebra $L(G)$. If T is a maximal torus of G and W its Weyl group, then it is well known that W-orbits in $L(T)$ correspond to G-orbits in $L(G)$. Now fix a G-invariant metric on $L(G)$, so that we can define orthogonal projection. Then Kostant's result is†

(1.2) *The orthogonal projection of a G-orbit onto $L(T)$ coincides with the convex hull of the corresponding W-orbit.*

Clearly (1.2) reduces to (1.1) when G is the unitary group $U(n)$: we replace the Hermitian matrix A by iA, and note projection here amounts to taking the diagonal part.

The purpose of this paper is to put (1.2) into the more general context of symplectic geometry. Thus we shall prove a general result which reduces to (1.2) in the homogeneous case. For Kähler manifolds we shall go further and prove a more delicate result which, in the homogeneous case, yields an interesting refinement of (1.1) and (1.2).

Our proofs are quite different from those of Kostant (at least for the difficult part) and depend on some Morse theory ideas first employed by T. Frankel [2]. There is also an interesting connection with the ideas of toroidal compactification as developed by Mumford and others [8].

To formulate our main result we recall the rudiments of symplectic geometry. First a symplectic manifold M is a differentiable manifold of even dimension with an exterior differential 2-form ω which satisfies

(i) $d\omega = 0$,

(ii) ω is of maximal rank.

Received 27 March, 1981.

† Actually Kostant deals with a somewhat more general situation.

A real-valued smooth function f on M defines a Hamiltonian vector-field X_f which corresponds to the 1-form df using the duality defined by ω. The Poisson bracket $\{f, g\}$ of two functions is defined by

$$\{f, g\} = X_f(g)$$

and is skew-symmetric: f and g are said to Poisson commute if $\{f, g\} = 0$. A vector field X is said to be *periodic* if it generates a circle action (so that all orbits are closed with periods dividing some number T). It is said to be *almost periodic* if it generates a torus action. Our main result is then the following

THEOREM 1. *Let M be a compact connected symplectic manifold and let $f_1, \ldots, f_n$ be n real-valued functions which Poisson commute and whose Hamiltonian vector fields are almost-periodic. Then the map $f: M \to R^n$ given by the f_i satisfies*

(A_n) *all (non-empty fibres) $f^{-1}(c)$ are connected ($c \in R^n$);*

(B_n) *the image $f(M)$ is convex.*

Moreover if $Z_1, \ldots, Z_N$ are the connected components of the set $Z \subset M$ of common critical points of the f_i then $f(Z_j) = c_j$ is a single point and $f(M)$ is the convex hull of $c_1, \ldots, c_N$.

REMARKS. 1. The essential case of Theorem 1 is when all f_i have periodic Hamiltonian fields so that we have a symplectic action of the n-torus T^n on M. Conversely such an action leads to functions $f_1, \ldots, f_n$ as in Theorem 1 provided the first Betti number of M vanishes. The map f is usually called the moment map and is canonically a map $M \to L(T)^*$.

2. Replacing f_1 by f_1^2 (say) destroys the convexity of the image and shows that the periodicity assumption is essential.

3. Although we are mainly interested in (B_n) and the subsequent statement about the points $c_1, \ldots, c_N$, we have formulated (A_n) also because (A_n) actually implies (B_{n+1}). Note that (B_1) is trivial, since any compact connected set on the real line is an interval, while (A_1) and (B_2) are non-trivial, and represent the first interesting cases.

4. For any Lie group G any orbit M in the dual $L(G)^*$ of the adjoint representation has a natural homogeneous symplectic structure. Any element $u \in L(G)$ is a linear function on $L(G)^*$ and so defines a function on M. The corresponding Hamiltonian vector field is just the natural action of u on M (see **[6]**). Taking $u_1, \ldots, u_n$ to lie in the Lie algebra of a torus $T \subset G$ we are then in the situation of Theorem 1 provided M is compact. If G is compact, so is M, and we can identify $L(G)$ with $L(G)^*$ by an invariant metric. This now reduces to Kostant's situation yielding (1.2): it is easy to identify our points $c_1, \ldots, c_N$ with the W-orbit. Note that (A_n) asserts (for $G = U(n)$) that *the set of $n \times n$ Hermitian matrices with given eigenvalues and given admissible diagonal elements is connected.*

We turn now to the case when M is a compact Kähler manifold. Thus, in addition to being symplectic, M is also Riemannian and complex analytic, these three structures being compatible in various ways. When we consider an action of the torus T^n on M we shall now require it to preserve all structures: in fact because of their compatibility it is enough to preserve just the Riemannian metric and either of the other two structures (as proved in [2] even this last requirement can be dropped). Now the automorphism group of any compact complex manifold is a complex Lie group and so the action of T^n extends to a holomorphic action of its complexification T_c^n: this "complex torus" is a product of n copies of the multiplicative group C^* of complex numbers. Our next result refines Theorem 1 by describing the behaviour of the functions $f_1, \ldots, f_n$ on the orbits of T_c^n:

THEOREM 2. *Let M be a compact Kähler manifold and let $f_1, \ldots, f_n$ be n real-valued functions which Poisson commute and whose Hamiltonian vector-fields are periodic and generate an n-dimensional torus T. Let T_c be its complexification, Y an orbit of T_c acting on M and $\bar{Y}$ its closure. Finally let Z_j $(j = 1, \ldots, p)$ be those components of the common critical points of the f_i which intersect $\bar{Y}$ and put $c_j = f(Z_j) \in R^n$. Then the map $f\colon \bar{Y} \to R^n$ satisfies*

(a) *$f(\bar{Y})$ is the convex polytope P with vertices $c_1, \ldots, c_p$,*

(b) *For each open face σ of P, the inverse image $f^{-1}(\sigma)$ in $\bar{Y}$ consists of a single T_c-orbit,*

(c) *f induces a homeomorphism of $\bar{Y}/T$ onto P.*

REMARKS. 1. Theorem 2 applies whenever we have a complex torus acting holomorphically on a compact connected Kähler manifold with at least one fixed point (see [2]), since by averaging over the real torus we can produce an invariant Kähler metric. The interest is then not in the functions f_i (which depend on the metric) but in the implications about the structure of the closure $\bar{Y}$ of an orbit Y. We shall explain in Section 3 the relation of this result with Mumford's toroidal compactifications [8].

2. Our proof of the convexity of $f(\bar{Y})$ in Theorem 2(a) will be rather different from the proof of convexity of $f(M)$ in Theorem 1. This could be used to give an alternative proof of Theorem 1 in the Kähler case by showing that there is always an orbit Y with $f(\bar{Y}) = f(M)$. More generally by choosing other orbits we can identify the images under f of various sub-varieties of M. Applied to the homogeneous case this will prove that the images of the Bruhat cells are convex sub-polyhedra of $f(M)$: this is the content of Theorem 3 in Section 4 and it can be viewed as a refinement of (1.1) and (1.2).

In Section 4 we also prove a result (Theorem 4) concerning the restriction of harmonic forms to T_c-orbits. This result, which arises naturally in our context, was motivated by a special case of Gel'fand and Macpherson [3] connected with the study of the di-logarithm.

In recent years there has been considerable interest in various infinite-dimensional Hamiltonian systems. One may ask whether Theorem 1 can be extended in any interesting way to infinite-dimensions. In a subsequent paper it will be shown that Theorems 1 and 3 extend to the special infinite-dimensional case when $M = \Omega(G)$ is the loop space of a compact Lie group. This case turns out to be very similar to the homogeneous case studied by Kostant.

Since writing this paper I have learnt that V. Guillemin and S. Sternberg have arrived independently and almost simultaneously at similar results, though with mildly different proofs.

§2. *The symplectic case*

In this section we shall give the proof of Theorem 1. As already mentioned (A_n) in Theorem 2 implies (B_{n+1}). To see this consider the diagram

$$\begin{array}{ccc} M & \xrightarrow{\ f\ } & R^{n+1} \\ & \searrow^{g} & \downarrow{\scriptstyle \pi} \\ & & R^n \end{array}$$

where $f = (f_1, \ldots, f_{n+1})$, π is any linear projection and $g = \pi f$. Then, for any $c \in R^n$, we have

$$f(M) \cap \pi^{-1}(c) = f\big(g^{-1}(c)\big).$$

Applying (A_n) to g it follows that this is empty or connected. Thus $f(M)$ is a compact subset of R^{n+1} meeting every line in a closed (possibly empty) interval and this is equivalent to the convexity of $f(M)$, that is, (B_n) holds.

To prove Theorem 1 it will be sufficient therefore to prove (A_n) by induction on n (recall that (B_1) is trivially true). For this we shall first need a simple application of Morse theory. We recall that a function ϕ on a manifold N is said to be non-degenerate (in the sense of Morse) if at every critical point the Hessian $H(\phi)$, given by the second derivatives, is a non-degenerate quadratic form. The critical points of such a non-degenerate function are necessarily discrete. A more general notion of non-degeneracy, due to Bott [1], allows sub-manifolds of critical points. The Hessian of ϕ is necessarily degenerate in the directions tangent to such a sub-manifold, but Bott's requirement is that $H(\phi)$ is non-degenerate in the remaining normal directions. The index of ϕ along a critical manifold Z is then defined as the number of negative terms in the diagonalization of $H(\phi)$. The result which we need is the following:

LEMMA (2.1) *Let $\phi : N \to R$ be a non-degenerate function (in the sense of Bott) on the compact connected manifold N, and assume that neither ϕ nor $-\phi$ has a critical manifold of index* 1. *Then $\phi^{-1}(c)$ is connected (or empty) for every $c \in R$.*

Proof. By continuity, it is enough to consider non-critical values of c. Let

$$N_c^+ = \{x \in N \mid \phi(x) \geqslant c\}, \quad N_c^- = \{x \in N \mid \phi(x) \leqslant c\}.$$

The Morse theory, as generalized by Bott [1], tells us that, as we increase c across a critical level, N_c^- is altered (homotopically) by attaching the negative normal bundle of the critical manifold. If the index is zero (i.e. we have a local minimum) then we get a new component, and such a component can disappear later only by crossing a level of index one. Hence if ϕ never has index one, and N is connected, it follows that ϕ has a unique locally minimal manifold and that N_c^- is always connected (for

$c \geqslant \operatorname{Min}\phi$). Similarly, if $-\phi$ never has index one, ϕ has a unique locally maximal manifold and N_c^+ is always connected (for $c \leqslant \operatorname{Max}\phi$). Now $\phi^{-1}(c)$ is the boundary of N_c^-, so that if $\phi^{-1}(c)$ is not connected we get a non-trivial $(n-1)$-cycle for N_c^- from a boundary component (where $n = \dim N$). But for $\operatorname{Min}\phi < c < \operatorname{Max}\phi$ we meet only critical manifolds with index $(-\phi) \geqslant 2$ so that the negative bundle has total dimension $\leqslant n-2$ (index $(-\phi) = 0$ would be a local maximum, and index $(-\phi) = 1$ is excluded by hypothesis). Hence we can never produce an $(n-1)$-cycle in N_c^-. Thus $\phi^{-1}(c)$ is connected, as required.

In view of this Lemma the key step towards proving Theorem 1 is the observation that Hamiltonian functions of the type occurring in Theorem 1 always satisfy the hypothesis of Lemma (2.1). More precisely we have

LEMMA (2.2). *Let M be a compact symplectic manifold, $\phi : M \to R$ a function whose Hamiltonian vector field X_ϕ is almost-periodic. Then ϕ is non-degenerate in the sense of Bott and has only critical manifolds of even index.*

Proof. Let T be the torus generated by X_ϕ so that the zero set Z of X_ϕ is precisely the fixed point set of T. It is well known that Z is then the union of submanifolds Z_j. We recall that this follows easily by picking a Riemannian metric on M which is T-invariant and using normal coordinates. Since, in our case, M also has a T-invariant symplectic structure it acquires a T-invariant almost-complex structure. Thus if V is the tangent space to M at $z \in Z$ it has a complex structure and decomposes under T as:

$$V = V_0 \oplus V_1 \oplus \dots \oplus V_p$$

where V_0 is fixed by T and is the tangent space to Z while each V_j, for $j > 0$, corresponds to a non-trivial character of T. Since X_ϕ generates T it acts on each V_j by some scalar $i\lambda_j$ with λ_j real and non-zero. The Hessian $H(\phi)$ at z is then the corresponding Hermitian form $\sum_{j=1}^{p} \lambda_j |v_j|^2$ which is non-degenerate and necessarily has even index.

REMARK. The above proof shows that the critical manifold $Z \subset M$ is also non-degenerate with respect to the symplectic structure of M, so that it inherits a symplectic structure itself. In the presence of a T-invariant metric Z is actually an almost-complex submanifold.

Combining (2.1) and (2.2) (applied to $\pm\phi$) we then get

LEMMA (2.3). *Let M be a compact symplectic manifold, $\phi : M \to R$ a function whose Hamiltonian vector field is almost periodic. Then for any $c \in R$ the level surface $\phi^{-1}(c)$ is connected (or empty).*

This Lemma is just the case (A_1) of Theorem 1 and we shall now use it inductively to prove (A_n). Assume therefore that (A_n) is true and let $f_1, \dots, f_{n+1}$ be functions on M satisfying the hypotheses of Theorem 1. We have to show that if $c = (c_1, \dots, c_{n+1}) \in R^{n+1}$ then

$$f^{-1}(c) = f_1^{-1}(c_1) \cap \dots \cap f_{n+1}^{-1}(c_{n+1})$$

is connected. Clearly we may suppose $f_1, \ldots, f_{n+1}$ linearly independent, otherwise we can drop one of them and apply (A_n). Moreover it will be sufficient by continuity to consider only regular values c, that is, values so that the df_i are linearly independent for all $x \in f^{-1}(c)$ (as we shall see later the irregular values actually form a finite number of hyperplanes in R^{n+1}). With this assumption

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

is a submanifold of codimension n in M, and by (A_n) it is *connected*. We now want to show that the function f_{n+1}, restricted to N, satisfies the hypotheses of Lemma (2.1). For this we must consider its critical points, i.e. points x of N for which (on M)

$$df_{n+1} + \sum_{i=1}^{n} \lambda_i df_i = 0 ,$$

for some constants λ_i. Such a point is then critical on M for the function

$$\phi = f_{n+1} + \sum_{i=1}^{n} \lambda_i f_i .$$

Now by Lemma (2.2) we know that $\pm\phi$ has only non-degenerate critical manifolds of even index. Let Z be the component through x: we will show that Z intersects N transversally at x. From this it will follow that $\pm\phi \mid N$, and so also $\pm f_{n+1} \mid N$ (which only differs by the constant $\sum_1^n \lambda_i c_i$), has $Z \cap N$ as non-degenerate critical manifold of even index. Since this is true starting from any critical point x of $f_{n+1} \mid N$ we can apply Lemma (2.1) and deduce that

$$f^{-1}(c) = N \cap f_{n+1}^{-1}(c_{n+1})$$

is connected (or empty) for any c_{n+1}, which will establish (A_{n+1}) inductively.

It remains to check that Z and N intersect transversally or equivalently that $df_1, \ldots, df_n$ remain linearly independent when restricted to Z at x. Now the Hamiltonian fields $X_1, \ldots, X_n$ must preserve Z and so the independent vectors $X_1(x), \ldots, X_n(x)$ lie in the tangent space $\mathscr{Z}$ to Z at x. Now Z is non-degenerate relative to the symplectic structure of M (Remark after (2.2)). Hence, for any constant $\mu = (\mu_1, \ldots, \mu_n) \neq 0$ there exists a tangent vector $Y \in \mathscr{Z}$ such that

$$\omega(\sum \mu_i X_i(x), Y) \neq 0 ,$$

where ω is the symplectic form. But by definition of the X_i this is the same as

$$\{\sum \mu_i df_i(x)\}(Y) \neq 0 ,$$

which proves our assertion.

REMARK. If we factor N locally by the action of $X_1, \ldots, X_n$ then f_{n+1} is defined on the quotient which is symplectic and we are essentially in the position of (2.2). The only difficulty is that this may not work globally, which is why we proceeded slightly differently. Note also that the critical points of $f_{n+1} \mid N$ come from the critical points of a finite number of functions ϕ—not just one.

We have completed the proof of (A_n) and (B_n) in Theorem 1, and we now come to the last part. The set Z of common critical points of $f_1, \ldots, f_n$ is also the fixed-point set of the torus T generated by the Hamiltonian fields $X_1, \ldots, X_n$ and as we have already noted Z is a disjoint union of connected submanifolds Z_j. On each Z_j we have $df_i = 0$ for all i and so each f_i is constant. Thus $f(Z_j) = c_j$ is a single point in $\boldsymbol{R}^n$. Moreover if $\phi = \sum \lambda_i f_i$ is a generic linear combination, so that the corresponding Hamiltonian field generates T, then the critical set of ϕ is precisely Z, and in particular ϕ takes its maximum on Z. Hence the linear form $\sum \lambda_i x_i$ considered as a function on $f(M) \subset R^n$, takes its maximum at one of the points $c_1, \ldots, c_n$. Since this holds for almost all $(\lambda_1, \ldots, \lambda_n)$ it follows that†

$$f(M) \subset \text{convex hull of } c_1, \ldots, c_N .$$

But by (B_n) $f(M)$ is convex and since it contains the c_j it must coincide with the convex hull as stated in Theorem 1.

Theorem 1 implies of course that the set Z of common critical points is not empty. In fact it asserts considerably more because if $f_1, \ldots, f_n$ are linearly independent, then $f(M)$ is a full n-dimensional convex polytope and so it has at least $(n+1)$ vertices. Thus *the number N of components of Z is at least* $(n+1)$.

It is illuminating to consider the map $f\colon M \to R^n$ in some more detail. For this we should introduce not only the fixed-point set Z of the torus T but also fixed-points of its various sub-tori. By considering the representations of T on the normal bundles of the components Z_j of Z we get a finite set of characters, and their kernels give a finite set of codimension-one sub-tori of T. Taking intersections these generate a finite lattice of sub-tori. Without essential loss of generality, we may assume that $\dim T = n$ and that T acts effectively on M. Then the minimal non-zero elements of our lattice will be circles $S_1, \ldots, S_k$ and the quotient $(n-1)$-torus T/S_i acts effectively on the components Σ_{ij} of the fixed-point set of S_i. Restricting f to Σ_{ij} we see therefore that its image in R^n lies in the hyperplane $\sum \lambda_{ri} x_r = \text{constant}$, where $\phi_i = \sum \lambda_{ri} f_r$ is the Hamiltonian corresponding to S_i. Moreover $f(\Sigma_{ij})$ will contain (and is spanned by) the subset of $c_1, \ldots, c_N$ corresponding to the components of Z lying in Σ_{ij}. Thus the union of all these hyperplanes contains the set of critical values of the map f. A bounding face of the convex polytope $f(M)$ must arise from a maximal or minimal component of the corresponding function ϕ.

As already noted, our proof of Theorem 1 used a small amount of Morse theory. In fact Frankel [**2**] shows that a function as in (2.3) is a "perfect" Morse function, in the sense that the sum of the Betti numbers of its critical manifolds is equal to the sum of the Betti numbers of M. Frankel states his results for Kähler manifolds but the proofs work also in the symplectic case.

§3. *Kähler manifolds*

In this section we shall give the proof of Theorem 2. Thus M is now a compact Kähler manifold and our purpose is to study the closure $\bar{Y}$ of an orbit Y of the

† This is essentially the argument of Kostant [**1**].

complex torus T_c, and to describe the map $f\colon \bar{Y} \to R^n$. There is no loss of generality in assuming that $f_1, \dots, f_n$ are linearly independent on Y: if not we simply consider an independent subset. Thus dim $Y = 2n$ and T acts on Y with a finite isotropy group Γ, so that T/Γ acts freely.

We first note that, because on a Kähler manifold the various structures are compatible, for any $\phi = \sum \lambda_i f_i$ we have grad $\phi = JX_\phi$ where X_ϕ is the Hamiltonian field of ϕ, J is multiplication on tangent vectors by $\sqrt{-1}$ and grad ϕ is defined relative to the metric. More formally if $(\, ,)$ and $\langle\, , \rangle$ denote the symplectic and Riemannian pairings respectively, and η is any tangent vector then

$$\langle \eta, \operatorname{grad} \phi \rangle = d\phi(\eta) = (\eta, X_\phi) = \langle \eta, JX_\phi \rangle .$$

If we decompose T_c as $T \times H$ corresponding to the Lie algebra decomposition $L(T) \oplus JL(T)$, then the gradient flow ϕ_t of ϕ is just given by the action of the corresponding one-parameter subgroup H_ϕ of H. In particular ϕ is a strictly monotonic increasing function along any (non-trivial) orbit of H.

As a first step we shall prove

LEMMA (3.1). *For any $y \in Y$ let y_∞ be a limit point of $\phi_t(y)$ as $t \to \infty$. Then*

(i) *y_∞ lies on a critical manifold N of ϕ,*

(ii) *$\lim_{t \to \infty} \phi(\phi_t(y))$ exists and is a constant $d(\phi)$ independent of y,*

(iii) $d(\phi) = \sup_Y \phi$.

Proof. (i) is an elementary general fact about the flow of any gradient vector field. For (ii) we note that $\phi(\phi_t(y))$ is, for fixed y, a bounded monotonic function of t and so the limit d exists and is equal to $\phi(y_\infty)$. Now ϕ is constant along N and T_c leaves N invariant, hence for any $u \in T_c$,

$$\lim_{t \to \infty} \phi(\phi_t(uy)) = \phi(uy_\infty) = \phi(y_\infty) = d ,$$

is independent of u. Clearly (iii) follows from (ii).

Taking ϕ generic so that $T_\phi = T$ then N is one of the components Z_j of the fixed-point set of T and so (3.1) implies

$$\sup_Y \phi = \sup_j \phi(Z_j)$$

where $j = 1, \dots, p$ runs over those indices for which Z_j meets $\bar{Y}$. Putting $c_j = f(Z_j)$ as before this implies

(3.2) *$f(\bar{Y})$ is contained in the convex hull P of $c_1, \dots, c_p$*

which is part of Theorem 2(a).

Our next step is to prove that $f(\bar{Y}) = P$. For this we consider any point $y \in Y$ and let δ be the Euclidean distance of $p = f(y)$ from the boundary ∂P of P. Since the Jacobian of f has rank n everywhere on Y, p is necessarily an interior point of P and so $\delta > 0$. We shall show that $f(Y)$ contains the ball $B(p, \frac{1}{2}\delta)$ with centre p and radius $\frac{1}{2}\delta$: clearly this will prove that $f(Y)$ is the whole interior of P, and so $f(\bar{Y}) = P$.

For convenience, we shall assume that p is the origin of R^n: this can be achieved by adding constants to the f_i. Now let $d(\lambda)$, for $\lambda \in R^n$ with $|\lambda| = 1$, be the support function of P, that is,

$$d(\lambda) = \sup_{x \in P} (\sum \lambda_i x_i) .$$

Then the distance δ from the origin to ∂P is

$$\delta = \inf_{\lambda} d(\lambda) .$$

If we put $\phi^\lambda = \sum \lambda_i f_i$ then we have

$$d(\phi^\lambda) = d(\lambda)$$

where, as before,

$$d(\phi^\lambda) = \sup_{Y} \phi^\lambda = \sup_{j} \sum \lambda_i c_{ij}$$

and c_{ij} $(i = 1, \ldots, n)$ are the coordinates of $c_j \in R^n$. Now consider the gradient flow ϕ_t^λ of ϕ^λ through y. As t goes from 0 to ∞ the function ϕ^λ increases monotonically from 0 to $d(\lambda)$. Since $\delta \leqslant d(\lambda)$ there is a unique value $t(\lambda)$ for which ϕ^λ takes the value δ. Moreover it is easy to check that $t(\lambda)$ varies continuously with λ. Now let us define a star-shaped neighbourhood U of 0 in $L(H)$ as the set of all points $r\lambda$ with $|\lambda| = 1$ and $r \leqslant t(\lambda)$. Then $(\exp U)y$ defines a neighbourhood V of 0 in $f(Y)$ and for any point $v \in \partial V$ we have

$$\sum \lambda_i v_i = \tfrac{1}{2}\delta \text{ for some } \lambda \text{ with } |\lambda| = 1 .$$

In particular $|v| \geqslant \frac{1}{2}\delta$ and so $V \supset B(0, \frac{1}{2}\delta)$, as required.

We have now proved most of part (a) of Theorem 2, namely that $f(\bar{Y}) = P$. To proceed further we have to examine $\bar{Y}$ more closely and for this we shall need to use more about gradient flows. More precisely we shall need the following facts which hold for any function ϕ which is non-degenerate in the sense of Bott. They are fairly routine extensions of standard results for non-degenerate functions in the sense of Morse (for the local theory see [**4**, p. 244]).

(3.3) For any $x \in M$ the gradient flow $\phi_t(x)$ has a unique limit point $\phi_\infty(x)$ as $t \to \infty$.

(3.4) The set of all points x in M for which $\phi_\infty(x)$ lies on a given (connected) critical manifold N forms a submanifold N^u such that

(i) ϕ restricted to N^u has N as a non-degenerate maximum.

(ii) $x \to \phi_\infty(x)$ gives a continuous map $N^u \to N$.

The manifold N^u is called the "unstable manifold" of ϕ at N. Note that, from (3.4)(i), points x of N^u for which $\phi(N) - \phi(x) \leqslant \varepsilon$ form a neighbourhood of N in N^u. This clearly implies

(3.5) If N_1 is another critical manifold of ϕ with $\phi(N_1) = \phi(N)$, then N_1 is not contained in the closure of N^u.

If C is any subset of N^u, $\bar{C}$ its closure in M and $\bar{C}(N^u)$ its closure in N^u, then N being compact implies that $\bar{C} \cap N = \bar{C}(N^u) \cap N$. Hence (3.4)(i) implies

(3.6) If C_∞ is the image of C under the map $N^u \to N$ then $\bar{C} \cap N = \bar{C}_\infty$.

With these general results about gradient fields assumed, we now return to the study of our orbit Y in M, and the map $f: M \to R^n$. For any $\phi = \sum \lambda_i f_i$ we consider its gradient flow ϕ_t. Fixing some $y \in Y$ we then have $\phi_\infty(y)$ in some critical manifold N of ϕ. Since $\operatorname{grad}\phi$ commutes with the action of T_c the manifolds N and N^u are invariant under T_c. Hence $Y \subset N^u$, and $Y_\infty \subset N$ is the T_c-orbit of y_∞ (and so is independent of our original choice of $y \in Y$). By (3.5) and (3.6) with $C = Y$ we have

$$\bar{Y} \cap N_1 = \varnothing, \quad \bar{Y} \cap N = \bar{Y}_\infty \tag{3.7}$$

where N_1 is any other critical manifold of ϕ with $\phi(N_1) = \phi(N)$. We shall now show that (3.7), for all ϕ, gives us all the information we need about $\bar{Y}$.

Let us first note that on an orbit Hy of the non-compact part H of T_c, f is a diffeomorphism onto its image (which we know is the interior of P). Clearly f is a local diffeomorphism, since it has rank n: it is globally injective because $\phi = \sum \lambda_i f_i$ is strictly monotonic along the orbits of $H\phi$. Thus on Y the map $f: Y \to P$ is a bundle projection with group T/Γ (Γ finite). This implies that $f(\partial Y) \subset \partial P$. Hence if $z \in \partial Y$, $f(z) \in \partial P$ and so lies on a support hyperplane S with equation $\sum \lambda_i x_i = d$, say. The intersection $S \cap \partial P$ is a "face" $\bar{\sigma}$ of ∂P and is the convex hull of the vertices of P lying in S. We can always choose S so that $f(z)$ belongs to the *interior* σ of this face (note that σ can have any dimension from 0 to $n-1$). If P lies in the half-space $\sum \lambda_i x_i \leqslant d$ then $\phi = \sum \lambda_i f_i$ has d as its supremum on $\bar{Y}$, and $\phi(z) = d$. Hence z lies on some critical manifold of ϕ at level d. Applying (3.7) we see that $z \in \bar{Y}_\infty$ is in the closure of the orbit $Y_\infty \in N$. Under f, $Y_\infty \to \sigma$ and, by the same argument as used earlier, $\partial Y_\infty \to \partial\sigma$. Since $f(z)$ was in σ (and not in its boundary) it follows that $z \in Y_\infty$. Thus the map $f: \bar{Y} \to P$ has $f^{-1}(\sigma)$ a single orbit Y_σ. Since we have earlier proved that $f(\bar{Y}) = P$ it follows that $\bar{Y}$ is the disjoint union of orbits Y_σ, one for each face σ of P, establishing Theorem 2(b). As with the map on Y the map $f: Y_\sigma \to \sigma$ is a bundle map for the appropriate torus group. This establishes part (c) of Theorem 2 and also shows that if $\dim\sigma > 0$ then Y_σ contains no fixed point of T. This proves that the points $c_1, \ldots, c_p$, which were the images $f(Z_j)$ where Z_j was a fixed component of T meeting $\bar{Y}$, coincide with the vertices (i.e. the extreme points) of P. This finishes the proof of part (a) and thus of the whole of Theorem 2.

Note that our proof has also shown that $Y_\sigma = \phi_\infty(Y)$ where ϕ is any linear combination $\sum \lambda_i f_i$ such that $\bar{\sigma} = P \cap S$ (where as before S is the hyperplane $\sum \lambda_i x_i = d$ and P is in the half space $\sum \lambda_i x_i \leqslant d$).

If M is a projective algebraic variety then Mumford and others [**8**] have studied the orbit-closure $\bar{Y}$ and shown that it has basically the structure described above. Their procedure is to study the affine case first and then to glue things together. The polyhedral cone associated in [**8**] to $\bar{Y}$ is abstractly dual to our P: it lies canonically in $L(T)$ whereas P lies in $L(T)^*$. Naturally the proofs in [**8**] are algebraic and are replaced here by our use of gradient flows. In [**8**] the polyhedron is "rational" relative to the integer lattice in $L(T)$, but our polyhedron P is not rational for general Kähler manifolds. In fact in the homogeneous case studied by Kostant any orbit in $L(G)^*$ has a natural Kähler metric whereas only "integral" orbits correspond to projective varieties (or, in the Kirillov sense, to representations of G). Thus the convex hull of any W-orbit in $L(T)^*$ occurs as a polytope P, whereas only integral ones occur in the projective case. This suggests that P should be rational for Hodge metrics, provided we normalize the arbitrary constant in the definition of f by requiring that $\int_M f = 0$.

§4. *Examples and Applications*

We return now to the homogeneous examples arising from orbits of G in $L(G)$. Any such orbit M is of the form $G/C(T_0)$ where $C(T_0)$ is the centralizer of some subtorus $T_0 \subset T$. Thus a generic orbit is always the flag manifold G/T but special orbits are quotients of G/T. If we fix a positive Weyl Chamber $\mathscr{C}$ in $L(T)$ this fixes a Borel subgroup B of the complexification G_c containing T (and a parabolic subgroup $P \supset C(T_0)$). We can then identify G/T with G_c/B and more generally $G/C(T_0)$ with G_c/P. In this way our orbit M becomes a complex manifold, homogeneous under the action of G_c. Together with the symplectic structure this then determines a Kähler metric on M (this is *not* the same as the metric induced from a G-invariant Euclidean metric on $L(G)$, see the Remark below). Thus we can apply Theorem 2 to describe the image $f(\bar{Y})$ for various T_c-orbits in M. We shall use this to describe the images under f of the *Bruhat cells* of M. We recall that these are the B-orbits in M and they correspond to the double cosets $B\omega P$ where ω runs over representatives for W/W_0, where W, W_0 are the Weyl groups of G and $C(T_0)$. We write M_ω for the cell $B\omega P$ and $[\omega]$ for the "point" ωP in M. Since M_ω is a B-orbit it is invariant under T_c. Moreover M_ω has a natural linear structure arising from the nilpotent part of B, the action of T_c is then linear and the weights of the representation are a subset of the positive roots. Now let $\lambda \in L(T)$ be an interior point of the positive Weyl chamber $\mathscr{C}$. Then $\alpha(\lambda) > 0$ for all positive roots α and so any orbit of $\exp 2\pi i(\mu + it\lambda)$ in M_ω tends to $[\omega]$ as $t \to \infty$ (μ being any element of $L(T)$). But such an orbit is a trajectory of grad ϕ where ϕ is the function corresponding to λ, that is, $\phi(x) = \langle x, \lambda \rangle$ for $x \in M \subset L(G)$. Hence M_ω is in the unstable manifold of ϕ at ω. Since this holds for all ω it follows that M_ω *coincides with this unstable manifold.*

There is a natural partial ordering on Bruhat cells, and so on W/W_0, in which

(4.1) $$\omega' \geqslant \omega \Leftrightarrow M_{\omega'} \subset \bar{M}_\omega)$$

By the same argument as used in Section 3, we see that

(4.2) $$f(\bar{M}_\omega) \subset \text{convex hull of the set of all } f(\omega') \text{ with } \omega' \geqslant \omega\,,$$

where f is orthogonal projection of $L(G)$ onto $L(T)$. To show that we have equality it is then only necessary, by Theorem 2, to prove that there is some point $y \in M_\omega$ with T^c-orbit Y, so that $[\omega'] \in \bar{Y}$ for all $\omega' \geqslant \omega$. Now there is just one cell M_τ of maximal dimension and its complement is an algebraic subvariety. Translating this cell by ω' we get a new cell of maximal dimension (for the Bruhat decomposition of a different Borel subgroup) centred on $[\omega']$. Points of M_τ converge to $[\tau]$ under the flow ϕ_t as $t \to \infty$. Hence points of $\omega'\tau^{-1} M_\tau$ converge to $[\omega']$ under the flow ψ_t as $t \to \infty$ where ψ is obtained from ϕ by the action of $\omega'\tau^{-1}$. Now by hypothesis M_ω gets arbitrarily close to $[\omega']$, hence it has non-zero intersection $M_\omega(\omega')$ with the open neighbourhood $\omega'\tau^{-1}M_\tau$. Since all these are algebraic varieties, $M_\omega(\omega')$ is Zariski open in M_ω. Hence the intersection of all $M_\omega(\omega')$ for $\omega' > \omega$ is non-empty. Take y to be in this intersection. Then by construction we see that for each $\omega' > \omega$ we have $\psi_t(y) \to [\omega']$ as $t \to \infty$ for some ψ. Hence $[\omega'] \in \bar{Y}$ as required. Thus we have proved the following refinement of Kostant's result.

THEOREM 3. *Let M be an orbit of G on $L(G)$, x the unique point of $M \cap L(T)$ in the positive Weyl chamber and M_ω the Bruhat cell determined by $\omega \in W$. Then the image of $\bar{M}_\omega$ under orthogonal projection to $L(T)$ is the convex hull of the set of all points $\omega'(x)$ with $\omega' \geqslant \omega$.*

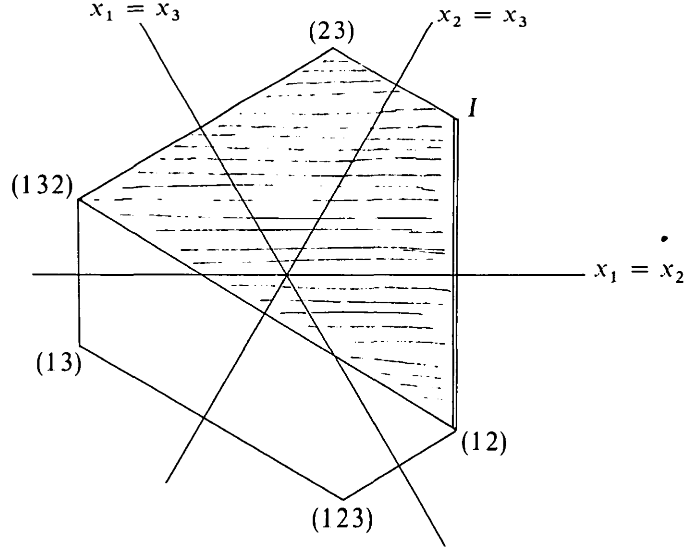
NOTE. This result has also been obtained by G. Heckman in his Leiden thesis.

The partial ordering defined geometrically in (4.1) can be given a direct definition in Weyl group terms. From our discussion it clearly follows that

(4.3) $$\omega' \geqslant \omega \Rightarrow \langle \lambda, \omega'(\zeta)\rangle \leqslant \langle \lambda, \omega(\zeta)\rangle \quad \text{for all } \lambda \in \mathscr{C}.$$

Here ζ is any point of $\mathscr{C}$ with W_0 as isotropy group. In fact it is well known and not hard to prove directly that the converse of (4.3) also holds. Note that the partial ordering on W/W_0 is induced by that on W, corresponding to the fact that the Bruhat cells of G_c/P are projections of those of G_c/B.

The simplest non-trivial example of Theorem 3 arises for $G = SU(3)$. The flag manifold has complex dimension 3, there are 6 Bruhat cells indexed by the elements of Σ_3. The trivial element gives a point, the transposition (13) gives the 6-cell (using the usual ordering to define the positive Weyl chamber), leaving two 2-cells corresponding to (12), (23) and two 4-cells corresponding to the 3-cycles (123), (321). In the diagram below the shaded region describes the image of the 4-cell (132), while the double line is the image of the 2-cell (12). The other images are obtained by interchanging the roles of 1 and 3.



REMARKS. 1. In the proof of Theorem 3 we used strongly the fact that there is one dense open Bruhat cell, or equivalently that each generic ϕ has a unique maximal component. On any Kähler manifold this is guaranteed by Morse theory as in Section 2. Thus deducing Theorem 1 from Theorem 2 still requires Morse theory.

2. It is perhaps helpful to make a comment on the various metrics which arise on the orbits M in $L(G)$. For simplicity we consider the general case when $M \cong G/T$, but our conclusion will apply to all cases. Any G-invariant metric on M is determined by a T-invariant metric on $L(G)/L(T) \cong \oplus E_\alpha$, where the E_α are the 2-dimensional root spaces. Thus the metric is given by choosing metrics on each E_α. There is a natural metric arising from any G-invariant metric on $L(G)$—the Killing form if G is semi-simple gives a preferred choice. Relative to this standard metric any other metric is defined by positive scalars ρ_α. Suppose now that M is the G-orbit of $x \in L(T)$ chosen in the positive Weyl chamber. Then for every positive root we have $\alpha(x) \geqslant 0$ and $\alpha(x) > 0$ if x is an interior point. Then simple computation shows that

(i) for the Kähler metric on M, $\rho_\alpha = \alpha(x)$;

(ii) for the metric on M, induced from $L(G)$, $\rho_\alpha = \alpha(x)^2$.

Alternatively we can say that, relative to the standard metric on M, any other one is given at each point by a positive self-adjoint operator on the tangent space. Then if A, B are the operators for the metrics (i) and (ii) respectively, we have $B = A^2$.

Finally we shall discuss the relation of our results to recent work of Gel'fand and Macpherson [3]. First let us observe that Theorem 2 can be reformulated in terms of the orbits of the non-compact part H of the complex torus T_c, to give

THEOREM 2'. *The map* $f: M \to R^n = L(H)^*$ *induces an* H*-homeomorphism* $\overline{Hy} \to P$.

Consider in particular the special case when M is the Grassmannian of complex p-subspaces in C^{p+q}. We take for T the diagonal matrices in $U(p+q)$. Then H is the group of positive real diagonal matrices. Clearly H acts on M_R, the real (unoriented) Grassmannian of p-subspaces in R^{p+q}, and so, taking $y \in M_R$, Theorem 2' gives us the structure of $\overline{Hy}$ in M_R. This is what Gel'fand and Macpherson call a hypersimplex. One of their main results is that *(for pq even) any* $O(p+q)$*-invariant differential form* θ *on* M_R *restricts to zero on any* H*-orbit.* Their proof is by direct computation and we shall now show that this result can be proved quite naturally and more generally in our context, without computation. The key idea is that one should work with the complex Grassmannian, exploiting to the full its Kähler structure. The result we shall prove is

THEOREM 4. *Let* M *be a compact connected Kähler manifold,* T *a toral group of automorphisms of* M *having at least one fixed point. Write* $T^c = T \times H$ *so that* $L(H) = JL(T)$. *Then* T^c *acts on* M *and, for* $l > 0$, *every harmonic* l*-form* θ *on* M *restricts to zero on any* H*-orbit.*

Consider any T-orbit Ty. Since T has a fixed point (and M is connected) the map $\gamma: T \to M$ given by $t \to ty$ is homotopic to a constant. Hence, for $l > 0$, the induced map on H^l is zero. Hence $\gamma^*(\theta)$ represents zero in $H^l(T)$. On the other hand

T preserves all harmonic forms on M and in particular θ. Hence $\gamma^*(\theta)$ is an invariant form on T. But an invariant form representing zero must itself be zero. Hence $\gamma^*(\theta) \equiv 0$ on T and so θ restricts to zero on the orbit Ty. Now we may, without loss of generality, assume that θ is a pure (r, s)-form with $r+s = l$, (because on a Kähler manifold the (r, s) components of a harmonic form are themselves harmonic). This implies that θ is an eigenvector of J: $J\theta = i^{r-s}\theta$. Now in the tangent space to M at y the H-directions are obtained from the T-directions by applying J. Since θ restricts to zero on Ty, $J\theta$ restricts to zero on Hy. But $J\theta = i^{r-s}\theta$ and so θ restricts to zero on Hy as required.

To deduce the special case of Gel'fand–Macpherson, we note that on any homogeneous symmetric space, such as the real or complex Grassmannians, invariant forms are harmonic. Finally, we need to observe that the map $H^l(M, R) \to H^l(M_R, R)$, induced by the inclusion of the real Grassmannian M_R into the complex Grassmannian M, is surjective for pq even : $H^*(M_R)$ is generated by the Pontrjagin classes and $H^*(M)$ by the Chern classes. Note that the surjectivity of cohomology is not true if pq is odd or if we consider *oriented* Grassmannians, and in these cases there are indeed invariant forms which are not zero on H-orbits.

The hypothesis in Theorem 4 that T have at least one fixed-point is of course satisfied for the Grassmannian. More generally it holds whenever X is *algebraic* and T_c acts *algebraically*. As noted in [2] it is also the condition which guarantees in general the existence of the Hamiltonian functions $f_1, \ldots, f_n$. Thus Theorem 4 and Theorem 2′ are both available in this context and so the integral-geometry formulae of [3] could in principle be generalized to other Kähler manifolds of the type considered here.

Our proof of the Gel'fand–Macpherson result indicates the advantage of looking at the real Grassmannian as the real part of the complex Grassmannian. In the same spirit one could extend Theorem 2′ to suitable real algebraic manifolds.

It is perhaps worth making a final remark in connection with Theorem 2′. If we introduce the natural coordinates $(y_1, \ldots, y_n)$ on the orbit Hy coming from the Lie algebra then we find that

$$\frac{\partial f_i}{\partial y_j} = g_{ij},$$

where $g_{ij}dy^i dy^j$ is the metric on the orbit induced by the Kähler metric on M. In particular since $g_{ij} = g_{ji}$ we see that

$$f_i = \frac{\partial \rho}{\partial y_i}$$

for some function ρ. This function is strictly convex and its graph has, by Theorem 2′, an asymptotic polyhedral cone with base ∂P.

Locally any Kähler metric can be written as $id'd''\rho$. In our situation we get a natural choice for ρ along an H-orbit, unique up to a constant and linear terms, and our map f is just $d\rho$.

References

1. R. Bott, 'Non-degenerate critical manifolds', *Ann. of Math.*, 60 (1954), 248–261.
2. T. Frankel, 'Fixed points on Kähler manifolds', *Ann. of Math.*, 70 (1959), 1–8.
3. I. M. Gel'fand and R. MacPherson, 'Geometry in Grassmannians and a generalization of the dilogarithm', preprint, Institut des Hautes Études Scientifiques.
4. P. Hartman, *Ordinary differential equations* (Wiley, New York, 1964).
5. A. Horn, 'Doubly stochastic matrices and the diagonal of a rotation matrix', *Amer. J. Math.*, 76 (1954), 620–630.
6. A. A. Kirillov, *Elements of the theory of representations* (Springer, Berlin, 1976).
7. B. Kostant, 'On convexity, the Weyl group and the Iwasawa decomposition', *Ann. Sci. Ec. Norm. Sup.*, 6 (1973), 413–455.
8. D. Mumford, G. Kempf, F. Knudsen and B. Saint-Donat, *Toroidal embeddings I*, Lecture Notes in Mathematics 339 (Springer, Berlin, 1973).
9. I. Schur, 'Über eine Klasse von Mittelbildungen mit Anwendungen auf der Determinantentheorie', *Sitzungsberichte der Berliner Mathematischen Gesellschaft*, 22 (1923), 9–20.

Mathematical Institute,
University of Oxford.