Convexity and integrability

A.M. Bloch[*] and T.S. Ratiu[**]

[*] Partially supported by NSF Grant DMS 9002136 and a Seed Grant from the Ohio State University

[**] Partially supported by NSF Grant DMS 8922699

[*] Dept. of Mathematics
The Ohio State University
Columbus, OH 43210
USA

[**] Dept. of Mathematics
University of California
Santa Cruz, CA 95064
USA

[**] Max–Planck–Institut
für Mathematik
Gottfried–Claren–Str. 26
5300 Bonn 3
FRG

MPI/90-103

B^x_n, daß $(u_n, v_{n,m}) < \frac{1}{n}$ ist. Ebenso kommt $(u_m, v_{n,m}) < \frac{1}{m}$, also $(u_n, u_m) < \frac{1}{n} + \frac{1}{m}$, d.h. die u_n bilden eine Fundamentalfolge. Wegen der Vollständigkeit von E konvergieren die u_n gegen einen Punkt $u_\omega \subset E$. Folglich müssen die x_n gegen den Bildpunkt $x_\omega \subset A$ von u_ω konvergieren. Die x_n konvergieren aber evidentermassen gegen ξ, also ist $\xi = u_\omega \subset A$,

$$\prod_{n=1}^{\infty} G_n \subset A \quad ; \quad \prod_{n=1}^{\infty} G_n = A$$

(w.z.b.w.)

Es wäre noch interessant zu entscheiden, ob die absolute Definition der G_δ (mit Hilfe der „geschlossenen Umgebungssysteme" = „système determinant clos") auch auf den allgemeinen Fall (aller vollständigen Räume) ausgedehnt werden kann. In diesem Ideenkreise ist kürzlich von P. Alexandroff eine absolute (intrinsèque) Definition aller Klassen der B-Mengen (in vollständigen Räumen mit abzählbarer dichter Teilmenge) gefunden worden, die aber naturgemäß viel komplizierter ist.

Vielleicht wird Sie auch der folgende Urysohnsche Metrisationssatz interessieren:

Convexity and integrability

A. M. Bloch* and T. S. Ratiu**

Abstract

In this review we present the main convexity results of Atiyah, Duistermaat, Guillemin, Kostant, and Sternberg and their relationship to the dynamic behavior of the generalized non-periodic Toda lattice. A "symplectic" proof of one of Kostant's convexity results at group level is also sketched. This review is an expanded version of the talk one of us (T.S.R.) presented at this conference and is based upon the papers, Bloch, Brockett and Ratiu [1990 a,b], Bloch, Flaschka and Ratiu [1990], and Lu and Ratiu [1990].

§1. Introduction

In symplectic and Poisson geometry one of the main objects of study is the momentum map associated to a Hamiltonian action. This map, with values in the dual of the Lie algebra of symmetries, was introduced in its modern formulation by Kirillov (see Kirillov [1976]), Kostant [1970], Souriau [1970] and Smale [1970] and it turns out to be at the center of many important geometrical facts that are useful in a variety of fields of both pure and applied mathematics. One of the most striking aspects of the momentum map are its convexity properties as formulated in the now famous theorems of Atiyah [1982], Guillemin and Sternberg [1982], [1984] and Kirwan [1984]. No serious relationship of this property with the dynamics of Hamiltonian systems has been forged so far even though already a superficial familiarity with these theorems suggest links to dynamic bifurcation behavior.

A specific branch of Hamiltonian dynamics, namely the theory of completely integrable systems, seems particularly suited to a study via the convexity theorems, since the hierarchy of flows comprise the action of a cylinder (often just a torus) whose momentum map

*Department of Mathematics, The Ohio State University, Columbus, OH 43210. Partially supported by NSF Grant DMS 9002136 and a Seed Grant from the Ohio State University.

**Department of Mathematics, University of California, Santa Cruz, CA 95064. Partially supported by NSF Grant DMS 8922699

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

is simply the collection of integrals of motion. To the disappointment of many symplectic geometers, the theory of integrable systems, which interacted so fruitfully both with algebraic geometry and representation theory, had no impact in its recent development on its own field of origin, namely symplectic geometry itself. It may turn out that the reasons are simply in the subtlety of such an interaction. In this review we want to outline several results in this direction, some of which we came across simply by accident. We shall deal here only with one class of examples: the non-periodic generalized Toda lattices. Judging from this example the convexity properties of the momentum map have a very subtle but decisive influence upon the dynamics of completely integrable systems and this whole relationship is worthwhile exploring both in general, as well as in specific examples.

The review is organized as follows. §2 is simply a collection of standard defintions and properties from Poisson geometry and the theory of momentum maps. Its sole purpose is to fix notation and conventions for the rest of the work. §3 reviews the relevant symplectic convexity results and §4 discusses their extensions in the important case of Kähler manifolds. The different metrics on orbits are also discussed since they are needed later. §5 introduces an abstract gradient system first studied in the $su(n)$-case by R. Brockett [1988]. This gradient system lives on adjoint orbits of compact Lie algebras and the dynamic behavior of its flow is entirely determined by the momentum map on this orbit. §6 shows that for a special choice of parameters, this gradient system becomes the Toda lattice on its isospectral set. Thus the famous scattering behavior of the Toda flow, so useful in numerical analysis, is simply an inheritance from its "partner", the so-far hidden gradient system. §7 discusses this relationship on a deeper level: since the Toda hierarchy is integrable, it naturally defines a non-compact torus action on an adjoint orbit which is canonically a Kähler manifold. Then why does the isospectral set not have a convex image under the momentum map? The answer is because the action is not the diagonal one. Can one embed the isospectral set in a different way so that its image under the projection is convex? The answer is positive and discussed in §7. Finally, §8 presents a symplectic proof of one of Kostant's non-linear convexity results. The relationship with the Toda flows is not clear here but several formulas and constructions are very similar and suggestive for the standard results one finds in the non-periodic Toda-lattice theory. We hope that a more serious link can be found. Also, the result in §8 suggests several generalizations.

As mentioned already at the beginning, this paper is a review. We hereby would like to thank our coauthors R. Brockett, H. Flaschka, and J.-H. Lu without whom the results presented here would not have been possible. §5 and §6 are based on Bloch, Brockett, Ratiu [1990a,b], §7 on Bloch [1990] and Bloch, Flaschka, Ratiu [1990], and §8 on Lu, Ratiu [1990]. These papers drew inspiration from the papers Brockett [1988, 1989]. Conversations with B. Kostant, J. Marsden, A. Reyman and A. Weinstein throughout the period the above mentioned papers were written have been of indispensable importance. Finally, T. Ratiu would like to thank the Max Planck Institute for Mathematics, Bonn, Germany, where part of this review was written, and the organizers of this conference for the opportunity to participate and enjoy both the talks and the lively discussions.

§2 Poisson manifolds and momentum maps.

The only purpose of this section is to introduce notations, specify conventions, and spell out definitions of concepts used throughout this review. Anyone familiar with this material can safely skip it.

A *Poisson manifold* is a pair $(P, \{,\})$ formed by a smooth manifold P whose ring of functions $\mathcal{F}(P)$ is a Lie algebra relative to an operation $\{,\}$ called a *Poisson bracket* which is a derivation in each argument. Thus, for fixed $H \in \mathcal{F}(P)$, the formula $F \in \mathcal{F}(P) \mapsto \{F, H\} \in \mathcal{F}(P)$ determines a vector field X_H called the *Hamiltonian vector field* determined by H, i.e. $\langle dF, X_H \rangle = \{F, H\}$. There are two main examples of Poisson manifolds:

a) *Symplectic manifold,* which are manifolds P carrying a closed non-degenerate two-form ω. Defining X_H by the requirement that $\omega(X_H, Y) = \langle dH, Y\rangle$ for any vector field Y on P and setting $\{F, H\} = \omega(X_F, X_H)$ one easily proves that (P, ω) becomes a Poisson manifold.

b) *Duals of Lie algebras.* Let $\mathfrak{g}$ be a Lie algebra and let $\mathfrak{g}^*$ be the space of all linear functionals on $\mathfrak{g}$. For $F : \mathfrak{g}^* \to \mathbf{R}$ define the *functional derivative* $\delta F/\delta\mu \in \mathfrak{g}$ by the requirement

$$\mathbf{D}F(\mu) \cdot \nu = \langle \nu, \frac{\delta F}{\delta \mu} \rangle$$

for all $\nu \in \mathfrak{g}^*$; $\mathbf{D}F$ denotes the usual Frechet derivative of F. $\mathfrak{g}^*$ is a Poisson manifold relative to the *Lie-Poisson structure*

$$\{F, H\}_{\pm}(\mu) = \pm\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu}\right]\rangle.$$

A key property satisfied by the Poisson bracket on $(P, \{,\})$ is that $F \mapsto X_F$ is a Lie algebra anti-homomorphism:

$$X_{\{F,H\}} = -[X_F, X_H]$$

for all $F, H \in \mathcal{F}(P)$. Poisson manifolds stratify into symplectic manifolds called the *symplectic leaves* of P. They are described in the following way. On P define an equivalence relation which identifies two points if they can be joined by a piecewise smooth path, each segment of which is a trajectory of a locally defined Hamiltonian vector field. The equivalence classes are connected immersed Poisson submanifolds (i.e. a Hamiltonian vector field on P at a point on such a submanifold is tangent to it) with the Poisson bracket defined by a symplectic form. The tangent space at p to the leaf containing p is $\{X_H(p) | H \in \mathcal{F}(P)\}$. We refer to Weinstein [1983] for more information on Poisson manifolds. If P is symplectic, its symplectic leaves are its connected components. If $P = \mathfrak{g}^*$, its leaves are the connected components of the coadjoint orbits of the underlying Lie group G of $\mathfrak{g}$. The symplectic form on these orbits is the *Kirillov-Kostant-Souriau form,*

$$\omega(\mu)(ad^*_\xi \mu,\ ad^*_\eta \mu) = \pm\langle \mu, [\xi, \eta]\rangle$$

for $\xi, \eta \in \mathfrak{g}$, $\mu \in \mathfrak{g}^*$, μ on the orbit in question.

Let G be a Lie group acting on $(P, \{,\})$ in a canonical fashion, i.e., preserving the Poisson bracket. For $\xi \in \mathfrak{g}$ denote by ξ_P the infinitesimal generator of the action, i.e.

$$\xi_P(p) = \frac{d}{dt}\Big|_{t=0} (\exp t\xi) \cdot p$$

where $\exp : \mathfrak{g} \to G$ is the exponential map and dot denotes the action. Since we consider left actions, $\xi \mapsto \xi_P$ is an anti-homomorphism, i.e.

$$[\xi, \eta]_P = -[\xi_P, \eta_P]$$

for all $\xi, \eta \in \mathfrak{g}$. We shall say that the G-action on P is *Hamiltonian* if there is a Lie algebra homomorphism $\rho : \mathfrak{g} \to \mathcal{F}(P)$ factoring $\xi \mapsto \xi_P$ through $H \mapsto X_H$, i.e.

$$\xi_P = X_{\rho(\xi)}$$

for all $\xi \in \mathfrak{g}$. If P is symplectic and G is semisimple such a homomorphism always exists; this is a corollary of the two Whitehead lemmas. The map $J : P \to \mathfrak{g}^*$ given by $\langle J(p), \xi \rangle = \rho(\xi)(p)$ for all $\xi \in \mathfrak{g}$, $p \in P$ is called the *momentum map* of the action. We shall denote $\rho(\xi) = J_\xi$ and therefore we have

$$J_{[\xi,\eta]} = \{J_\xi, J_\eta\}$$

for all $\xi, \eta \in \mathfrak{g}$.

§3 A brief review of symplectic convexity results.

Some of the most striking global properties of the momentum map are its convexity properties. In this section we briefly review the main results in the symplectic context.

Theorem 3.1. *(Atiyah [1982], Guillemin and Sternberg [1982]). Let P be a compact connected symplectic manifold and T a torus acting in a Hamiltonian fashion on P with momentum map $J : P \to \mathbf{R}^n$, $n = \dim T$. Let P^T denote the fixed point set of T. Then:*

(i) *$J(P^T)$ is finite and each point is the image of a connected compenent of the set where the derivative of J vanishes;*

(ii) *each fiber $J^{-1}(c)$, $c \in \mathbf{R}^n$, if non-empty, is connected;*

(iii) *$J(P)$ is the convex hull of $J(P^T)$*

This theorem generalizes to the symplectic setting the following remarkable result of Kostant:

Corollary 3.1. *(Kostant [1973]) Let G be a connected semisimple or a compact connected Lie group, T a maximal torus. Let $\mathfrak{g}, \mathfrak{t}$ denote the Lie algebras of G and T respectively.*

Fix a G-invariant metric on $\mathfrak{g}$. *Then the orthogonal projection of an adjoint orbit* $\mathcal{O}$ *onto* $\mathfrak{t}$ *is the convex hull of the corresponding Weyl group orbit* $\mathcal{O} \cap \mathfrak{t}$.

Indeed, by the G-invariant metric, adjoint and coadjoint orbits are identified and the orthogonal projection is just the momentum map of the T-action on $\mathcal{O}$. The fixed point set of the T-action on $\mathcal{O}$ is the Weyl group orbit $W \cap \mathcal{O}$ and the corollary is thus a direct consequence of Theorem 2.1. Kostant's result is itself a generalization of the following classical result:

Corollary 3.2. *(Schur [1923], Horn [1954]) Denote by* $\{a\}$ *the set of diagonals of all Hermitian* $n \times n$ *matrices with given eigenvalues* $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$. *Let the symmetric group* S_n *act on* $\mathbf{R}^n$ *by permutation of coordinates. Then* $\{a\}$ *is the convex hull of* $S_n \cdot \boldsymbol{\lambda}$, *the* S_n*-orbit through* $\boldsymbol{\lambda}$.

This follows from Corollary 3.1 by taking $G = U(n)$. Going in the opposite direction, Kirwan has generalized Theorem 3.1 to compact connected Lie groups.

Theorem 3.2. *(Kirwan [1984]) Let* P *be a compact connected symplectic manifold and let the compact connected Lie group* G *act in a Hamiltonian fashion on* P *with momentum map* $J : P \to \mathfrak{g}^*$. *Let* T *be a maximal torus in* G, $\mathfrak{t}$ *its Lie algebra and let* $\mathfrak{t}^*_+$ *denote the positive Weyl chamber relative to a fixed ordering. Then*

(i) *each fiber* $J^{-1}(\mu)$, *if non-empty, is connected;*

(ii) $J(P) \cap \mathfrak{t}^*_+$ *is convex.*

Theorems 3.1 and 3.2 can be proved together by a method different from the ones used in the original papers; this is done in Condevaux and Molino [1988]. A beautiful survey of applications of theorem 2.1 can be found in Atiyah [1983].

The question naturally arises whether the image of the momentum map characterizes the action. For so-called completely integrable toral actions the answer is positive. A torus action on a symplectic manifold P is called *completely integrable* if the action is Hamiltonian, faithful, and $\dim P$ is twice the dimension of the torus.

Theorem 3.3. *(Delzant [1988]) Let P_1, P_2 be two compact connected symplectic manifolds of dimension $2n$ and T an n-dimensional torus acting in a completely integrable fashion on both P_1 and P_2 with momentum maps J_1 and J_2 respectively. If $J_1(P_1) = J_2(P_2)$ then there is a T-equivariant symplectic diffeomorphism $\varphi : P_1 \to P_2$ such that $J_2 \circ \varphi = J_1$.*

Theorem 3.3 is a companion to the Theorem 3.1. The correspnding companion to Theorem 3.2 is not known in full generality but there are results for groups of low rank by Delzant.

In order to deal with the analog of projections of real flag manifolds as opposed to complex ones, Duistermaat has proved the following:

Theorem 3.4. *(Duistermaat [1983]) Let (P, ω) be a compact connected symplectic manifold on which the torus T acts in a Hamiltonian fashion with momentum map J. Assume that $\tau : P \to P$ is an antisymplectic involution (i.e. $\tau \circ \tau =$ identity and $\tau^* \omega = -\omega$) which leaves J invariant (and so it will necessarily invert the toral action: $\tau \circ g \circ \tau^{-1} = g^{-1}$ for all $g \in T$). Let Q be the fixed point set of P which is assumed to be non-empty. Then*

(i) $J(Q) = J(P)$

(ii) *$J(Q)$ is the convex hull of the finitely many points $J(P^T \cap Q)$, where P^T is the fixed point set of the T-action on P;*

(iii) *$P^T \cap Q = \{p \in Q |$ the derivative of $J|Q$ at p vanishes$\}$; each connected component of this set equals the intersection of a connected component of P^T with Q; it is mapped by $J|Q$ into one point;*

(iv) *all results above hold if Q is replaced by one of its connected components.*

There are many other convexity results in the literature. We have deliberately avoided the infinite dimensional case and those theorems for which a symplectic analog is not yet known. The notable exception is one of the non-linear convexity results of Kostant [1973] dealing with the Iwasawa projection which will be reviewed in a more natural setting in the last section of this paper; a "symplectic proof" for it will also be sketched. It is a challenge, in view of this, to try to "symplecticize" the great variety of convexity results available. The Kähler case is special and very important. It will be dealt with in the next section.

§4 The Kähler convexity theorem and coadjoint orbits.

The convexity results reviewed in §3 hold also for Kähler manifolds, since every Kähler manifold is symplectic. However, due to the extra structure there is an important refinement, particularly relevant to integrable systems, proved by Atiyah [1982]; see also Atiyah [1983]. In this section we review this result and discuss it also in the important context of coadjoint orbits of compact Lie groups. In addition, we will briefly recall the relevant facts about the possible invariant metrics on these orbits; this material is needed in §7.

On a *Kähler manifold* one has a Riemannian metric g, a symplectic form ω, and an integrable almost complex structure J (i.e. $\nabla \mathsf{J} = 0$ where ∇ is the covariant derivative of the Riemannian connection defined by g). We have $\mathsf{J}^{-1} = \mathsf{J}^* = -\mathsf{J}$ the adjoint being taken relative to g. The relationship between grad f and X_f, where X_f is the Hamiltonian vector field corresponding to $f : P \to \mathbf{R}$, is the following:

$$grad\, f = \mathsf{J} X_f.$$

This follows from the interdependence of (g, ω, J) established by the formula

$$\omega(u, \mathsf{J}v) = g(u, v)$$

for all $u, v \in T_pP$, $p \in P$. Any two of (g, ω, J) uniquely determine the third. We have $\mathsf{J}v = iv$ for all $v \in TP$.

Assume a torus T acts on the Kähler manifold P preserving all structures. Thus the action of T consists of holomorphic mappings. Since the automorphism group of a complex manifold is a complex Lie group, the T-action extends to a holomorphic action of the complexification $T_{\mathbf{C}}$ of T. The Lie algebra $\mathfrak{t}_{\mathbf{C}}$ of $T_{\mathbf{C}}$ is $\mathfrak{t} \oplus \mathfrak{a}$, where $\mathfrak{t}$ is the Lie algebra of T and $\mathfrak{a} = i\mathfrak{t}$. Define $A = \exp \mathfrak{a} \subset T_{\mathbf{C}}$; it is a vector Lie subgroup of $T_{\mathbf{C}}$ since exp is injective on $\mathfrak{a}$. Moreover $(t, a) \in T \times A \mapsto ta \in T_{\mathbf{C}}$ is a diffeomorphism. We phrase the following theorem both in terms of complex toral orbits and in terms of the noncompact real toral orbits. As usual, the standing assumption is that P *is compact and connected and that the* T*-action is Hamiltonian with momentum map* $J : P \to \mathfrak{t}^*$.

Theorem 4.1. *(Atiyah [1982]) (i) Let X be the closure in P of the complex toral orbit $T_{\mathbf{C}} \cdot p$, $p \in P$. The set of fixed points of the T-action in X, $P^T \cap X$, equals the intersection of X with the set of points in P where the derivative of J vanishes. Each connected component of this set is mapped by J to one point in $\mathfrak{t}^*$. The image $J(X)$ equals the convex hull of these points which are the vertices of a convex polytope. For each open face σ of this polytope $J^{-1}(\sigma) \subset X$ consists of a single $T_{\mathbf{C}}$-orbit. Moreover, J induces a homeomorphism of X/T onto the polytope.*

(ii) Let Y be the closure in P of the A-orbit $A \cdot p$, $p \in P$. $P^T \cap Y$ equals the intersection of Y with the set of points in P where the derivative of J vanishes. Each connected component of this set is mapped by J to one point and $J(Y)$ is the convex hull of these points which are the vertices of a convex polytope. For each open face σ of this polytope, $J^{-1}(\sigma)$ consists of a single A-orbit. Moreover, J induces a homeomorphism of Y onto the polytope.

(iii) For generic X and Y as in (i) and (ii), $J(P) = J(X) = J(Y)$.

The formulation (ii) was explicitly spelled out by Duistermaat [1983] and used to generalize Kostant's theorem on the convexity of certain projections of real flag manifolds.

If G is a compact Lie group it admits a complexification $G_{\mathbf{C}}$. Any coadjoint orbit of G is of the form $G/C(T')$ where $C(T')$ is the centralizer of a subtorus T' of the maximal torus T of G. So, generically, the orbits are flag manifolds. Fix a positive Weyl chamber in $\mathfrak{t}$ and hence a Borel subgroup B of $G_{\mathbf{C}}$, $B \supset T$, and a parabolic subgroup P of $G_{\mathbf{C}}$, $P \supset C(T')$. Since $G/C(T')$ and $G_{\mathbf{C}}/P$ (generically G/T and $G_{\mathbf{C}}/B$) are diffeomorphic, one can naturally induce a complex structure on $G/C(T')$ which is homogeneous under the $G_{\mathbf{C}}$-action. Thus on each coadjoint orbit of G in $\mathfrak{g}^*$ we get a homogeneous complex structure. Together with the Kirillov-Kostant-Souriau form this makes the coadjoint orbits Kähler. Theorem 4.1 immediately implies the following:

Corollary 4.1. *(Atiyah [1982]) Let A be the non-compact part of $T_{\mathbf{C}}$, i.e. $A = \exp \mathfrak{a}$, $\mathfrak{a} = i\mathfrak{t}$, $\mathfrak{t}$ the Lie algebra of T. Relative to the invariant metric on G identify coadjoint and adjoint orbits. Let Y be the closure of an A-orbit in $\mathfrak{g}^* = \mathfrak{g}$ which lies in the (co)adjoint orbit $\mathcal{O}$ of G. If $J : \mathcal{O} \to \mathfrak{t}$ is the orthogonal projection then $J(Y)$ is the convex hull of the* *points in $\mathfrak{t}$ which are the images of the points in Y where the derivative of J vanishes. J is a homeomorphism of Y to this closed convex polytope. For generic Y, $J(Y) = J(\mathcal{O})$.*

For our later purposes it will be useful to review here several of the many natural metrics on coadjoint orbits of compact Lie groups. Let G be a compact connected semisimple Lie group and $\mathcal{O}_{\mu_0}$ an adjoint orbit through $\mu_0 \in \mathfrak{g}$. Let κ denote the Killing form on $\mathfrak{g}$. Most of the material below may be found in Besse [1987].

(i) $\mathfrak{g}$ is a Riemannian manifold with the constant metric given by the negative of the Killing form. Pull this metric back to O_{μ_0} by the inclusion. One gets the *induced metric* given by

$$\langle [\mu, \eta], [\mu, \xi] \rangle_i = -\kappa([\mu, \eta], [\mu, \xi])$$

for all $[\mu, \eta]$, $[\mu, \xi] \in T_\mu O_{\mu_0}$, $\mu \in O_{\mu_0}$.

(ii) Left translate $-\kappa(\cdot, \cdot)$ from $\mathfrak{g} = T_e G$ to $T_g G$ to obtain a left-invariant metric on G. This metric is actually bi-invariant since κ is invariant under the adjoint action. The quotient of G by the stabilizer of μ_0 inhertis in this way a Riemannian structure. Thus $\mathcal{O}_{\mu_0}$ becomes a Riemannian manifold. In Besse [1987] the metric obtained in this fashion is called the *normal metric*; Atiyah [1982] calls it the *standard metric*. To write an explicit formula, let $\mu \in \mathfrak{g}$, $\mathfrak{g}_\mu = \{\xi \in \mathfrak{g} | [\xi, \mu] = 0\}$, $\mathfrak{g}^\mu = im\, ad\, \mu\}$, and denote for $\eta \in \mathfrak{g}$ by η^μ the $\mathfrak{g}^\mu$-component of η in the direct sum decomposition $\mathfrak{g} = \mathfrak{g}_\mu \oplus \mathfrak{g}^\mu$; the two summands are orthogonal relative to $-\kappa(\cdot, \cdot)$. Then for $[\mu, \eta], [\mu, \zeta] \in T_\mu \mathcal{O}_{\mu_0}$ the normal metric has the expression

$$\langle (\mu, \eta], [\mu, \xi] \rangle_n = -\kappa(\eta^\mu, \zeta^\mu).$$

(iii) Finally there is the large family of G-invariant Kähler metrics on $\mathcal{O}_{\mu_0}$ which are in bijective correpsondence with the points of the positive Weyl chamber $\mathfrak{t}^*_+$ by Borel's theorem. Among all of these we single out the one corresponding to the intersection point of $\mathcal{O}_{\mu_0}$ with $\mathfrak{t}^*_+$. An explicit formula in the spirit of the above two expressions is not possible. Let $A(\mu) = \sqrt{-(ad\,\mu)^2}$ be the positive square root of $-(ad\,\mu)^2$. Then the Kähler metric is given by

$$\langle (\mu, \eta], [\mu, \zeta] \rangle_K = \langle A(\mu)[\mu, \zeta], [\mu, \zeta] \rangle_n$$

whereas the induced metric and the normal metric are related by

$$\langle[\mu,\eta],[\mu,\zeta]\rangle_i = \langle A(\mu)^2[\mu,\eta],[\mu,\zeta]\rangle_n$$

as can be easily checked. From the formula of the Kähler metric one sees that the root space decomposition of $\mathfrak{g}$ enters quite explicitly.

Generically, if $\mathcal{O}_{\mu_0}$ is diffeomorphic to G/T, the Kähler metric is determined by a T-invariant metric on $\mathfrak{g}/\mathfrak{t} = \bigoplus_\alpha \mathfrak{g}_\alpha$, where $\mathfrak{g}_\alpha$ is the root space determined by the root α. Let $\mu \in \mathcal{O}_{\mu_0}$ be in the interior of the positive Weyl chamber, i.e. $\alpha(\mu) > 0$ for any positive root α. On each two-dimensional space $\mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha}$, for α a positive root, the negative of the Killing form determines an inner product. Any other inner product is given by multiplying it by a scalar. Then these scalars are all equal to 1 for the normal metric, equal to $\alpha(\mu)$ for the Kähler metric, and equal to $\alpha(\mu)^2$ for the induced metric; the spaces $\mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha}$ and $\mathfrak{g}_\beta \oplus \mathfrak{g}_{-\beta}$ for different positive roots α, β are orthogonal.

§5. A special gradient system for a linear function on coadjoint orbits.

In this section we begin by discussing a gradient system on a compact Lie group and its induced system on a specific coadjoint orbit. We wil show that this new system is also gradient relative to the restriction of a linear function to the orbit. The dynamic properties of its flow turn out to be intimately connected with the convexity results of §3. The material in this section reviews the work in Bloch, Brockett, Ratiu [1990a], [1990b] some of which was in turn influenced by the results in Brockett [1988], [1989] dealing with the $SU(n)$-case.

Let K be a compact semisimple Lie group, $\mathfrak{k}$ its Lie algebra and κ is Killing form. Fix two elements $Q, N \in \mathfrak{k}$ and define the function $F : K \to \mathbf{R}$ by

$$F(\theta) = \kappa(Q, Ad_\theta N). \tag{5.1}$$

Endow K with the bi-invariant Riemannian metric $\langle\cdot,\cdot\rangle$ whose value at the identity is minus the Killing form. Denote by $\theta \cdot P$ the left translate of $P \in \mathfrak{k}$ by $\theta \in K$ to $T_\theta K$. If

$v_\theta = \theta \cdot R \in T_\theta K$, $R \in \mathfrak{k}$, then

$$\begin{aligned} dF(\theta) \cdot v_\theta &= \frac{d}{dt} \Big|_{t=0} F(\theta \exp tR) \\ &= \frac{d}{dt} \Big|_{t=0} \kappa(Ad_{\theta^{-1}} Q, Ad_{\exp tR} N) \\ &= \kappa(Ad_{\theta^{-1}} Q, [R, N]) \\ &= \kappa([N, Ad_{\theta^{-1}} Q], R) \\ &= -\langle \theta \cdot [N, Ad_{\theta^{-1}} Q], v_\theta \rangle \end{aligned}$$

whence

$$\nabla F(\theta) = \theta \cdot [Ad_{\theta^{-1}} Q, N] \tag{5.2}$$

and we get

Proposition 5.1. *The gradient flow on K relative to $\langle \cdot, \cdot \rangle$ and the function (5.1) is given by $\dot{\theta} = \theta \cdot [Ad_{\theta^{-1}} Q, N]$.*

Let $\mathcal{O}$ be the adjoint orbit containing Q. The projection of K to $\mathcal{O}$ is given by

$$\theta \mapsto Ad_{\theta^{-1}} Q. \tag{5.3}$$

Via this map, the gradient flow in proposition 5.1 transforms to

$$\dot{L} = [L, [L, N]] \tag{5.4}$$

Indeed, put $L(t) = Ad_{\theta(t)^{-1}} Q$, where $\theta(t)$ is an integral curve of $\dot{\theta} = \theta \cdot [Ad_{\theta^{-1}} Q, N]$; then

$$\begin{aligned} \dot{L}(t) &= -[\theta^{-1}(t)\dot{\theta}(t), Ad_{\theta(t)^{-1}} Q] \\ &= [L(t), \theta(t)^{-1}\dot{\theta}(t)] = [L(t), [L(t), N]]. \end{aligned}$$

Next, decompose orthogonally $\mathfrak{k} = \mathfrak{k}_L \oplus \mathfrak{k}^L$, $\mathfrak{k}_L = \{x \in \mathfrak{k} | [L, X] = 0\}$, $\mathfrak{k}^L = im(ad\, L)$, and denote by $X = X_L + X^L$ the decomposition of $X \in \mathfrak{k}$ into its $\mathfrak{k}_L$ and $\mathfrak{k}^L$-components (see the end of §4). Endow $\mathcal{O}$ with the normal metric

$$\langle [L, X], [L, Y] \rangle_n = \langle X^L, Y^L \rangle \tag{5.5}$$

and define . .

$$H : L \in \mathfrak{k} \mapsto \kappa(L, N) \in \mathbf{R} \tag{5.6}$$

By definition of the gradient relative to $\langle \cdot, \cdot \rangle_n$ on $\mathcal{O}$ we have for the restriction of H to $\mathcal{O}$

$$\begin{aligned}
\langle gradH(L), [L, \delta L] \rangle_n &= dH(L) \cdot [L, \delta L] \\
&= -dH(L) \cdot [\delta L, L] \\
&= -\frac{d}{dt} \Big|_{t=0} H(\exp t\delta L \cdot L) \\
&= \frac{d}{dt} \Big|_{t=0} \langle \exp t\delta L \cdot L, N \rangle \\
&= \langle [\delta L, L], N \rangle \\
&= \langle [L, N], \delta L \rangle \\
&= \langle [L, N], (\delta L)^L \rangle.
\end{aligned}$$

Therefore, if we set $grad\, H(L) = [L, X]$, the above equality and (5.5) say that

$$\langle X^L, (\delta L)^L \rangle = \langle [L, N], (\delta L)^L \rangle$$

whence $X^L = ([L, N])^L = [L, N]$. Thus

$$grad\, H(L) = [L, [L, N]]$$

which coincides with (5.4). We summarize these results in the following:

Theorem 5.1. *The projection (5.3) of the gradient flow $\dot{\theta} = \theta \cdot [Ad_{\theta^{-1}} Q, N]$ to the orbit $\mathcal{O}$ through Q is the gradient flow $\dot{L} = [L, [L, N]]$ on $\mathcal{O}$ relative to the normal metric and the function H given by (5.6).*

The connection of the dynamics defined by this gradient system and the convexity results of §3 is given by the following.

Theorem 5.2. *On the orbit $\mathcal{O}$ of K in $\mathfrak{k}$ consider the gradient flow $\dot{L} = [L, [L, N]]$ for N a fixed regular element. Let F_t be the flow of this vector field on $\mathcal{O}$. The set of equilibria equals $\mathcal{O} \cap \mathfrak{t}$ where $\mathfrak{t}$ is the Cartan subalgebra of $\mathfrak{k}$ containing N. This set $\mathcal{O} \cap \mathfrak{t}$ consists of a single Weyl group orbit. The convex hull of these equilibria is a compact-polytope $\mathcal{P}$*

which is the image of $\mathcal{O}$ under the momentum map $\pi : \mathcal{O} \to \mathfrak{t}$ (the orthogonal projection) defined by the adjoint T-action on $\mathcal{O}$, where T is the maximal torus in $\mathfrak{k}$ obtained by exponentiating $\mathfrak{t}$. $\pi(F_t(\mathcal{O}))$ lies entirely in $\mathcal{P}$.

Proof. Indeed, by Kostant's theorem, stated in Corollary 3.1, the theorem is proved provided we show that the equilibria of $\dot{L} = [L, [L, N]]$ necessarily lie in $\mathfrak{t}$. Since this gradient field is defined on the compact manifold $\mathcal{O}$, its flow is complete, so all that needs to be shown is that $\lim_{t \to \infty} L(t) \in \mathfrak{t}$ for any integral curve $L(t)$. But

$$\frac{d}{dt}(\kappa(L, N)) = \kappa(N, [L, [L, N]]) = -\kappa([L, N], [L, N]) \geq 0$$

so that $\kappa(L(t), N)$ is increasing as a function of time. It must also be bounded since $L(t) \in \mathcal{O}$ and $\mathcal{O}$ is compact. Thus $\kappa(L(t), N)$ has a limit and its derivative vanishes if and only if $L(\infty)$ and N commute, i.e. the equilibrium $L(\infty)$ must lie in $\mathfrak{t}$. □

By taking advantage of this theorem, Bloch Brockett, Ratiu [1990b] prove the following

Theorem 5.3. *In the hypotheses and notation of Theorem 5.2 we have:*

(i) The only stable equilibrium (a sink) of this gradient vector field is the one that lies in the Weyl chamber of $-N$.

(ii) The only source of this vector field is the unique equilibrium which lies in the same Weyl chamber as N.

(iii) The dimension of the stable manifold at any of the equilibria equals the length of the Weyl group element which maps the Weyl chamber of N to the Weyl chamber containing this equilibrium.

§6. The Toda equations as a gradient system.

In this section we show the relationship between the non-periodic Toda lattice equations and the gradient system in §5. It turns out that for very special choices of N, the two systems coincide on the isospectral set. Thus, the Toda flow naturally inherits all the scattering behavior of its "brother", the gradient system discussed in §5. This property of the Toda flow has found remarkable applications in numerical analysis, especially regarding the QR-algorithm; see Deift, Nanda, Tomei [1983], Lagarias [1988], Symes [1982]. The material in this section reviews the results in Bloch, Brockett, Ratiu [1990a], [1990b], Bloch, Flaschka, Ratiu [1990],

We begin with a slight generalization of the non-periodic Toda lattice associated to an arbitrary Dynkin diagram. These equations will lie on orbits of subalgebras in a *compact* Lie algebra which is the compact real form of a complex semisimple Lie algebra. For certain values of the parameters, the usual generalized non-periodic Toda lattice is recovered as we shall explicitly point out later on. This set-up is the natural one in which one links the Toda equations with the gradient system in §5 and also gives rise to several interesting questions, some of which will be addressed in §7.

We build on the Lie algebraic notations and conventions at the end of §4 which are in agreement with Humphreys [1972]. $\mathfrak{g}$ denotes a complex semi-simple Lie algebra of rank ℓ. Fix a Cartan subalgebra of $\mathfrak{g}$ and denote by $\Phi, \Phi^+, \Phi^-, \Delta$ the systems of roots, of positive roots, of negative roots, and of simple roots $\Delta = \{\alpha_1, \ldots, \alpha_\ell\}$ respectively. The Killing form on $\mathfrak{g}$ is denoted by κ, and the inner product on roots by $(\cdot,\cdot)$. The notation $\langle\alpha,\beta\rangle = (\alpha,\beta)/(\beta,\beta)$ for $\alpha,\beta \in \Phi$ will also be employed. $\mathfrak{g}_\alpha$ denotes the α-root space of $\mathfrak{g}$. Fix a Chevalley basis $\{h_i, e_\alpha | i = 1,\ldots,\ell, \alpha \in \Phi\}$ and recall that e_α is the basis vector of the one-dimensional complex vector space $\mathfrak{g}_\alpha$, $h_\alpha = [e_\alpha, e_{-\alpha}]$ for $\alpha \in \Phi^+$, $h_i = h_{\alpha_i}$, and $\kappa(h_\alpha, h) = 2\alpha(h)/(\alpha,\alpha)$ for all $h \in \mathfrak{g}$ and $\alpha \in \Phi^+$. Let $x_\alpha = e_\alpha - e_{-\alpha}$, $y_\alpha = i(e_\alpha + e_{-\alpha})$ for $\alpha \in \Phi^+$ and define:

(i) the *compact real form* of $\mathfrak{g}$

$$\mathfrak{k} = \{i\sum_{j=1}^{\ell} b_j h_j + \sum_{\alpha\in\Phi^+} i(c_\alpha x_\alpha + d_\alpha y_\alpha) | b_j, c_\alpha, d_\alpha \in \mathbf{R}\}$$

(ii) the *normal real form* of $\mathfrak{g}$

$$\mathfrak{g}_n = \{\sum_{j=1}^{\ell} i b_j h_j + \sum_{\alpha\in\Phi^+} c_\alpha e_\alpha | b_j, c_\alpha \in \mathbf{R}\}$$

(iii) the *compact toral subalgebra* of $\mathfrak{g}$

$$\mathfrak{t} = \{i\sum_{j=1}^{\ell} b_j h_j | b_j \in \mathbf{R}\}$$

(iv) the *non-compact toral subalgebra* of $\mathfrak{g}$

$$\mathfrak{a} = i\mathfrak{t} = \{\sum_{j=1}^{\ell} b_j h_j | b_j \in \mathbf{R}\}$$

(v) $\mathfrak{b} = \mathfrak{a} \oplus \bigsqcup_{\alpha \in \Phi^+} \mathfrak{g}_\alpha, \quad \mathfrak{n}^{\pm} = \bigsqcup_{\alpha \in \Phi^{\pm}} \mathfrak{g}_\alpha$

Think of $\mathfrak{g}$ as a *real* Lie algebra and thus $\dim_{\mathbf{R}} \mathfrak{g} = 2 \dim_{\mathbf{C}} \mathfrak{g}$. We shall always denote by $\mathfrak{g}$ this real Lie algebra and will write $\mathfrak{g}_{\mathbf{C}}$ whenever $\mathfrak{g}$ is thought of as a complex Lie algebra. We have $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{b}$. The adjoint groups corresponding to these Lie algebras are denoted by $G, K, G_n, T, A, B, N^{\pm}$ repsectively. For the Iwasawa decomposition $G = KAN = KB$ we will write the factorization of $g \in G$ as $g = \mathbf{k}(g)\mathbf{b}(g)$ for $\mathbf{k}(g) \in K$, $\mathbf{b}(g) \in B$.

There are two Poisson structures on $\mathfrak{k}$. The first one is obtained by pulling back the Lie-Poisson structure of $\mathfrak{k}^*$ to $\mathfrak{k}$ via the restriction of the Killing form κ to $\mathfrak{k}$ which is negative definite. The other Poisson structure comes from a construction used by Lu and Weinstein [1990]. The non-degenerate bilinear form $Im\,\kappa$ on $\mathfrak{g}$ vanishes on $\mathfrak{k}$ since $\mathfrak{k}$ is a real form of $\mathfrak{g}_{\mathbf{C}}$. Also $\kappa(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$ unless $\alpha + \beta = 0$ and thus $\kappa(\mathfrak{b}, \mathfrak{b}) = \kappa(\mathfrak{a}, \mathfrak{a}) \in \mathbf{R}$; thus $Im\,\kappa$ vanishes on $\mathfrak{b}$. The annihilator of $\mathfrak{k}$ relative to $Im\,\kappa$ is $\mathfrak{k}$ itself, and is isomorphic to $\mathfrak{b}^*$ as a real vector space. We summarize these observations together with the Adler-Kostant-Symes theorem in the following:

Proposition 6.1. *(i) $\mathfrak{k}$ is equipped with the usual Lie-Poisson bracket. The coadjoint orbits $\mathcal{O}$ are the symplectic leaves of this Poisson manifold. The symplectic form is given by the Kirillov-Kostant-Souriau formula (see §2). The coadjoint orbits $\mathcal{O}$ are Kähler relative to the canonical Kähler structure (see §4).*

(ii) $\mathfrak{k} \cong \mathfrak{b}^*$ has a second Lie-Poisson bracket given by

$$\{f, h\}(\xi) = -Im\,\kappa(\xi, [\pi_{\mathfrak{b}} \nabla \tilde{f}(\xi),\ \pi_{\mathfrak{b}} \nabla \tilde{h}(\xi)])$$

where $\tilde{f}, \tilde{h}$ are arbitrary extensions to $\mathfrak{g}$ of $f, h : \mathfrak{k} \to \mathbf{R}$, ∇ is the gradient relative to $Im\,\kappa$, i.e.

$$d\tilde{f}(\xi) \cdot \delta\xi = Im\,\kappa(\nabla \tilde{f}(\xi), \delta\xi),$$

$\pi_{\mathfrak{b}}$ is the projection onto $\mathfrak{b}$ corresponding to the decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{b}$, and $\xi, \delta\xi \in \mathfrak{k}$.

(iii) If f is an invariant function on $\mathfrak{g}$, i.e. $[\nabla f(\zeta), \zeta] = 0$ for all $\zeta \in \mathfrak{g}$, then the Hamiltonian vector field relative to the Poisson structure in (ii) defined by $f|\mathfrak{k}$ is given by the Lax equation

$$X_{f|\mathfrak{k}}(\xi) = [\pi_{\mathfrak{k}} \nabla f(\xi), \xi], \quad \xi \in \mathfrak{k}.$$

Hamiltonian vector fields defined by invariant functions commute.

The Toda equations are a Hamiltonian system on a coadjoint orbit in $\mathfrak{b}^* = \mathfrak{k}$ of dimension 2ℓ. The following is a straightforward verification following the pattern in Kostant [1979]:

Proposition 6.2. *Let $\delta = (\delta_1, \ldots, \delta_\ell)$ be a vector of 0's and 1's and fix $(\theta_1, \ldots, \theta_\ell)$, $\theta_j \in [0, 2\pi]$. Let $L(\delta, \theta) = \sum_{j=1}^{\ell} \delta_j e^{i\theta_j} x_{\alpha_j}$. Then the coadjoint B-orbit through $L(\delta, \theta)$ is given by*

$$\begin{aligned} Jac := &\{\pi_{\mathfrak{k}} Ad_b^G L(\delta, \theta) | b \in B\} \\ = &\{L = \sum_{j=1}^{\ell} [b_j h_j + \delta_j (a_j e_{\alpha_j} - \overline{a}_j e_{-\alpha_j})] | ib_j \in \mathbf{R}, \\ &a_j \in \mathbf{C} \backslash \{0\}, \quad arg\, a_j = \theta_j \text{ if } \delta_j \neq 0\}, \end{aligned} \tag{6.1}$$

where Ad^G denotes the usual adjoint action of G on $\mathfrak{g}$.

If one takes $\theta_j = \pi/2$ and $\delta_j = 1$, for all $j = 1, \ldots, \ell$, then L is, up to a factor of i, a typical element of a so-called "Toda orbit" in $\mathfrak{g}_n$ of Symes [1982]. We will call all elements L appearing in (6.1) *Jacobi elements*.

Let $I_1, \ldots, I_\ell$ be a set of homogeneous generators of the ring of invariant polynomials on $\mathfrak{g}$ chosen such that their restrictions to $\mathfrak{k}$ are real and generate the invariant polynomials on $\mathfrak{k}$. The Hamiltonian equations

$$\dot{L} = [\pi_{\mathfrak{k}} \nabla I_j(L),\ L] \tag{6.2}$$

on the B-orbit Jac of Jacobi elements (5.1) will be called the *Toda hierarchy*. Fix an element Λ in the interior of the positive Weyl chamber of $\mathfrak{t}$ and let $\mathcal{O}_\Lambda$ be the K-orbit through Λ. The the Toda hierarchy leaves $\mathcal{O}_\Lambda \cap Jac$ invariant. For $A_\ell, \mathcal{O}_\Lambda$ consists of matrices with fixed spectrum Λ and $\mathcal{O}_\Lambda \cap Jac$ is the isospectral manifold. After multiplying by i (and setting $\theta_j = \pi/2$, $\delta_j = 1$ for $j = 1, \ldots, \ell$) the first vector field in the Toda hierarchy for A_ℓ takes the familiar form $\dot{L} = [B, L]$, where

$$L = \begin{bmatrix} b_1 & a_1 & & & \\ a_1 & b_2 - b_1 & a_2 & & \\ & \ddots & \ddots & \ddots & \\ & & & b_\ell - b_{\ell-1} & a_\ell \\ & & & a_\ell & -b_\ell \end{bmatrix}, B \begin{bmatrix} 0 & a_1 & & & & \\ -a_1 & 0 & a_2 & & & \\ & \ddots & \ddots & \ddots & & \\ & & & & 0 & a_\ell \\ & & & & -a_\ell & 0 \end{bmatrix} \tag{6.3}$$

For an arbitrary Lie algebra

$$B = \sum_{j=1}^{\ell} a_j (e_{\alpha_j} - e_{-\alpha_j}) \tag{6.4}$$

Theorem 6.1. *The Toda hierarchy is a completely integrable system and the 2ℓ integrals in involution are $I_1, \ldots, I_\ell$.*

For a proof of this and a wealth of additional material on the Toda lattice see Kostant [1979] and Symes [1980], [1982a]. We will return and sketch how a solution of the Toda lattice is found in terms of the Iwasawa decomposition in the next section. In the sequel we shall fix all $\theta_j = \pi/2$ and all $\delta_j = 1$ and set $arg\, a_j = \pi/2$.

Theorem 6.2. *If H is minus i times the sum of the simple coweights of $\mathfrak{g}$, then for*

$$L = \sum_{j=1}^{\ell} i b_j h_j + \sum_{j=1}^{\ell} i a_j (e_{\alpha_j} + e_{\alpha_j}),$$

b_j, $a_j \in \mathbf{R}$, equation $\dot{L} = [L, [L, N]]$ gives the generalized Toda flow on the isospectral set $Jac \cap O_\Lambda$. Explicitly, $N = \sum_{j=1}^{\ell} i x_j h_j$ where $(x_1, \ldots, x_\ell)$ is the unique solution of the system $\sum_{j=1}^{\ell} x_j \alpha_k(h_j) = -1$, $k = 1, \ldots, \ell$.

For the proof, it is enough to require that $[L, N] = -B$ with B given by (5.4) and to observe that the solution of the system in $(x_1, \ldots, x_\ell)$ gives $N = -ih_\delta$, where $h_\delta = \sum_{r=1}^{\ell} \check{\lambda}_i$ since $(h_\delta, e_{\alpha_i}) = 1$, (,) being the inner product. Therefore we can conclude

Corollary 6.1. *The generalized Toda lattice equations are a gradient flow on $Jac \cap \mathcal{O}_\Lambda$ when $\mathcal{O}_\Lambda$ is equipped with the normal metric.*

The coefficients x_i in N are listed for all simple Lie algebras below

$\underline{\mathcal{G}}$	$\underline{x}$
A_ℓ	$x_i = -\frac{1}{2}i(\ell - i + 1),\ i = 1, 2, \ldots, \ell$
B_ℓ	$x_i = -\frac{1}{2}i(2\ell - i + 1),\ i = 1, 2, \ldots, \ell - 1$
	$x_\ell = -\frac{1}{4}\ell(\ell + 1)$
C_ℓ	$x_i = -\frac{1}{2}i(2\ell - i)\ i = 1, 2, \ldots, \ell$
D_ℓ	$x_i = -\frac{1}{2}i(2\ell - 1 - i),\ i = 1, 2, \ldots, \ell - 2$
	$x_{\ell-1} = x_\ell = -\frac{1}{4}\ell(\ell - 1)$
G_2	$x_1 = -3,\ x_2 = -5$
F_4	$x_1 = -11,\ x_2 = -21,\ x_3 = -15,\ x_4 = -8$
E_6	$x_1 = -8,\ x_2 = -11,\ x_3 = -15,$
	$x_4 = -21,\ x_5 = -15,\ x_6 = -8$
E_7	$x_1 = -17,\ x_2 = -\frac{49}{2},\ x_3 = -33,$
	$x_4 = -48,\ x_5 = -\frac{75}{2},\ x_6 = -26,$
	$x_7 = -\frac{27}{2}$
E_8	$x_1 = -46,\ x_2 = -68,\ x_3 = -91,$
	$x_4 = -135,\ x_5 = -110,\ x_6 = -84,$
	$x_7 = -57,\ x_8 = -29$

For the classical simple Lie algebras represented as in Sattinger and Weaver [1986], for example, we have

$A_\ell : h_1 = \text{diag}(1, -1, 0, 0, \ldots),\ h_2 = \text{diag}(0, 1, -1, 0, \ldots)\ldots,$

$h_\ell = \text{diag}(0, \ldots, 0, 1 - 1)$ and hence

$$N = i \begin{bmatrix} -\frac{\ell}{2} & & & & \\ & \frac{-\ell+2}{2} & & & \\ & & \ddots & & \\ & & & \frac{\ell-2}{2} & \\ & & & & \frac{\ell}{2} \end{bmatrix},$$

$B_\ell : h_1 = \mathrm{diag}(0,1,-1,-1,1,0,\dots),\ h_2 = \mathrm{diag}(0,0,0,1,-1,-1,1,\dots)\dots,$

$h_\ell = \mathrm{diag}(0,\dots,0,2,-2)$ and

$$N = \begin{bmatrix} 0 &&&&&&& \\ & -\ell &&&&&& \\ && \ell &&&&& \\ &&& -\ell+1 &&&& \\ &&&& 1-\ell &&& \\ &&&&& \ddots && \\ &&&&&& -1 & \\ &&&&&&& 1 \end{bmatrix}.$$

$C_\ell : h_1 = \mathrm{diag}(1,-1,-1,1,0,\dots),\ h_2 = \mathrm{diag}(0,0,1,-1,-1,1,\dots)\dots,$

$h_\ell = (0,0,\dots,0,1,-1)$ and

$$N = i\begin{bmatrix} \frac{1}{2}(1-2\ell) &&&&&&&& \\ & -\frac{1}{2}(1-2\ell) &&&&&&& \\ && \frac{1}{2}(3-2\ell) &&&&&& \\ &&& -\frac{1}{2}(3-2\ell) &&&&& \\ &&&& \ddots &&&& \\ &&&&& -\frac{3}{2} &&& \\ &&&&&& \frac{3}{2} && \\ &&&&&&& -\frac{1}{2} & \\ &&&&&&&& \frac{1}{2} \end{bmatrix}.$$

D_ℓ: h_i as for C_ℓ and

$$N = \begin{bmatrix} -\ell+1 &&&&&&&& \\ & \ell-1 &&&&&&& \\ && -\ell+2 &&&&&& \\ &&& 2-\ell &&&&& \\ &&&& \ddots &&&& \\ &&&&& -1 &&& \\ &&&&&& 1 && \\ &&&&&&& 0 & \\ &&&&&&&& 0 \end{bmatrix}.$$

Corollary 6.1 and Theorem 5.2 imply

Corollary 6.2. *The projection of the flow of the Toda lattice to* $\mathfrak{t}$ *lies in the interior of the polytope which is the convex hull of* $O_\Lambda \cap \mathfrak{t}$.

This polytope is, however, not filled by all flows of the Toda hierarchy. For A_3 the image is polygonal but lies inside the polytope. However, in general the image is not even polygonal as shown by some explict cases studied by M. Zou at the University of Arizona.

§7 A convexity theorem for Jacobi elements

This section surveys the main results of Bloch, Flaschka, Ratiu [1990]. The starting point is the observation at the end of §6 that the Toda hierarchy flows do *not* fill out a convex polytope by orthogonal projection. Return to the notations and conventions of §6; set $\delta_j = 1$, $\theta_j = \pi/2$ and so $\arg a_j = \frac{\pi}{2}$ for all $j = 1, \ldots, \ell$. The latter condition is inessential since the arguments of the a_j are always constant on Jac. Thus in this case

$$Jac = \{L = i \sum_{j=1}^{\ell} [b_j h_j + a_j(e_{\alpha_j} + e_{j-\alpha_j})] \mid a_j, b_j \in \mathbf{R}\}$$

and we shall fix once and for all the B-orbit in $\mathfrak{b}^* \cong \mathfrak{k}$. Define

$$\mathcal{J}_\Lambda^0 = \{L \in \mathcal{O}_\Lambda \cap Jac \mid a_j > 0 \text{ for all } j = 1, \ldots, \ell\}$$

$$\mathcal{J}_\Lambda = \{L \in \mathcal{O}_\Lambda \cap Jac \mid a_j > 0 \text{ for all } j = 1, \ldots, \ell\}.$$

The closure of $\mathcal{J}_\Lambda^0$ is $\mathcal{J}_\Lambda$ and the boundary of $\mathcal{J}_\Lambda$ is a disjoint union of $2^\ell - 1$ strata each stratum being given by the vanishing of some of the a_j's. It is straightforward to see (as in Kostant [1979] for example) that $\mathcal{J}_\Lambda^0$ is preserved by the flows of the Toda hierarchy. For the case of $K = SU(\ell+1)$ the Jacobi matrix L is

$$L = i \begin{bmatrix} b_1 & a_1 & & & \\ a_1 & b_2 - b_1 & a_2 & & \\ & \ddots & \ddots & \ddots & \\ & & & b_\ell - b_{\ell-1} & a_\ell \\ & & & a_\ell & -b_\ell \end{bmatrix},$$

and

$$\mathcal{J}_\Lambda^0 = \{L \mid L \text{ is conjugate to } \Lambda \text{ and } a_j > 0 \text{ for all } j\},$$

$$\mathcal{J}_\Lambda = \{L \mid L \text{ is conjugate to } \Lambda \text{ and } a_j \geq 0 \text{ for all } j\},$$

$\mathcal{J}_\Lambda$ is the closure of $\mathcal{J}_\Lambda^0$ which in this case is the *isospectral set* of the Jacobi matrices L. We shall refer to $\mathcal{J}_\Lambda$ as the *closed isospectral set* and to $Jac \cap \mathcal{O}_\Lambda$ as the *isospectral manifold.* We shall use the same names when dealing with a general Lie algebra. Since the Toda lattice is integrable and the integrals are $I_1, \ldots, I_\ell$ (see §6), $\mathcal{J}_\Lambda^0$ is an orbit of the non-compact torus given by the flows of the Hamiltonian vector fields corresponding to these functions. Thus $\mathcal{J}_\Lambda$ is the closure of a non-compact torus orbit in the Kähler manifold $\mathcal{O}_\Lambda$.

In view of Atiyah's Theorem 4.1, the map $(I_1, \ldots, I_\ell)$ *to* $\mathbf{R}^\ell$ *must have convex image. But we just saw at the end of §6 that the orthogonal projection does not have convex image. Therefore, the question naturally arises: can one embed* $\mathcal{J}_\Lambda$ *in a different way in* $\mathcal{O}_\Lambda$ *such that its projection is convex?* The answer to this question is positive and solved in Bloch, Flaschka, Ràtiu [1990]; we will describe below the main ideas.

There has been prior work in which $\mathcal{J}_\Lambda$ and $\mathcal{J}_\Lambda^0$ appear in the context of convexity properties. Moser [1975] proved that $\mathcal{J}_\Lambda^0$ is diffeomorphic to $\mathbf{R}^\ell$ and Tomei [1984] proved that $\mathcal{J}_\Lambda$ is homeomorphic to a convex closed polyhedron. The first pictures of convex polyhedra appear in van Moerbeke [1976] and then again in Deift, Nanda, Tomei [1983]. These polyhedra are used in these works as a useful tool to describe certain topological relationships. In view of the convexity results of §4, it should be expected that these polyhedra have a symplectic significance as we shall outline below. The construction employed to achieve this goal is relatively involved and we shall omit it in this review. But it is based on a mysterious relationship between diagonal non-compact toral actions and left dressing transformations. Together with the ideas of §6, this work has relationships to Fried's [1986] cohomology computations and to the work of Davis [1987] and Davis and Janusiewicz [1990] on aspherical manifolds. The work of Lu and Weinstein [1990] and Lu [1990] on dressing transformations seems particularly relevant to this circle of ideas as will also be seen in §8.

Before outlining the results, a very simple example is in order. Let $\mathfrak{k} = su(2)$ and identify $su(2)^*$ with $\mathbf{R}^3$. The Lie-Poisson structure of $su(2)^*$ has the concentric spheres and the origin as symplectic leaves. Represent $\Lambda = diag(i\lambda, -i\lambda)$, $\lambda > 0$ by the point $(0, 0, \lambda)$ so that $\mathcal{O}_\Lambda$ is a sphere of radius λ centered at the origin. The symplectic form is, up to a factor of $-1/\lambda$, the area form, and the complex structure, making $\mathcal{O}_\Lambda$ into a Kähler manifold, is essentially that of the Riemann sphere. The orbit $\mathcal{O}_\Lambda$ consists in this simple example of Jacobi elements. The other Poisson structure of $\mathbf{R}^3 \cong \mathfrak{b}^*$ has leaves which are open half-planes containing the vertical axis and the points of the vertical axis, if one lets B be the set of upper triangular matrices in $SU(2)$. The compact torus $T = \{diag(e^{i\theta}, e^{-i\theta}) \mid \theta \in \mathbf{R}\}$

acts by rotating the sphere $\mathcal{O}_\Lambda$ about the vertical axis; namely

$$L \in \begin{pmatrix} ib & a \\ -\overline{a} & -ib \end{pmatrix} \in \mathcal{O}_\Lambda \subset su(2),$$

$b \in \mathbf{R}$, $a \in \mathbf{C}$, is sent to

$$\begin{pmatrix} ib & ae^{2\theta} \\ -\overline{a}e^{-2\theta} & -ib \end{pmatrix}.$$

The momentum map J of this action maps $L \in \mathcal{O}_\Lambda$ to b, i.e. it projects the point on the sphere $\mathcal{O}_\Lambda$ parallel to the horizontal plane onto the vertical axis. The image of the sphere $\mathcal{O}_\Lambda$ and of the "meridian", which is the intersection of $\mathcal{O}_\Lambda$ with a B-orbit (a "page" of the "book" $\mathfrak{b}^*$), coincide and equal the convex interval $[-\lambda, \lambda]$.

This example generalizes completely, except for the very difficult projection part, which does not give any problems here due to the low dimensionality of the example. What is true is that $\mathcal{J}_\Lambda$ (here the "meridian") can be embedded in a special way in $\mathcal{O}_\Lambda$ such that its projection is a convex polytope.

The convexity result alluded to above is built upon two main ingredients. The first one is the explicit solution of the Toda lattice equation in terms of the Iwasawa decomposition as it can be found in Kostant [1979], Reyman, Semenov-Tijan-Shanskii [1979], Symes [1980], or Goodman, Wallach [1984]. The goal is to solve

$$\dot{L}(t) = [\pi_{\mathfrak{k}} \bigtriangledown I_j(L(t)), L(t)], \quad L(0) = L, \tag{7.1}$$

where $I_1, \ldots, I_\ell$ are the homogeneous generators of the ring of real K-invariant polynomials on $\mathfrak{k}$. First one shows that if $\varphi : \mathfrak{k} \to \mathbf{R}$ is a homogeneous polynomial and one extends φ to $\mathfrak{g}_{\mathbf{C}}$ in the obvious way using homogeneity, then if $\chi = Re\,\varphi$ and the gradient is taken relative to $Im\,\kappa$, $\bigtriangledown\chi(\xi) \in i\mathfrak{k}$ for all $\xi \in \mathfrak{k}$. From here it follows that $\bigtriangledown I_j(L)$, $j = 1, \ldots, \ell$, generate $iC(L)$, where $C(L)$ is the centralizer of L. Secondly, carry out the factorization

$$\exp(t \bigtriangledown I_j(L)) = \mathbf{k}(\exp(t \bigtriangledown I_j(L)))\mathbf{b}(\exp(t \bigtriangledown I_j(L))); \tag{7.2}$$

then the solution of (7.1) is

$$L(t) = Ad_{\mathbf{k}(\exp(t\bigtriangledown I_j(L)))^{-1}} L = Ad_{\mathbf{b}(\exp(t\bigtriangledown I_j(L)))} L. \tag{7.3}$$

where $g = \mathrm{k}(g)\mathrm{b}(g)$ is the factorization of $g \in G$ defined by $G = KB$. Since $L \in \mathcal{O}_\Lambda$, we can write $L = Ad_k\Lambda$ for some $k \in K$. By our first comments, $Ad_{k^{-1}} i \nabla I_j(L) \in \mathfrak{t}$, or, equivalently, $Ad_{k^{-1}} \nabla I_j(L) \in \mathfrak{a} = i\mathfrak{t}$. Consequently, if $\mu \in \mathfrak{a}$, $X := Ad_k\mu$, $L := Ad_k\Lambda$, the curve $L(t) = Ad_{\mathrm{k}(\exp tX)^{-1}}L$ is a solution of a linear combination of the equations (7.3). But $L(t)$ can clearly be rewritten as

$$L = Ad_k\Lambda \mapsto L(t) = Ad_{\mathrm{k}(\exp(t\mu)k^{-1})^{-1}}\Lambda, \quad \exp t\mu \in A, \tag{7.4}$$

which defines the *left dressing action* of the non-compact real torus A on $\mathcal{O}_\Lambda$ (see Lu, Weinstein [1990], Lu [1990]). Unfortunately, the action of A defined by (4) is *not* the diagonal action

$$L = Ad_k\Lambda \mapsto Ad_{\mathrm{k}(hk)}\Lambda, \quad h \in T, \tag{7.5}$$

and Atiyah's theorem stated in Corollary 4.1 does not apply. Let J denote the momentum map of the diagonal T-action, i.e. the orthogonal projection onto $\mathfrak{t}$ relative to the invariant metric coming from the Killing form on $\mathfrak{k}$. The image $J(\mathcal{J}_\Lambda)$ is neither convex, nor a polytope, as shown by examples by M. Zou. One can turn, however, (7.4) into (7.5) by inversion. Namely, define $\iota : \mathcal{T}_\Lambda \to \mathcal{O}_\Lambda$ by $\iota(k\Lambda k^{-1}) = k^{-1}\Lambda k$ and the Toda flow (4) becomes

$$\iota(L) = \iota(Ad_k\Lambda) \mapsto Ad_{\mathrm{k}(\exp(t\mu)k^{-1})}\Lambda$$

which is exactly the diagonal action. Therefore $J(\iota(\mathcal{J}_\Lambda))$ will be a convex polytope by Atiyah's theorem. The trouble is that the "inversion" $Ad_k\Lambda \mapsto Ad_{k^{-1}}\Lambda$ makes no sense on $\mathcal{O}_\Lambda$: when one replaces k by kh, $h \in T$, the image under ι *will* depend on h.

More precisely , given $L \in \mathcal{O}_\Lambda \cap Jac$, $L = Ad_k\Lambda$, determines k only up to the right multiplication of an element of T; remember, Λ is chosen once and for all in the interior of the positive Weyl chamber of $\mathfrak{t}$. And this leads to the second main ingredient, namely, k must be chosen in a smooth way and uniquely for each $L \in \mathcal{J}_\Lambda$. That this is possible is ultimately based on the Bruhat decomposition of K into cells and constitutes the main technical part of the proof. It was inspired by ideas that show up also in Flaschka, Haine [1990] and Ercolani, Flaschka, Haine [1990]. The upshot of this work is that *the "inversion* *map" ι can be extended continuously to the boundary $\mathcal{J}_\Lambda \backslash \mathcal{J}_\Lambda^0$ and that it is a diffeomorphism on $\mathcal{J}_\Lambda^0$ and a homeomorphism on $\mathcal{J}_\Lambda$.* Here is its very simple and concrete description in the case of $SU(\ell+1)$. Let $L = k\Lambda k^{-1}$. The columns of k are orthonormalized eigenvectors of L. Let now $L \in \mathcal{J}_\Lambda^0$; then the first row entries r_j of k do not vanish. Fix k by the requirement that $r_j/r_1 \in \mathbf{R}$. This is the smooth unique choice of k for a given L. The "inversion" ι is now easy to describe. The Toda hierarchy acts on $\iota(\mathcal{J}_\Lambda^0)$ as the noncompact torus A as follows: Consider the element

$$a = \exp(p_1 c_1 \Lambda t_1 + \dots p_{\ell+1} c_{\ell+1} \Lambda^{\ell+1} t_{\ell+1}) \in A$$

where $c_k = i$ for k odd, $c_k = 1$ for k even, and $p_1, \dots, p_{\ell+1}$ are arbitrary real constants. We have the mapping

$$a : \iota(L) \mapsto \mathbf{k}(ak^{-1})\Lambda \mathbf{k}(ak^{-1})^{-1}.$$

It is a straightforward check to see that this action preserves the normalization of the r_j's described above.

Now, several conclusions are possible.

Theorem 7.1. *The image of $\mathcal{J}_\Lambda$ under the map $J \circ \iota : \mathcal{J}_\Lambda \to \mathfrak{t}$ is the convex hull of the Weyl group orbit through Λ.*

Combined with the results of §6 and those in Duistermaat, Kolk, Varadarajan [1983] who show that the action of the one-parameter group $\exp t\mu$ for $\mu \in \mathfrak{a}$, on $\mathcal{O}_\Lambda$, is a gradient flow generated by $-\kappa(\mu, \cdot)$ relative to the Kähler metric, Theorem 7.1 now implies

Theorem 7.2. *The Toda flows in $\mathcal{J}_\Lambda$ are gradient flows in the metric induced by pulling back the Kähler metric of $\iota(\mathcal{J}_\Lambda)$ to $\mathcal{J}_\Lambda$.*

It should be emphasized that the functions which generate the gradient flows are *not* the Toda Hamiltonians. Theorem 7.2 is complementary to Theorem 6.2, where the Toda flow is seen to be gradient relative to another metric which, geometrically, is more natural. Shades of the "inversion map" ι will appear also in §8 and there seems to be a connection between these results and the ones of next section.

We close with some comments on the relationship between the results in §6, §7 and Moser's [1975] gradient flows associated with the Toda lattice.

We recall from Sections 5 and 6 that the Toda lattice equations are gradient with respect to the normal metric on an orbit $\mathcal{O}_\Lambda$ and hence may be written in the form $\dot{L} = [L, [L, N]]$.

Now one can project the flow on $\mathcal{O}_\Lambda$ (which is diffeomorphic to K/T) to K/P, P a parabolic subgroup of K.

Consider specifically the case $K = SU(\ell+1)$ and let $P_0 = \text{diag}(1, 0, \dots 0) - \frac{I}{\ell+1}$. Now take the K-orbit $\mathcal{O}_{iP_0}$ through iP_0 with points $k(iP_0)k^{-1}$, $k \in K$. This orbit consists of points of the form: rank 1 projection matrix minus a multiple of the identity, and may thus be identified with $\mathbf{CP}^\ell$. Now we can check that for $iP \in \mathcal{O}_{iP_0}$, $-(ad\, iP)^2([iP, \eta]) = [iP, \eta]$ and hence (see section 4) the normal and Kähler metric coincide. (This also applies to a general Grassmannian; see Bloch, Flaschka and Ratiu [1990].)

Now consider the gradient flow with respect to the Kähler (or normal !) metric of $\phi_k(iP) = -cTr(iP\Lambda^k)$ on $\mathbf{CP}^\ell$. Here $\Lambda = \text{diag}(\lambda_1, \dots \lambda_\ell)$ and c is 1 or i as needed to make ϕ_k real. The gradient flow is

$$i\dot{P} = [iP, [iP, c\Lambda^k]].$$

Taking $P = r \otimes \bar{r}$, $r = (r_1, \dots r_{\ell+1})$ with $\sum |r_j|^2 = 1$ yields

$$|\dot{r}_j| = -|r_j|(\lambda_j^k - \sum_{i=1}^{\ell+1} \lambda_i^k |r_i|^2), \quad j = 1 \dots \ell+1$$

which is Moser's equation for the Toda lattice.

Note that this flow is just conjugate to the Hamiltonian flow $i\dot{P} = [iP, c\Lambda^k]$ – just apply the complex structure $[iP, \cdot]$!

Note also that Atiyah's theorem also applies to K/P. Ignoring the constant shift in P_0, the momentum mapping $J_P : K/P \to \mathfrak{t}$ is given by k^{-1} mod $P \to i\, diag(r \otimes \bar{r})$. The Jacobi matrix L is thus sent to $(i|r_1|^2, \dots, i|r_{\ell+1}|^2)$ and the image in $\mathfrak{t} = \mathbf{R}^\ell$ is the standard simplex $\sum_{i=1}^{\ell+1} \xi_i = 1$.

§8 The non-linear convexity theorem of Kostant

This section reviews the results in Lu, Ratiu [1990]. It is not directly related to the previous ideas on the relationship between integrable systems and convexity. However, the dressing transformation that appeared so prominently in §7 and certain aspects of the "inversion map" on $\mathcal{O}_\Lambda$ can be found here too. There seems to be a much deeper connection between the ideas in this section and the Toda lattice. On the symplectic geometric side, we think, the result discussed below raises a host of other questions. For example, are there other convexity theorems that have ultimately a symplectic origin? There are several one can think of, among which are other convexity results proved in the same paper of Kostant [1973], as Kostant himself has pointed out to us. On the other hand, are there meaningful generalizations of the result below to a symplectic group setting? Finally, do the constructions below shed some light on the second Hamiltonian structure for the A_ℓ-Toda lattice (as opposed to the well-known results on the $g\ell(\ell+1)$-Toda lattice)?

We return to §3 on the symplectic convexity results and set the stage for the statement of one of Kostant's non-linear convexity results at Lie group level.

G is a real connected semisimple Lie group with Lie algebra $\mathfrak{g}$ and $\mathfrak{k}$ is the Lie subalgebra corresponding to a maximal compact subgroup of the adjoint group of G. $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ is the *Cartan decomposition*, i.e. $\mathfrak{p}$ is the orthogonal complement of $\mathfrak{k}$ in $\mathfrak{g}$ relative to the Killing form. The elements of $\mathfrak{p}$ are all semisimple and $P = \exp \mathfrak{p}$ is a closed submanifold of G; *exp* is a diffeomorphism between $\mathfrak{p}$ and P. At the group level, the map $(k,p) \in K \times P \mapsto kp$ is a diffeomorphism, where K is a connected subgroup of G with Lie algebra $\mathfrak{k}$; one writes $G = KP$ and calls this the *Cartan decomposition of* G.

By the invariance of the Killing form, $[\mathfrak{k}, \mathfrak{p}] \subset \mathfrak{p}$ and therefore K acts on $\mathfrak{p}$ by the adjoint action and also on P by conjugation; the exponential map $\exp : \mathfrak{p} \to P$ is K-equivariant. Let $\mathfrak{a}$ denote a maximal abelian subalgebra of $\mathfrak{p}$. The elements of $\mathfrak{a}$ are all semisimple and thus the adjoint representation of $\mathfrak{a}$ on $\mathfrak{g}$ decomposes as $\mathfrak{g} = \mathfrak{l} \oplus \bigsqcup_{\alpha\in\Phi} \mathfrak{g}_\alpha$ where $\mathfrak{l}$ is the centralizer of $\mathfrak{a}$ in $\mathfrak{g}$, and for a linear functional α on $\mathfrak{a}$, $\mathfrak{g}_\alpha = \{\xi \in \mathfrak{g} \mid [\eta, \xi] = \alpha(\eta)\xi$ for all $\eta \in \mathfrak{a}\}$; Φ is the set of all such non-zero α's. Fix a basis in $\mathfrak{a}$ and introduce the lexicographic ordering on functionals on $\mathfrak{a}$. Put $\mathfrak{n} = \bigsqcup_{\alpha>0} \mathfrak{g}_\alpha$ and then $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}$ is the *Iwasawa* *decomposition of* $\mathfrak{g}$. Let A, N be the connected subgroups of G with Lie algebras $\mathfrak{a}$ and $\mathfrak{n}$ respectively. The *Iwasawa decomposition of* G states that $(k, a, n) \in K \times A \times N \mapsto kan \in G$ is a diffeomorphism; one writes $G = KAN$. By W we denote the relative Weyl group of $(K, \mathfrak{a})$, i.e. W is the quotient of the normalizer of $\mathfrak{a}$ in K by the centralizer of $\mathfrak{a}$ in K.

If G is a complex semisimple Lie group, but thought of as a real Lie group, the above choices are easier. $\mathfrak{k}$ is a compact real form of $\mathfrak{g}$, the Killing form on $\mathfrak{g}$ is the complex linear extension of that on $\mathfrak{k}$, $\mathfrak{p} = i\mathfrak{k}$, $\mathfrak{a} = i\mathfrak{t}$, where $\mathfrak{t}$ is a Cartan subalgebra of $\mathfrak{k}$. The Weyl group of $(K, \mathfrak{t})$ is that of $(K, \mathfrak{a})$. Fixing a basis of $\mathfrak{t}$ will ultimately define an Iwasawa decomposition of $\mathfrak{g}$ and G.

Now let G be a semisimple Lie group, real or complex. The Cartan and Iwasawa decompositions define a diffeomorphism between P and AN, namely, if $p = kan$, then associate to p the element an. Let ρ_A be the projection from G, P, or AN to A according to the Iwasawa decomposition.

Theorem 8.1. *(Kostant [1973]) For $a \in A$ denote by $\mathcal{O}_a$ the K-orbit of a in P. Then $\rho_A(\mathcal{O}_a)$ is the convex hull of the Weyl group orbit $W \cdot a$ in A. (The Lie group A is identified with its Lie alegbra $\mathfrak{a}$ via the exponential map, so convexity makes sense.)*

This statement is one of Kostant's nonlinear convexity theorems. The term is justified by the following remarks and the example to be discussed below. The differential (tangent map) of $\rho_A : P \to A$ gives the orthogonal projection $\rho_{\mathfrak{a}} : \mathfrak{p} \to \mathfrak{a}$ relative to the Killing form. Let $log : A \to \mathfrak{a}$ be the inverse of exp and define $J_1 := log \circ \rho_A \circ exp : \mathfrak{p} \to \mathfrak{a}$. We compare J_1 to $\rho_{\mathfrak{a}}$. J_1 is non-linear and its differential at zero is $\rho_{\mathfrak{a}}$. *What Kostant's theorem states is the remarkable fact that each K-orbit in $\mathfrak{p}$ has the same convex image under both J_1 and $\rho_{\mathfrak{a}}$ and this polytope is the convex hull of the corresponding Weyl group orbit.*

Let's sketch the simplest example: $G = SL(n, \mathbb{C})$, $K = SU(n)$, $P = \{n \times n$ positive definite Hermitian matrices with determinant $1\}$, $N = \{n \times n$ strictly upper triangular complex matrices$\}$, $\mathfrak{g} = s\ell(n, \mathbb{C})$, $\mathfrak{k} = su(n)$, $\mathfrak{a} = \{n \times n$ real diagonal traceless matrices$\}$, $\mathfrak{p} = \{n \times n$ traceless Hermitian matrices $\}$, $\mathfrak{n} = \{n \times n$ strictly upper triangular complex matrices$\}$. The Cartan decompositions are $SL(n, \mathbb{C}) = SU(n)P$ and $s\ell(n, \mathbb{C}) = su(n) \oplus \mathfrak{p}$.

The Iwasawa decompositions are $SL(n,\mathbf{C}) = SU(n)AN$, $s\ell(n,\mathbf{C}) = su(n) \oplus \mathfrak{a} \oplus \mathfrak{n}$. If $X \in \mathfrak{p}$, $\rho_{\mathfrak{a}}(X)$ is the diagonal part of X. On the other hand

$$\begin{aligned} J_1(X) &= (log \circ \rho_A \circ exp)(X) \\ &= \frac{1}{2}\left(log\Delta_1(e^{2x}), log\frac{\Delta_2(e^{2x})}{\Delta_1(e^{2x})}, \ldots, log\frac{\Delta_n(e^{2x})}{\Delta_{n-1}(e^{2x})} \right) \end{aligned}$$

where Δ_k is the determinant of the left upper corner $k \times k$ matrix. For $X = (x_1, \ldots, x_n) \in \mathfrak{a}$, the flag manifold $\mathcal{O}_X$ (the K-orbit in $\mathfrak{p}$ through X) is the set of all Hermitian matrices with $x_1, \ldots, x_n$ as their eigenvalues. The relative Weyl group is in this example the permutation group on n letters.

The expressions appearing in $J_1(X)$ remind one of the explicit solution of the Toda lattice equations (see Kostant [1979]); we don't believe this to be an accident but cannot explain it so far.

Let us outline below the main ideas going into the "symplectic proof" of Kostant's theorem. We begin with the complex case, i.e. G is a complex semisimple Lie group, K is a real compact form of G, and $G = KAN$ is the Iwasawa decomposition. We let $\mathfrak{b} := \mathfrak{a} \oplus \mathfrak{n}$, $B := AN$ and remark as in §7 that $\mathfrak{k}$ is isomorphic to $\mathfrak{b}^*$ via the imaginary part of the Killing form κ. Let $\rho_{\mathfrak{k}}$ and $\rho_{\mathfrak{b}}$ be the projections of $\mathfrak{g}$ on $\mathfrak{k}$ and $\mathfrak{b}$ respectively relative to the Iwasawa decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{b}$. We will denote by R_b the right translation in B by b, its differential, and all its tensorially induced maps. The following theorem is deduced form abstract considerations in Lu, Weinstein [1990], Lu [1990] and carried out "by hand" in Lu, Ratiu [1990].

Theorem 8.2. *On B define a bivector field π by*

$$(R_{b^{-1}}\pi(b))(X,Y) = (Im\,\kappa)(\rho_{\mathfrak{k}}(Ad_{b^{-1}}X), \rho_{\mathfrak{b}}(Ad_{b^{-1}}Y))$$

where $b \in B$, $X, Y \in \mathfrak{k} \cong \mathfrak{b}^*$. Then π defines a Poisson structure on $B = AN$ which is multiplicative.

The property of *multiplicativity* means that π satisfies

$$\pi(b_1 b_2) = L_{b_1}\pi(b_2) + R_{b_2}\pi(b_1)$$

where L_b is left translation by b. A ***Poisson Lie group*** is a Lie group endowed with a multiplicative Poisson structure. Thus theorem 8.2 gives a concrete natural formula for a Poisson Lie group structure on B.

The Iwasawa decomposition at group level $G = KB$ induces a smooth projection $\rho_B : G \to B$, $\rho_B(kb) = b$. Define the K-action on B by

$$\sigma : (k, b) \in K \times B \mapsto \rho_B(kbk^{-1}) \in B.$$

Theorem 8.3. *The symplectic leaves of the Poisson structure π in Theorem 8.1 are the K-orbits in B relative to the action σ.*

Let $\mathfrak{t} = i\mathfrak{a}$ and T be the connected subgroup of K with Lie algebra $\mathfrak{t}$. T is a maximal torus in K.

Theorem 8.4. *The restriction σ to T leaves the Poisson structure π on B invariant and the map*

$$J = \log \circ \rho_A : B = AN \to \mathfrak{a}$$

$$J(an) = \log(a)$$

is the momentum map for this T-action.

It should be noted that the Poisson structure π on B is not K-invariant, but only T-invariant. In Lu, Weinstein [1990] it is shown that there is a natural Poisson structure on K such that the map $\sigma : K \times B \to B$ is a Poisson map if one thinks of $K \times B$ as a product Poisson manifold (see Weinstein [1983]). This Poisson structure on K alluded to above happens to vanish on T and this is why this T-action is Hamiltonian.

The Poisson structure π on B and the T-momentum map $J : B \to \mathfrak{a}$ are the first ingredient in the proof of Kostant's theorem in the complex case. The second ingredient is the convexity Theorem 3.1. P and $AN = B$ are identified via ρ_B (and both of them are diffeomorphic to $K\backslash G$). The action σ of K on B becomes the conjugation of K on P. We will still denote by π the push-forward to P of the Poisson structure on B and so the symplectic leaves are the K-orbits on P. Every K-orbit in P intersects A and the fixed point set of the T-action on the K-orbit is the corresponding Weyl group orbit. Now apply Theorem 3.1 to the equivariant momentum map J of Theorem 8.4 to immediately conclude Kostant's convexity result in Theorem 8.1.

To prove Theorem 8.1 for real Lie groups, a last ingredient is needed: namely Duistermaat's Theorem 3.4. For this, we need to identify an antisymplectic involution on the K-orbits in P. We will do better and find an anti-Poisson involution; it will be induced by a Cartan involution. Here is the construction. Let G be a connected real semisimple Lie group with Lie algebra $\mathfrak{g}$. Without loss of generality we may assume G has trivial center so it admits a complexificaton $G_{\mathbf{C}}$. Then the Lie alegbra of $G_{\mathbf{C}}$ is $\mathfrak{g}_{\mathbf{C}} = \mathfrak{g} \oplus i\mathfrak{g}$, the complexification of $\mathfrak{g}$. Let $\tau(X + iY) = X - iY$ be conjugation in $\mathfrak{g}_{\mathbf{C}}$, $X, Y \in \mathfrak{g}$, and denote also by τ the unique automorphism of $G_{\mathbf{C}}$ whose differential at the identity is the conjugation. The identity component of the fixed point set of τ is G. If $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ is the Cartan decomposition of $\mathfrak{g}$, let $\mathfrak{k}_{\mathbf{C}} = \mathfrak{k} \oplus i\mathfrak{p}$, $\mathfrak{p}_{\mathbf{C}} = i\mathfrak{k} \oplus \mathfrak{p} = i\mathfrak{k}_{\mathbf{C}}$ so that both $\mathfrak{k}_{\mathbf{C}}$ and $\mathfrak{p}_{\mathbf{C}}$ are τ-invariant and the fixed point sets of $\tau \mid \mathfrak{k}_{\mathbf{C}}$, $\tau \mid \mathfrak{p}_{\mathbf{C}}$ are $\mathfrak{k}$ and $\mathfrak{p}$ respectively. $\mathfrak{k}_{\mathbf{C}}$ is a compact real form of $\mathfrak{g}_{\mathbf{C}}$ and $\mathfrak{g}_{\mathbf{C}} = \mathfrak{k}_{\mathbf{C}} \oplus \mathfrak{p}_{\mathbf{C}}$ is a Cartan decomposition of $\mathfrak{g}_{\mathbf{C}}$. Let $\mathfrak{a}$ be a maximal abelian subalgebra of $\mathfrak{p}$, $\mathfrak{a}'$ a maximal abelian subalgebra of $\mathfrak{g}$ containing $\mathfrak{a}$ and let $\mathfrak{a}_{\mathbf{C}} = \mathfrak{a} \oplus i(\mathfrak{a}' \cap \mathfrak{k})$. Then $\mathfrak{a}_{\mathbf{C}}$ is a maximal abelian subalgebra of $\mathfrak{p}_{\mathbf{C}}$ and $\mathfrak{a}$ is the fixed point set of $\tau \mid \mathfrak{a}_{\mathbf{C}}$. Bases in $\mathfrak{a}$ and $\mathfrak{a}_{\mathbf{C}}$ are chosen such that the basis in $\mathfrak{a}_{\mathbf{C}}$ is a prolongation of the basis in $\mathfrak{a}$. Relative to such a choice (and a lexicographic ordering), we have the Iwasawa decomposition

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}, \ \mathfrak{g}_{\mathbf{C}} = \mathfrak{k}_{\mathbf{C}} \oplus \mathfrak{a}_{\mathbf{C}} \oplus \mathfrak{n}_{\mathbf{C}}.$$

$\mathfrak{n}_{\mathbf{C}}$ is τ-invariant and, as before, the fixed point set of $\tau|\mathfrak{n}_{\mathbf{C}}$ is $\mathfrak{n}$. Now pass to the group level and use the same notational scheme.

Theorem 8.5. *Let π be the multiplicative Poisson structure on $B_{\mathbf{C}} = A_{\mathbf{C}}N_{\mathbf{C}}$ defined in Theorem 8.2. Then $\tau|B_{\mathbf{C}}$ is an anti-Poisson automorphism, i.e. $\tau^*\pi = -\pi$.*

Now Kostant's Theorem 8.1 for real Lie groups is a direct consequence of Duistermaat's Theorem 3.4 applied to a K-orbit in P viewed as the fixed point set of τ on the $K_{\mathbf{C}}$-orbit in $P_{\mathbf{C}}$ through the same point. The only missing link is the momentum map $P_{\mathbf{C}} \to \mathfrak{a}$, which, after a suitable identification becomes the Iwasawa projection $P_{\mathbf{C}} \to A$. This is constructed in the following way. Let $T_{\mathbf{C}}$ be the maximal torus of $K_{\mathbf{C}}$ with Lie algebra $i\mathfrak{a}_{\mathbf{C}}$. By Theorem 8.4, $T_{\mathbf{C}}$ leaves the Poisson structure on $P_{\mathbf{C}}$ invariant (recall, we identify $P_{\mathbf{C}}$ and $B_{\mathbf{C}} = A_{\mathbf{C}}N_{\mathbf{C}}$). Let T be the subtorus generated by $\mathfrak{t} = i\mathfrak{a}$. Apply theorem 8.4 to conclude that this action has equivariant momentum map

$$J_T : (k_{\mathbf{C}}a_{\mathbf{C}}n_{\mathbf{C}}) \in P_{\mathbf{C}} \mapsto proj_{\mathfrak{a}}(\log a_{\mathbf{C}})$$

where $k_{\mathbf{C}} \in K_{\mathbf{C}}$, $a_{\mathbf{C}} \in A_{\mathbf{C}}$, $n_{\mathbf{C}} \in N_{\mathbf{C}}$ and $proj_{\mathfrak{a}} : \mathfrak{a}_{\mathbf{C}} \to \mathfrak{a}$ is the orthgonal projection relative to the Killing form. Let $\mathfrak{a}_0$ be the orthogonal complement of $\mathfrak{a}$ in $\mathfrak{a}_{\mathbf{C}}$ and define $A_0 = \exp \mathfrak{a}_0$. Then $A_{\mathbf{C}} = AA_0$ and we can regard J_T as

$$J_T : (k_{\mathbf{C}}aa_0n_{\mathbf{C}}) \in P_{\mathbf{C}} \mapsto a \in \mathfrak{a}$$

for $k_{\mathbf{C}} \in K_{\mathbf{C}}$, $a \in A$, $a_0 \in A_0$, $n_{\mathbf{C}} \in N_{\mathbf{C}}$. As before, this map is τ-invariant. Now restrict the T-action and the momentum map to a $K_{\mathbf{C}}$-orbit in $P_{\mathbf{C}}$. By Theorem 3.4 the image of the τ-fixed point set, which is the K-orbit in P through the same point, has J_T-image in $\mathfrak{a}$ equal to the convex hull of the image of the fixed point set of the T-action on the K-orbit. But this fixed point set is the intersection of the K-orbit with A, i.e. the corresponding Weyl group orbit in A. Finally, J_T restricts to the identity map on A. This proves Theorem 8.1 in the real case.

It is worthwhile noting that the symplectic leaves of π in $\mathfrak{p}$ and the symplectic leaves of the Lie-Poisson structure on $\mathfrak{k}^* = \mathfrak{p}$ coincide. Moreover, a theorem of Conn [1985] guarantees that locally, around zero, these two Poisson structures are isomorphic (the key assumption in Conn's theorem, that $\mathfrak{k}$ is compact semisimple, is automatically fulfilled in our case). We suspect that these Poisson structures are globally isomorphic and hope that the isomorphism is relevant to questions regarding the multi-Hamiltonian structure of the Toda lattice equations. Duistermaat [1984] has already shown that the momentum maps for these two structures can be obtained from each other by a homotopy argument. The hoped for Poisson isomorphism between π and the Lie-Poisson structure cannot be K-invariant since π is not whereas the Lie-Poisson structure is; it must, however, be T-invariant.

References

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

M. F. Atiyah [1983], Angular momentum, convex polyhedra and algebraic geometry, *Proc. Edinburgh Math. Soc.* **26**, 121-138.

A. Besse [1987], *Einstein Manifolds*, Ergebenisse der Mathematik und ihrer Grenzgebiete, 3 Folge, Springer-Verlag, New York.

A. M. Bloch [1990], The Kähler structure of the total least squares problem, Brockett's steepest descent equations, and constrained flows, in (M. Kaashoek, J. H. van Schuppen and A. C. M Ran eds.) *Realization and Modelling in Systems Theory,* Birkhauser, Boston.

A. M. Bloch, R. W. Brockett, T. S. Ratiu [1990a], A new formulation of the generalized Toda lattice equations and their fixed point analysis via the momentum map, *Bull. Amer. Math. Soc.* **23**(2), 477-486.

A. M. Bloch, R. W. Brockett, T. S. Ratiu [1990b], Completeley integrable gradient flows, (to appear).

A. M. Bloch, H. Flaschka, T. S. Ratiu [1990], A convexity theorem for isospectral manifolds of Jacobi matrices in a compact Lie algebra, *Duke Math. Journ.* **61**(1), 41-66.

R. W. Brockett [1988], Dynamical systems that sort lists and solve linear programming problems, *Proc. 27th IEEE Conf. on Decision and Control*, 799-803.

R. W. Brockett [1989], Least squares matching problems, *Lin. Alg. and Appl.* **122/123/124**, 761-777.

M. Condevaux, P. Molino [1988], Géometrie du moment, Preprint, Université des Sciences et Techniques du Languedoc. Montpellier.

J. Conn [1985], Normal forms for smooth Poisson structures, *Ann. of Math.* **121**, 565-593.

M. Davis [1987], Some aspherical manifolds, *Duke Math. Journ.* **55**, 105-138.

M. Davis, T. Janusiewicz [1990], Convex polytopes, Coxeter polytopes and torus action, *Duke. Math. Journ.* (to appear).

P. Deift, T. Nanda, C. Tomei [1983], Differential equations for the symmetric eigenvalue problem, *SIAM Journ. Num. Anal.* **20**, 1-22.

J. J. Duistermaat [1983], Convexity and tightness for restrictions of Hamiltonian functions to fixed point sets of an anitsymplectic involution, *Trans. Amer. Math. Soc.* **275**, 417-429.

J. J. Duistermaat [1984], On the similarity between the Iwasawa projection and the diagonal part, *Soc. Mat. de France, 2e Série, Mémoire* **15**, 129-138.

J. J. Duistermaat, J. A. Kolk, V. S. Varadarajan [1983], Functions, flows, and oscillatory integrals on flag manifolds and conjugacy classes in real semisimple Lie groups, *Compositio Math.*, **49**, 309-398.

N. Ercolani, H. Flaschka, L. Haine [1990], The level varieties of the complex Toda equations, preprint.

H. Flaschka, L. Haine [1990], Torus orbits in G/P, *Pac. Journ. Math.* (to appear).

D. Fried [1986], The cohomology of an isospectral flow, *Proc. Amer. Math. Soc.* **98**, 363-368.

R. Goodman, N. Wallach [1984], Classical and quantum mechanical systems of Toda lattice type, II, *Comm. Math Phys.* **94**, 177-217.

V. Guillemin, S. Sternberg [1982], Convexity properties of the moment mapping, *Invent. Math.* **67**, 491-513.

V. Guillemin, S. Sternberg [1984], Convexity properties of the moment mapping II, *Invent. Math.* **77**, 533-546.

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

J. Humphreys [1972], *Introduction to Lie Algebras and Representation Theory*, Springer-Verlag, New York.

A. Kirillov [1976], *Elements of the Theory of Representations*, Springer-Verlag, Berlin.

F. Kirwan [1984], Convexity properties of the moment mapping III, *Invent. Math.* **77**, 547-552.

B. Kostant [1970], Quantization and unitary representations, in *Lecture Notes in Math.* **700**, Springer-Verlag, New York.

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

B. Kostant [1979], The solution to a generalized Toda lattice and representation theory, *Adv. in Math.* **34**, 195-338.

J. Lagarias [1988], Monotonicity properties of the generalized Toda flow and QR flow, *SIAM Journ. Matrix Analysis and Applications*, (to appear).

J.-H. Lu [1990], Mulitplicative and affine Poisson structures on Lie groups, Ph. D. Thesis, U. C. Berkeley.

J.-H. Lu, T. Ratiu [1990], On the nonlinear convexity theorem of Kostant, preprint.

J.-H. Lu, A. Weinstein [1990], Poisson Lie groups, dressing transformations and Bruhat decomposition, *Journ. Diff. Geom.* **31**, 501-526.

J. Moser [1975], Finitely many mass points on the line under the influence of an exponential potential–an integrable system, *Springer Lecture Notes in Physics* **38**, 467-497.

A. Reyman, M. Semenov-Tijan-Shanskii [1979], Reduction of Hamiltonian system, affine Lie algebras, and Lax equations; I, *Invent. Math.* **54**, 81-100; II *ibid* **63** (1981), 423-432.

D. H. Sattinger and D. L. Weaver [1986], *Lie Groups and Lie Algebras with Applications to Physics, Geometry and Mechanics,* Applied Mathematical Sciences 69, Springer-Verlag, New York.

I. Schur [1923], Über eine Klasse von Mittelbildungen mit Anwendungen auf der Determinantentheorie, *Stizungsberichte der Berliner Math. Gesellschaft* **22**, 9-20.

S. Smale [1970], Topology and Mechanics; I, *Invent. Math.* **10**, 305-311; II *ibid.* **11**, 45-64.

J.-M. Souriau [1970], *Structure des Sytemes Dynamiques*, Dunod, Paris.

W. W. Symes [1980], Hamiltonian group actions and integrable systems, *Physica D* **1**, 39-376.

W. W. Symes [1982a], Systems of Toda type, inverse spectral problems and representation theory, *Invent. Math.* **59** (1982), 13-51.

W. W. Symes [1982b], The QR algorithm and scattering for the nonperiodic Toda lattice, *Physica D* 4, 275-280.

C. Tomei [1984], The topology of isospectral manifolds of tridiagonal matrices, *Duke Math. Journal* **51**, 981-996.

P. van Moerbeke [1976], The specturm of Jacobi matrices, *Invent. Math.* **37**, 45-81.

A. Weinstein [1983], The local structure of Poisson manifold, *J. Diff. Geom.* **18**, 523-557.