

Completely Integrable Gradient Flows

Anthony M. Bloch^{1,*}, Roger W. Brockett^{2,**} and Tudor S. Ratiu^{3,***}

¹ Department of Mathematics, The Ohio State University, Columbus, OH 43210, USA

² Division of Applied Sciences, Harvard University, Cambridge, MA 02138, USA

³ Department of Mathematics, University of California, Santa Cruz, CA 95064, USA

Received December 6, 1990, in revised form July 11, 1991

Abstract. In this paper we exhibit the Toda lattice equations in a double bracket form which shows they are gradient flow equations (on their isospectral set) on an adjoint orbit of a compact Lie group. Representations for the flows are given and a convexity result associated with a momentum map is proved. Some general properties of the double bracket equations are demonstrated, including a discussion of their invariant subspaces, and their function as a Lie algebraic sorter.

0. Introduction

In this paper we present details of the proofs and applications of the results on the generalized Toda lattice equations announced in [7].

One of our key results is exhibiting the Toda equations in a double bracket form which shows directly that they are gradient flow equations (on their isospectral set) on an adjoint orbit of a compact Lie group. This system, which is gradient with respect to the normal metric on the orbit, is quite different from the representation of the Toda flow as a gradient system on \mathbb{R}^n in Moser's fundamental paper [27]. In our representation the same set of equations is thus Hamiltonian and a gradient flow on the isospectral set.

While the double bracket equations mentioned above are fundamental in demonstrating properties of the Toda flow, they are in fact more general and are of interest in their own right. These equations arose originally in the study by Brockett (see [10 and 11]) of the steepest descent equations corresponding to certain least squares matching and sorting problems. In [5] it was noticed that a

* Supported in part by NSF Grant DMS-90-02136, NSF PYI Grant DMS-9157556, and a Seed Grant from Ohio State University

** Supported in part by AFOSR grant AFOSR-96-0197, by U.S. Army Research Office grant DAAL03-86-K-0171 and by NSF grant CDR-85-00108

*** Supported in part by NSF Grant DMS-8922699

particular case of these equations give the Toda ($sl(n)$) equations in the Flaschka ([15]) form. In fact it can be shown that these equations have many different invariant subspaces as we prove here.

Just as it can be shown that the time-1 map of the generalized Toda flows is equivalent to the QR algorithm for diagonalizing symmetric matrices (see [14, 25, 30 and 31], for example), these double bracket equations provide a method for solving the symmetric eigenvalue problem, which coincides with the QR algorithm in the tridiagonal case. Recall also that the Toda equations in Flaschka's form sort the eigenvalues of a given symmetric matrix into ascending order. Here we show that the double bracket equations give a general Lie-algebraic sorter, which will sort in *any* prescribed fashion (corresponding to a choice of Weyl chamber).

We also show here how the equilibria of the double bracket equations are the vertices of a convex polytope which is the image of a momentum mapping associated with the problem (see the work of Schur–Horn [19], Kostant [21], Atiyah [3] and Guillemin and Sternberg [17]). The image of the isospectral Toda orbit is shown to be a complex object lying in the interior of the polytope defined by the momentum mapping. That a convex polytope may be associated with the Toda flows, was observed originally by Deift Nanda and Tomei [14] and Tomei [33] (see also van Moerbeke [34], Fried [16] and Davis [13]). In a related paper [6], by exploiting the Kähler structure of the problem, we show that the isospectral Toda orbit may in fact be mapped onto the polytope by a momentum mapping, thus putting the topological observation of Tomei et al. into a symplectic context.

The outline of this paper is as follows. In Sect. 1 we discuss the double bracket equations, indicate how they arise as gradient flows and describe certain of their invariant subspaces. In Sect. 2 we show how a special case of these equations gives the generalized Toda lattice equations (see Kostant [22, 23], Symes [30] and others). A Lie algebraic explanation is given for the double bracket form of the equations in the Toda case. In Sect. 3 we give explicit representations for the Toda flows in double bracket form for the classical Lie algebras. In Sect. 4 we discuss the convexity and momentum map result and the sorting result mentioned above.

1. The Double Bracket Equations and Gradient Flows

In this section we begin by discussing gradient flows on compact Lie groups. We derive an explicit equation for a particular gradient flow relative to a left-invariant Riemannian metric. We then show that projecting the flow onto an adjoint orbit of the Lie algebra by taking the natural projection yields a flow which is itself gradient with respect to the “normal” metric on the orbit. This flow is shown to be of the “double bracket” form $\dot{L}(t) = [L(t), [L(t), N]]$ where N is a given constant matrix. This flow is isospectral and we show that in fact this double bracket flow is the compatibility condition for the “spectral equations” for the flow at the group level. This is related to the classical theory of integrability of Hamiltonian systems as we shall see in the next section. Finally, we discuss some invariant subspaces of the double bracket equation.

Let G_u be a compact Lie group with Lie algebra \mathcal{G}_u and let $\kappa(\cdot)$ denote the Killing form on \mathcal{G}_u . Define the smooth function $F(\theta)$ on G_u by

$$F(\theta) = \kappa(Q, \text{Ad}_\theta N) \tag{1.1}$$

where Q and N are fixed elements of \mathcal{G}_u and Ad denotes the adjoint action of G_u on \mathcal{G}_u . Then we have (see [10] for the unitary case)

Proposition 1.1. *The gradient flow of $F(\theta)$ relative to the left-invariant Riemannian metric on G_u whose value at the identity is minus the Killing form is given by*

$$\dot{\theta}(t) = \theta(t) \cdot [\text{Ad}_{\theta(t)^{-1}} Q, N], \quad (1.2)$$

where $\theta \cdot P$ denotes the left translation of P by θ .

Proof. Let $v_\theta = \theta \cdot R$ where $R \in \mathcal{G}_u$ and denote by $\langle \cdot, \cdot \rangle$ the metric on G_u obtained by left-translation of $-\kappa(\cdot, \cdot)$. Then

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

whence

$$\nabla F(\theta) = \theta \cdot [\text{Ad}_{\theta^{-1}} Q, N]. \quad \square$$

Let \mathcal{G}_u be the compact real form of a complex semisimple Lie algebra \mathcal{G} . Then the equation (1.2) have the normal real form of \mathcal{G} intersect the compact real form as an invariant submanifold. The gradient flow can then be formulated as a gradient flow with respect to the normal metric on this intersection. Equation (1.2) first appeared in this form on $SO(n)$ (see [10]).

Corollary 1.2. *The projection of the flow (1.2) onto the adjoint orbit containing Q , obtained by setting $L(t) = \text{Ad}_{\theta(t)^{-1}} Q$ is given by*

$$\dot{L}(t) = [L(t), [L(t), N]]. \quad (1.3)$$

Indeed, by the Chain Rule and Proposition 1.1 we get

$$\dot{L} = [L, \theta^{-1} \dot{\theta}] = [L, [L, N]].$$

It is interesting to notice that Corollary 1.2 can be interpreted as a compatibility condition between an ‘‘eigenvalue equation’’ and the time evolution of the ‘‘eigenfunction.’’ Indeed $L(t) = \text{Ad}_{\theta(t)^{-1}} Q$ says that $L(t)$ and Q have the same spectrum, if they are matrices. In other words, this relation can be thought of as the totality of eigenvalue equations for L , with the columns of θ as eigenfunctions. The time evolution for θ is given by (1.2). The usual cross-differentiation argument repeats the proof of Corollary 1.2.

Corollary 1.3. *The Eq. (1.3) is the compatibility condition for the system $L(t) = \text{Ad}_{\theta(t)^{-1}} Q$ and $\dot{\theta}(t) = \theta(t) \cdot [L(t), N]$.*

Note that the flow (1.3) is the image of (1.2) under the quotient map $G_u \rightarrow G_u/H$, H the stabilizer of $L(0) = L_0$, where G_u/H is identified with the adjoint orbit \mathcal{O}

through L_0 . This “projected” flow turns out to be itself a gradient flow on the adjoint orbit \mathcal{O} endowed with the “standard” or “normal” metric (see Atiyah [3]). Explicitly this metric is given as follows.

Decompose orthogonally, relative to $-\kappa(\cdot, \cdot) = \langle \cdot, \cdot \rangle$, $\mathcal{G}_u = \mathcal{G}_u^L \oplus \mathcal{G}_{uL}$, where \mathcal{G}_{uL} is the centralizer of L and $\mathcal{G}_u^L = \text{Im ad } L$. For $X \in \mathcal{G}_u$ denote by $X^L \in \mathcal{G}_u^L$ the orthogonal projection of X on \mathcal{G}_u^L . Then set the inner product of the tangent vectors $[L, X]$ and $[L, Y]$ to be equal to $\langle X^L, Y^L \rangle$. Denote this metric by $\langle \cdot, \cdot \rangle_N$. Then we have

Proposition 1.4. *The flow (1.3) is the gradient vector field of $H(L) = \kappa(L, N)$, κ the Killing form, on the adjoint orbit \mathcal{O} of \mathcal{G}_u containing the initial condition $L(0) = L_0$, with respect to the normal metric $\langle \cdot, \cdot \rangle_N$ on \mathcal{O} .*

Proof. We have, by the definition of the gradient,

$$dH \cdot [L, \delta L] = \langle \text{grad } H, [L, \delta L] \rangle_N, \tag{1.4}$$

where \cdot denotes the natural pairing between 1-forms and tangent vectors and $[L, \delta L]$ is a tangent vector at L . Set $\text{grad } H = [L, X]$. Then (1.4) becomes

$$-\langle [L, \delta L], N \rangle = \langle [L, X], [L, \delta L] \rangle_N$$

or

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

Thus

$$X^L = ([L, N])^L = [L, N]$$

and

$$\text{grad } H = [L, [L, N]]$$

as required. \square

The “double bracket” flow (1.3) is in the “Lax pair” form $\dot{L} = [B, L]$ and is thus isospectral. This of course plays a key role in our Toda lattice analysis.

Now for the Toda lattice analysis we need to consider the invariant subspaces of the equation $\dot{L} = [L, [L, N]]$. We prove here some general results regarding the invariant subspaces of this equation. For this analysis we may take L and N to be the elements of any Lie algebra \mathcal{G} over the real or complex field. The differential equation $\dot{L} = [L, [L, N]]$ then evolves in \mathcal{G} .

Proposition 1.5. *Let $r \in \mathbb{R}$ and let \mathcal{G} be a Lie algebra. Let $N \in \mathcal{G}$ and suppose $(\text{ad } N)^2$ has an eigenvalue r^2 with corresponding eigenspace V in \mathcal{G} . Let \mathcal{G}_N be the centralizer of N . Then $\mathcal{G}_N \oplus V$ is an invariant subspace for the nonlinear operator on \mathcal{G} given by $L \rightarrow [L, [L, N]]$, $L \in \mathcal{G}$.*

Proof. Set $V = V_r \oplus V_{-r}$, where V_r and V_{-r} are the $+r$ and $-r$ eigenspaces, respectively, of the operator $\text{ad } N$. We can write a typical element of $\mathcal{G}_N \oplus V$ as $v_+ + v_- + h$ with $[h, N] = 0$, $[N, v_+] = rv_+$ and $[N, v_-] = -rv_-$. By the Jacobi identity, we have $[N, [v_+, v_-]] = 0$, and $[N, [h, v_+]] = r[h, v_+]$ so that $[h, v_+] \in v_r$. Similarly, $[h, v_-] \in v_{-r}$. Together these identities imply

$$[v_+ + v_- + h, [v_+ + v_- + h, N]] \in \mathcal{G}_N \oplus V. \quad \square$$

We also have the following.

Proposition 1.6. *Suppose e is an involutive Lie algebra automorphism of \mathcal{G} . Let $\mathcal{G}_{\text{even}}$ and \mathcal{G}_{odd} be the $+1$ and -1 eigenspaces of e . Then if $\text{ad } N$ maps \mathcal{G}_{odd} into $\mathcal{G}_{\text{even}}$, it follows that $(\mathcal{G}_N \oplus V) \cap \mathcal{G}_{\text{odd}}$ is an invariant subspace for the operator $[\cdot, [\cdot, N]]$.*

This follows from Proposition 1.5 and the obvious relation $[\mathcal{G}_{\text{odd}}, \mathcal{G}_{\text{even}}] \subseteq \mathcal{G}_{\text{odd}}$.

The Toda flows will be seen to evolve in a subspace of the form defined in Proposition 1.6 with $r^2 = 1$.

2. The Double Bracket Equations and the Toda Lattice

In this section we show that the generalized Toda lattice equations may be written in the double bracket form (1.3) and hence are a gradient flow on their isospectral set.

Let \mathcal{G}_n be the normal real form of \mathcal{G} .

Then the Flaschka form of the generalized Toda lattice equations associated to \mathcal{G}_n are given by $\dot{A} = [B, A]$ where

$$A = \sum_{j=1}^l b_j h_j + \sum_{j=1}^l a_j (e_{\alpha_j} + e_{-\alpha_j}), \quad (2.1)$$

$$B = \sum_{j=1}^l a_j (e_{\alpha_j} - e_{-\alpha_j}), \quad a_j > 0. \quad (2.2)$$

This form of the Toda lattice equations is the one adopted in Moser [27] for $\mathcal{G}_n = \mathfrak{sl}(n, \mathbb{R})$, and is equivalent to, but different from, the conventions used in Kostant [22] and Symes [30, 31]. This is a Hamiltonian system on the coadjoint orbit of the lower Borel subalgebra of \mathcal{G}_n through $\sum_{j=1}^l (e_{\alpha_j} + e_{-\alpha_j})$ called the *Jacobi orbit*, with Hamiltonian function $\frac{1}{2}\kappa(A, A)$. All the elements of this orbit which are of the form (2.1) we call Jacobi elements. This system is integrable; the integrals are the basis of the ring of invariants of \mathcal{G}_n restricted to the orbit. Level sets of the integrals of motion are called isospectral sets.

In what follows, let \mathcal{G}_u be a compact Lie algebra, \mathcal{A} a maximal abelian subalgebra and \mathcal{G} the complexification of \mathcal{G}_u , and choose $\mathcal{H} = \mathcal{A} \oplus i\mathcal{A}$ as the Cartan subalgebra of \mathcal{G} . We let Φ denote the system of roots of \mathcal{G} defined by \mathcal{H} , let Δ denote the simple roots and choose $\{h_j, e_\alpha | j = 1, \dots, l, \alpha \in \Phi\}$, a Chevalley basis of \mathcal{G} , with $h_j \in \mathcal{A}$. $\alpha_1, \dots, \alpha_l$ denote this simple roots.

Let G and G_u denote the Lie groups with Lie algebra \mathcal{G} and \mathcal{G}_u respectively.

Theorem 2.1. *If N is i times the sum of the simple coweights of \mathcal{G} , then, for*

$$L = \sum_{j=1}^l i b_j h_j + \sum_{j=1}^l i a_j (e_{\alpha_j} + e_{-\alpha_j}), \quad (2.3)$$

Eq. (1.3) gives the generalized Toda lattice equations on the level set of all the integrals of the Toda flow. Explicitly, N is given by

$$N = \sum_j i x_j h_j, \quad (2.4)$$

where (x_1, x_2, \dots, x_l) is the unique solution of the system

$$\sum_j x_j \alpha_i(h_j) = -1, \quad i = 1, \dots, l. \quad (2.5)$$

Proof. From (2.3) and (2.4),

$$[L, N] = \sum_{i=1}^l \sum_{j=1}^l x_j a_i \alpha_i(h_j)(e_{\alpha_i} - e_{-\alpha_i}).$$

But for the generalized Toda flow, $\dot{L} = [B, L]$ where B is given by (2.2)

It is known that there is a unique element N in the Cartan subalgebra (see [20], p. 1014) such that $[N, e_\alpha] = \text{height}(\alpha)e_\alpha$. This element is found in the following way. Let $\tilde{\alpha}_i = \frac{2\alpha_i}{(\alpha_i, \alpha_i)}$, where $(,)$ denotes inner product, be a dual root and let $\tilde{\lambda}_i$,

where $(\tilde{\lambda}_i, \alpha_j) = \delta_{ij}$, be the corresponding dual weight. Set $h_{\tilde{\delta}} = \sum_{i=1}^l \tilde{\lambda}_i$. Then $(h_{\tilde{\delta}}, e_{\alpha_i}) = 1$ and we set $N = -ih_{\tilde{\delta}}$. Note that N is the Cartan part of a principal $sl(2)$.

Explicitly, $[L, N] = -B$ if and only if the coefficients x_i in (2.4) are the solution of the system

$$\sum_{j=1}^l x_j \alpha_i(h_j) = -1.$$

It is a standard result that $\alpha_i(h_j)$ is the Cartan matrix of \mathcal{G} . \square

We list now the coefficients x_i (i.e. the solutions of the equation $Cx = -1$, C the Cartan matrix for all the semi-simple Lie algebras (see [8]):

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

Table 1

Now we can show quite intrinsically that Eq. (1.3) with L and N given by (2.3) and (2.4) respectively are the Toda lattice equations; that is, the Toda Hamiltonian on a coadjoint orbit of the lower Borel subalgebra of \mathcal{G}_n gives rise to the Eq. (1.3)

We first derive an explicit formula for the Kostant–Kirillov–Souriau (orbit) form on the coadjoint orbit of the lower Borel subalgebra. Before we do this we need some notation and definitions:

Let Φ , Φ^+ , and Φ^- denote the system of roots, the positive and negative roots of \mathcal{G} respectively. Let

$$\mathcal{G}_n = \left\{ \sum_{j=1}^l b_j h_j + \sum_{\alpha \in \Phi} c_\alpha e_\alpha \mid b_j, c_\alpha \in \mathbb{R} \right\} \quad (2.6)$$

be the normal real form of \mathcal{G} .

Now consider the decomposition

$$\mathcal{G}_n = \mathcal{L} \oplus \mathcal{K}$$

where

$$\mathcal{L} = i\mathcal{A} \oplus \bigsqcup_{\alpha \in \Phi^-} \mathcal{G}_\alpha, \quad (2.7)$$

\mathcal{G}_α the \mathbb{R} span of the α root space, and let

$$\mathcal{K} = \left\{ \sum_{\alpha \in \Phi^+} c_\alpha (e_\alpha - e_{-\alpha}) \mid c_\alpha \in \mathbb{R} \right\}. \quad (2.8)$$

Then the complements of \mathcal{L} and \mathcal{K} under $-\kappa(\cdot)$ are

$$\mathcal{L}^\perp = \bigsqcup_{\alpha \in \Phi^-} \mathcal{G}_\alpha \quad (2.9)$$

and

$$\mathcal{K}^\perp = i\mathcal{A} \oplus \left\{ \sum_{\alpha \in \Phi^+} d_\alpha (e_\alpha - e_{-\alpha}) \mid d_\alpha \in \mathbb{R} \right\}. \quad (2.10)$$

Also

$$\mathcal{G} \cong \mathcal{G}^* = \mathcal{L}^\perp \oplus \mathcal{K}^\perp \quad (2.11)$$

and

$$\mathcal{L}^* \cong \mathcal{K}^\perp, \quad \mathcal{K}^* \cong \mathcal{L}^\perp. \quad (2.12)$$

Let $\pi_{\mathcal{L}}$, $\pi_{\mathcal{K}}$, $\pi_{\mathcal{K}^\perp}$ and $\pi_{\mathcal{L}^\perp}$ denote the associated projections.

Now consider the tri-diagonal elements of \mathcal{K} and \mathcal{K}^\perp , i.e. if Δ denotes the set of simple roots of \mathcal{G} , set

$$\bar{\mathcal{K}} = \left\{ \sum_{\alpha \in \Delta} c_\alpha (e_\alpha - e_{-\alpha}) \mid c_\alpha \in \mathbb{R} \right\} \quad (2.13)$$

$$\bar{\mathcal{K}}^\perp = i\mathcal{A} \oplus \left\{ \sum_{\alpha \in \Delta} a_\alpha (e_\alpha + e_{-\alpha}) \mid a_\alpha \in \mathbb{R} \right\}. \quad (2.14)$$

Let $M = -iN \in i\mathcal{A}$, i.e. $M = \sum_{j=1}^l x_j h_j$, where x_1, \dots, x_l is the solution of the linear system (2.5).

Then $\text{ad}_M: \bar{\mathcal{K}} \rightarrow \bar{\mathcal{K}}^\perp \ominus i\mathcal{A}$ is the isomorphism sending an element of $\bar{\mathcal{K}}$ to minus its symmetrization and $\text{ad}_M: \bar{\mathcal{K}}^\perp \rightarrow \bar{\mathcal{K}}$ sends an element $A \in \bar{\mathcal{K}}^\perp$ to minus its skew-symmetrization (i.e. to $-B$ in (2.22)) and has kernel equal to $i\mathcal{A}$. Thus $\text{ad}_M: \bar{\mathcal{K}}^\perp \ominus i\mathcal{A} \rightarrow \bar{\mathcal{K}}$ is also an isomorphism. The map $\text{ad}_M^2: \bar{\mathcal{K}}^\perp \rightarrow \bar{\mathcal{K}}^\perp \ominus i\mathcal{A}$ is the projection.

For a fixed $A \in \bar{\mathcal{K}}^\perp$ with all $a_\alpha > 0$, $\alpha \in \Delta$ (see (2.1)), the operator $\text{ad}_A: i\mathcal{A} \rightarrow \bar{\mathcal{K}}$ is an isomorphism. Indeed, proceeding as in the proof of Theorem 2.1 and denoting

$$P = \sum_{\alpha \in \Delta} p_\alpha (e_\alpha - e_{-\alpha}), \quad p_\alpha \in \mathbb{R}, \quad (2.15)$$

$$Z = \sum_{i=1}^l z_i h_i, \quad z_i \in \mathbb{R}, \quad (2.16)$$

we have $[A, Z] = P$ if and only if

$$\sum_{j=1}^l z_j \alpha_j(h_j) = -p_{\alpha_i}/a_{\alpha_i}, \quad i = 1, \dots, l. \quad (2.17)$$

Denote by $\text{ad}_A^{-1}: \bar{\mathcal{K}} \rightarrow i\mathcal{A}$ the inverse of this isomorphism.

The operator $\text{ad}_{[A, M]}: i\mathcal{A} \rightarrow \bar{\mathcal{K}}^\perp \ominus i\mathcal{A}$ is also an isomorphism, for if Z is as in (2.16) and

$$Q = \sum_{\alpha \in \Delta} q_\alpha (e_\alpha + e_{-\alpha}), \quad q_\alpha \in \mathbb{R}, \quad (2.18)$$

then $[[A, M], Z] = Q$ if and only if (2.17) holds with p_{α_i} replaced by q_{α_i} . Let $\text{ad}_{[A, M]}^{-1}: \bar{\mathcal{K}}^\perp \ominus i\mathcal{A} \rightarrow i\mathcal{A}$ denote the inverse of the isomorphism.

Theorem 2.3. *The orbit symplectic form on a co-adjoint orbit of the lower Borel subalgebra of \mathcal{G}_n is given by*

$$\omega = \frac{1}{2}(dA \wedge_\kappa (\text{ad}_A^{-1} + \text{ad}_{[M, A]}^{-1}) \text{ad}_M) \text{ad}_M dA) \quad (2.19)$$

where \wedge_κ is the wedge product associated with the bilinear form κ .

Proof. For $X, Y \in \mathcal{L}$, the orbit symplectic form on a Jacobi orbit is given by

$$\omega(\pi_{\mathcal{X}^\perp}[A, X], \pi_{\mathcal{X}^\perp}[A, Y]) = -\kappa(A, [X, Y]).$$

Consider then

$$\begin{aligned} & \frac{1}{2}(dA \wedge_\kappa (\text{ad}_A^{-1} + \text{ad}_{[M, A]}^{-1}) \text{ad}_M) \text{ad}_M dA) (\pi_{\mathcal{X}^\perp}[A, X], \pi_{\mathcal{X}^\perp}[A, Y]) \\ &= \frac{1}{2} \kappa(\pi_{\mathcal{X}^\perp}[A, X], (\text{ad}_A^{-1} + \text{ad}_{[M, A]}^{-1}) \text{ad}_M) \text{ad}_M \pi_{\mathcal{X}^\perp}[A, Y]) \\ & \quad - \frac{1}{2} \kappa(\pi_{\mathcal{X}^\perp}[A, Y], (\text{ad}_A^{-1} + \text{ad}_{[M, A]}^{-1}) \text{ad}_M) \text{ad}_M \pi_{\mathcal{X}^\perp}[A, X]) \end{aligned}$$

Since

$$\text{ad}_M \pi_{\mathcal{X}^\perp}[A, Y] = -\pi_{\mathcal{X}}[A, Y] \in \bar{\mathcal{K}}, \quad \text{ad}_A^{-1} \pi_{\mathcal{X}}[A, Y] \in i\mathcal{A} \subset \bar{\mathcal{K}}^\perp,$$

and

$$(\text{ad}_M)^2 \pi_{\mathcal{X}^\perp}[A, Y] = \pi_{\mathcal{X}^\perp}[A, Y] - \text{diag}[A, Y] \in \bar{\mathcal{K}}^\perp \ominus i\mathcal{A}$$

the above equals:

$$\begin{aligned} & \frac{1}{2} \kappa([A, X], -(\text{ad}_A^{-1}) \pi_{\mathcal{X}}[A, Y] + \text{ad}_{[M, A]}^{-1} \text{ad}_M^2 \pi_{\mathcal{X}^\perp}[A, Y]) \\ & \quad - \frac{1}{2} \kappa([A, Y], -(\text{ad}_A^{-1}) \pi_{\mathcal{X}}[A, X] + \text{ad}_{[M, A]}^{-1} \text{ad}_M^2 \pi_{\mathcal{X}^\perp}[A, X]) \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{2} \kappa(X, \pi_{\mathcal{X}}[A, Y] - \text{ad}_A \text{ad}_{[M, A]}^{-1} \text{ad}_M^2 \pi_{\mathcal{X}^\perp}[A, Y]) \\
 &\quad - \frac{1}{2} \kappa(Y, \pi_{\mathcal{X}}[A, X] - \text{ad}_A \text{ad}_{[M, A]}^{-1} \text{ad}_M^2 \pi_{\mathcal{X}^\perp}[A, X]) \\
 &= \frac{1}{2} \kappa(\pi_{\mathcal{Z}^\perp} X, \pi_{\mathcal{X}}[A, Y]) - \frac{1}{2} \kappa(\pi_{\mathcal{Z}^\perp} Y, \pi_{\mathcal{X}}[A, X]) \\
 &\quad - \frac{1}{2} \kappa(X, [A, Z]) + \frac{1}{2} \kappa(Y, [A, Z']),
 \end{aligned}$$

where $\text{ad}_{[M, A]}^{-1} \text{ad}_M^2 \pi_{\mathcal{X}^\perp}[A, Y] = Z$, and $\text{ad}_{[M, A]}^{-1} \text{ad}_M^2 \pi_{\mathcal{X}^\perp}[A, X] = Z' \in i\mathcal{A}$, and we also have $Z = -\text{diag } Y$, $Z' = -\text{diag } X$, projections into $i\mathcal{A}$ of Y and X respectively. Therefore we get

$$\begin{aligned}
 &= \frac{1}{2} \kappa(\pi_{\mathcal{Z}^\perp} X, [A, Y]) - \frac{1}{2} \kappa(\pi_{\mathcal{Z}^\perp} Y, [A, X]) \\
 &\quad + \frac{1}{2} \kappa(X, [A, \text{diag } Y]) - \frac{1}{2} \kappa(Y, [A, \text{diag } X]) \\
 &= \frac{1}{2} \kappa(A, [Y, \pi_{\mathcal{Z}^\perp} X] + [Y, \text{diag } X]) \\
 &\quad - \frac{1}{2} \kappa(A, [X, \pi_{\mathcal{Z}^\perp} Y] + [X, \text{diag } Y]) \\
 &= -\kappa(A, [X, Y])
 \end{aligned}$$

as required. \square

Then we have the

Theorem 2.4. *For the orbit symplectic form given in Theorem 2.3 the equations of motion on the Jacobi coadjoint orbit of the lower Borel subalgebra of \mathcal{G}_n corresponding to the Hamiltonian $\frac{1}{2} \kappa(A, A)$ are $A = [A, [A, -M]]$.*

Proof. $\omega([A, [A, -M]], \pi_{\mathcal{X}^\perp}[A, X])$

$$\begin{aligned}
 &= \frac{1}{2} \kappa([A, [A, -M]], (\text{ad}_A^{-1} + \text{ad}_{[M, A]}^{-1} \text{ad}_M) \text{ad}_M (\pi_{\mathcal{X}^\perp}[A, X]) \\
 &\quad - \frac{1}{2} \kappa(\pi_{\mathcal{X}^\perp}[A, X], (\text{ad}_A^{-1} + \text{ad}_{[M, A]}^{-1} \text{ad}_M) \text{ad}_M ([A, [A, -M]]) \\
 &= \frac{1}{2} \kappa([A, M], [M, \pi_{\mathcal{X}^\perp}[A, X]]) - \frac{1}{2} \kappa(\pi_{\mathcal{X}^\perp}[A, X], Z) \\
 &\quad + \frac{1}{2} \kappa([A, [A, -M]], Z') - \frac{1}{2} \kappa(\pi_{\mathcal{X}^\perp}[A, X], Z''),
 \end{aligned}$$

where $Z, Z', Z'' \in i\mathcal{A}$ are uniquely determined by the conditions

$$[A, Z] = [M, [A, [A, -M]]], \quad [[M, A], Z'] = \text{ad}_M^2 \pi_{\mathcal{X}^\perp}[A, X],$$

and $[[M, A], Z''] = \text{ad}_M^2 [A, [A, -M]]$. Since $[A, M] = B$ given by (2.2) and ad_M^2 erases the Cartan part of elements in $\tilde{\mathcal{X}}^\perp$, the last equation is equivalent to $[B, Z''] = \text{ad}_M^2 [A, B]$ which in turn determines $Z'' = -\text{diag } A$. Thus we get

$$\begin{aligned}
 &\frac{1}{2} \kappa(\pi_{\mathcal{X}^\perp}[A, X], A - \text{diag } A) + \frac{1}{2} \kappa([X, A], Z) \\
 &\quad + \frac{1}{2} \kappa(A, [[M, A], Z']) + \frac{1}{2} \kappa(\pi_{\mathcal{X}^\perp}[A, X], \text{diag } A) \\
 &= \frac{1}{2} \kappa(\pi_{\mathcal{X}^\perp}[A, X], A) + \frac{1}{2} \kappa(X, [M, [A, [A, -M]]]) + \frac{1}{2} \kappa(A, \text{ad}_M^2 \pi_{\mathcal{X}^\perp}[A, X]) \\
 &= \frac{1}{2} \kappa(A, \pi_{\mathcal{X}^\perp}[A, X]) + \frac{1}{2} \kappa(X, [A, \text{ad}_M^2 A]) + \frac{1}{2} \kappa(A, \pi_{\mathcal{X}^\perp}[A, X] - \text{diag } [A, X]) \\
 &= \kappa(A, \pi_{\mathcal{X}^\perp}[A, X]) + \frac{1}{2} \kappa(X, [A, A - \text{diag } A]) - \frac{1}{2} \kappa(A, \text{diag } [A, X]) \\
 &= \kappa(A, \pi_{\mathcal{X}^\perp}[A, X]) - \frac{1}{2} \kappa([X, A], \text{diag } A) - \frac{1}{2} \kappa(\text{diag } A, [A, X]) \\
 &= \kappa(A, \pi_{\mathcal{X}^\perp}[A, X]) = dH(A) \cdot \pi_{\mathcal{X}^\perp}[A, X]. \quad \square
 \end{aligned}$$

We remark that the Toda lattice is gradient on its restriction to the isospectral set, since it is the restriction of a gradient vector field on a G_u orbit. There appears to be no appropriate metric off the isospectral set. Similarly, Moser's form of the gradient flow (see [27]) also occurs on the isospectral set and is given by

$$\frac{dr_k}{dt} = -\frac{\partial V}{\partial r_k}, \quad k = 1, \dots, n,$$

where

$$V = \frac{\sum_{k=1}^n \lambda_k r_k^2}{2 \sum_{k=1}^n r_k^2}.$$

Here the λ_i are constant and satisfy $\lambda_1 < \lambda_2 < \dots < \lambda_n$, and the r_k are strictly positive and satisfy $\sum_{k=1}^n r_k^2 = 1$. For the precise relationship between Moser's flow and the double bracket flow see [6].

We remark also that for $sl(l+1)$ our function $\kappa(L, N)$ is (modulo a trivial normalization) the Morse function discussed in [13, 16 and 33] and used to analyze the topology of isospectral sets of Jacobi (symmetric, tridiagonal) matrices. We have shown that this Morse function gives the Toda flow as its gradient flow with respect to a suitable metric, and, moreover, we have given a natural generalization of this function to arbitrary semisimple Lie algebras.

We note also that in view of the above analysis, it is natural to write the Poisson bracket on Jacobi elements as

$$\{f(L), g(L)\} = -\kappa(L, [\nabla f - [N, \nabla f], \nabla g - [N, \nabla g]),$$

where ∇f is the gradient of f in \mathcal{G}_n defined by $df(L) \cdot \delta L = -\kappa(\nabla f, \delta L)$, $\delta L \in \mathcal{G}_n$.

This is equivalent to the bracket given in Symes [29].

The overall picture we have developed is as follows. The Toda flow is Hamiltonian in the Jacobi elements embedded in \mathcal{G}_u via (2.3). The level sets of the integrals of motion are Lagrangian submanifolds of the set of Jacobi elements. Also the level sets lie on orbits of \mathcal{G}_u . These level sets are invariant submanifolds for the gradient flow of the function $\kappa(L, N)$ with respect to the normal metric. On the level sets the Hamiltonian and gradient flows are identical for N as in Theorem 2.1.

Remark. We note also that one can compute quite explicitly the gradient flow with respect to the normal metric of an arbitrary function $\phi(A)$ on an orbit \mathcal{O} passing through $A_0 \in \mathcal{G}_{un} = \mathcal{G}_u \cap \mathcal{G}_n$. Let ∇ denote the gradient relative to the non-degenerate bilinear form $\langle \cdot, \cdot \rangle = -\kappa(\cdot, \cdot)$ on \mathcal{G}_n . Then the flow is given by

$$\dot{A} = [A, [A, -\nabla\phi]]. \quad (2.20)$$

The proof is essentially the same as that of Proposition 1.4.

Further, if $I: \mathcal{G}_n \rightarrow \mathbb{R}$ is an invariant function, i.e. $[\nabla I(A), A] = 0$, then the gradient flow of $\phi(A) = -\kappa(\nabla I(A), M)$, where $M = -iN$, is

$$\dot{A} = [A, [\nabla I(A), -M]]. \quad (2.21)$$

This follows from (2.20) together with the observation that taking the derivative

on \mathcal{G}_n at A in the direction of $M \in i\mathcal{A}$ of the relation $[\nabla I(A), A] = 0$ gives $[d(\nabla I)(A) \cdot M, A] + [\nabla I(A), M] = 0$.

3. Representations for the Toda Flows

In this section we consider representations for the Toda flow in the double bracket form for the classical Lie algebras. The key new ingredient here is of course the determination of the matrix N . For convenient bases for the classical Lie algebras, see, for example, Sattinger and Weaver [28], or Humphries [24].

A_I . We take here $h_1 = \text{diag}(1, -1, 0, 0, \dots)$, $h_2 = \text{diag}(0, 1, -1, 0, \dots)$. If E_{ij} is the matrix with 1 in the ij^{th} entry and zeros elsewhere, the positive simple roots are $e_{\alpha_1} = E_{12}$, $e_{\alpha_2} = E_{23} \dots$.

From $N = \sum_j ix_j h_j$, Table 1, and (2.3), we find;

$$N = i \text{diag} \left(-\frac{l}{2}, \frac{-l+2}{2}, \dots, \frac{l-2}{2}, \frac{l}{2} \right), \tag{3.1}$$

$$L = i \begin{bmatrix} b_1 & a_1 & & & & \\ a_1 & b_2 - b_1 & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & & a_l & \\ & & & & a_l & -b_l \end{bmatrix}. \tag{3.2}$$

C_l . Here we take $h_1 = \text{diag}(1, -1, -1, 1, 0, \dots, 0)$, $h_2 = (0, 0, 1, -1, -1, 1, \dots), \dots$, $h_l = (0, 0, \dots, 0, 0, 1, -1)$. Now, define the functionals α_i on the Cartan subalgebra by $\alpha_i(H) = \lambda_i$, where $H = \text{diag}(\lambda_1 a_0, \dots, \lambda_l a_0)$, a_0 as in (3.4) below. (Note that α_i here is not the same as that used in the definition of Chevalley basis.)

Calculating as before, we find

$$N = i \text{diag} \left(\frac{1}{2}(1-2l), -\frac{1}{2}(1-2l), \frac{1}{2}(3-2l), -\frac{1}{2}(3-2l), \dots, -\frac{3}{2}, \frac{3}{2}, -\frac{1}{2}, \frac{1}{2} \right) \tag{3.3}$$

We now calculate L explicitly. Following [28], define

$$a_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad a_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad a_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.4}$$

The simple positive roots are $\alpha_1 - \alpha_2, \dots, \alpha_{l-1} - \alpha_l, 2\alpha_l$ with corresponding root vectors

$$\begin{aligned} & \alpha_i - \alpha_j: \begin{pmatrix} \mathbf{1} + a_0 \\ 2 \end{pmatrix}_{ij} - \begin{pmatrix} \mathbf{1} - a_0 \\ 2 \end{pmatrix}_{ji}, \quad i < j \\ & \qquad \qquad \qquad 2\alpha_l: (a_+)_l \qquad \qquad \qquad \text{and} \\ & -(\alpha_i - \alpha_j): \begin{pmatrix} \mathbf{1} - a_0 \\ 2 \end{pmatrix}_{ij} - \begin{pmatrix} \mathbf{1} + a_0 \\ 2 \end{pmatrix}_{ji} \\ & \qquad \qquad \qquad -2\alpha_l: (a_-)_l \end{aligned}$$

with the double subscript indicating the position of the (i, j) 2×2 block. $\mathbf{1}$ is the 2×2 identity matrix.

Then for C_3 we have

$$L = i \begin{bmatrix} b_1 & 0 & a_1 & 0 & 0 & 0 \\ 0 & -b_1 & 0 & a_1 & 0 & 0 \\ -a_1 & 0 & b_2 - b_1 & 0 & a_2 & 0 \\ 0 & -a_1 & 0 & b_1 - b_2 & 0 & a_2 \\ 0 & 0 & -a_2 & 0 & b_2 - b_3 & a_3 \\ 0 & 0 & 0 & -a_2 & a_3 & b_2 - b_3 \end{bmatrix} \quad (3.5)$$

D_l . Here the h_i are as for C_l

From Table 1

$$N = i \operatorname{diag}(-l+1, l-1, -l+2, 2-l, \dots, -1, 1, 0). \quad (3.6)$$

We again calculate L .

Defining the α_i as for C_l , the simple positive roots are $\alpha_1 - \alpha_2, \alpha_2 - \alpha_3, \alpha_{l-1} - \alpha_l$, and $\alpha_{l-1} + \alpha_l$ with corresponding root vectors

$$\begin{aligned} \alpha_i - \alpha_j: & \left(\frac{\mathbf{1} + a_0}{2} \right)_{ij} + \left(\frac{a_0 - \mathbf{1}}{2} \right)_{ji} \quad i \neq j \\ \alpha_{l-1} + \alpha_l: & (a_+)_{l-1} - (a_+)_{l-1l} \\ -(\alpha_{l-1} + \alpha_l): & (a_-)_{l-1l} - (a_-)_{l-1}. \end{aligned}$$

Then for D_3 we have

$$L = i \begin{bmatrix} b_1 & 0 & a_1 & 0 & 0 & 0 \\ 0 & -b_1 & 0 & -a_1 & 0 & 0 \\ a_1 & 0 & b_2 - b_1 & 0 & a_2 & -a_3 \\ 0 & -a_1 & 0 & -b_1 + b_2 & a_3 & -a_2 \\ 0 & 0 & a_2 & a_3 & b_3 - b_2 & 0 \\ 0 & 0 & -a_3 & -a_2 & 0 & b_2 - b_3 \end{bmatrix}. \quad (3.7)$$

B_l . Here we take $h_1 = \operatorname{diag}(0, 1, -1, 0, \dots)$, $h_2 = \operatorname{diag}(0, 0, 0, 1, -1, \dots)$, etc. Then

$$N = i \operatorname{diag}(0, -l, l, -l+1, 1-l, \dots, -1, 1). \quad (3.8)$$

We now calculate L .

Defining α_i here by $\alpha_i(H) = \lambda_i$, where $H = \operatorname{diag}(0, \lambda_1 a_0, \dots, \lambda_l a_0)$, the simple positive roots are

$$\alpha_1 - \alpha_2, \dots, \alpha_{l-1} - \alpha_l, \alpha_l$$

with root vectors:

$$\begin{aligned} \alpha_i - \alpha_j: & \left(\frac{\mathbf{1} + a_0}{2} \right)_{ij} - \left(\frac{\mathbf{1} - a_0}{2} \right)_{ji} \quad i \neq j \\ \alpha_l: & E_{0,2l} - E_{2l-1,0} \\ -\alpha_l: & E_{2l,0} - E_{0,2l-1}. \end{aligned}$$

Thus for B_3

$$L = i \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -a_3 & a_3 \\ 0 & b_1 & 0 & a_1 & 0 & 0 & 0 \\ 0 & 0 & -b_1 & 0 & -a_1 & 0 & 0 \\ 0 & a_1 & 0 & -b_1 + b_2 & 0 & a_2 & 0 \\ 0 & 0 & -a_1 & 0 & b_1 - b_2 & 0 & -a_2 \\ -a_3 & 0 & 0 & a_2 & 0 & -b_2 + 2b_3 & 0 \\ a_3 & 0 & 0 & 0 & -a_2 & 0 & b_2 - 2b_3 \end{bmatrix} \quad (3.9)$$

4. The Double Bracket Equation, Momentum Maps, and Toda Flows

In this section we discuss how the theory of convexity of the image of the momentum maps, as developed by Schur–Horn [19], Kostant [21], Atiyah [3] and Guillemin and Sternberg [17], may be used to deduce properties of the double bracket equation and hence the Toda lattice flows. We also discuss Lie-algebraic sorting.

As observed by Brockett in [10], one can in fact view the equation $\dot{L}(t) = [L(t), [L(t), N]]$ in the Hermitian case as solving a linear programming problem on the convex set defined by the Schur–Horn theorem. This also makes contact with the problem of minimizing the Total Least Squares function (see [4 and 11]).

Theorem 4.1. (i) *Let \mathcal{O} be a (co)adjoint orbit of G_u in \mathcal{G}_u and consider the gradient vector field $\dot{L}(t) = [L(t), [L(t), N]]$, $L(0) \in \mathcal{O}$, with flow F_t for N a fixed regular element in \mathcal{G}_u . The set of equilibria of this vector field coincides with $\mathcal{O} \cap \mathcal{A}$, where \mathcal{A} is the Cartan subalgebra of \mathcal{G}_u containing N . This set $\mathcal{O} \cap \mathcal{A}$ consists of a single Weyl group orbit. By Kostant’s convexity theorem ([21]), the convex hull of these equilibria is a compact polytope \mathcal{P} which is the image of \mathcal{O} under the momentum map $\pi: \mathcal{O} \rightarrow \mathcal{A}$ defined by the adjoint T -action on \mathcal{O} , where T is the maximal torus in \mathcal{G}_u obtained by exponentiating \mathcal{A} and the momentum map π is the orthogonal projection of \mathcal{O} onto \mathcal{A} . $\pi(F_t(\mathcal{O}))$ lies entirely in \mathcal{P} . The number of equilibria of F_t equals the order of the Weyl group for $L(0)$ regular.*

(ii) *If N and L are chosen as in Theorem 1(ii), $\dot{L}(t) = [L(t), [L(t), N]]$ becomes the Toda lattice equations defined on the level set of all constants of the motion and the projection of the Toda flow on \mathcal{A} lies in the interior of the polytope given in (i).*

Proof. T acts on \mathcal{O} by the (co)adjoint action, defining the momentum map $\pi: \mathcal{O} \rightarrow \mathcal{A} = \mathbb{R}^l$ which is the orthogonal projection of \mathcal{O} onto \mathcal{A} (relative to the metric $-\kappa(\cdot, \cdot)$ on \mathcal{G}_u). The image of π is a convex compact polytope which is the convex hull of the critical values of π . These critical values coincide with the Weyl group orbit $\mathcal{O} \cap \mathcal{A}$ (see [21]).

To prove the theorem we will show that the equilibria of the vector field $\dot{L}(t) = [L(t), [L(t), N]]$ on \mathcal{G}_u necessarily lie in \mathcal{A} . (See [10] for the unitary case.) Note firstly that the solutions of the equation $\dot{L}(t) = [L(t), [L(t), N]]$ exists for all time since the flow is the projection of a flow from a compact group. We will now

show that $\lim_{t \rightarrow \infty} L(t)$ must be in \mathcal{A} . Observe that

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

Hence $\kappa(L, N)$ is monotonic increasing along the flow and is bounded above since $L(t)$ lies in the coadjoint orbit of a compact group. Thus $\kappa(L, N)$ has a limit and its derivative goes to zero. But we can see that its derivative vanishes only if L and N commute and, since N is regular, L must lie in \mathcal{A} . \square

We remark that the Toda lattice equations on the level set of all constants of the motion have no equilibria, but if one considers these equations defined on the closure of the Toda phase space, then the equilibria are given by the Weyl group orbit as stated in Theorem 4.1.

We remark that in [14 and 33], it was shown that an isospectral manifold of Jacobi (symmetric, tridiagonal) matrices with the off diagonal elements taken to be nonzero – i.e. a Toda orbit – is homeomorphic to a convex polytope. In a related paper [6] we show, by studying the Kähler geometry of the Toda flows, that this polytope is in fact the image of a momentum map.

In this paper we restrict ourselves to observing that while the image of the momentum map discussed in Theorem 4.1(i) is indeed convex, when one restricts the domain to the Toda orbit, one obtains an image that lies in the interior of a convex polytope, with equilibria at its vertices, but the image is far from convex. This is illustrated for A_2 below, where the projection onto the Cartan subalgebra of the integral curves of the system with conserved eigenvalues $\lambda_1 = 1$, $\lambda_2 = 2$, and $\lambda_3 = 3$, are shown. (This image has also been constructed by H. Flaschka and M. Zou.)

It is well known (see Symes [31] Deift, Nanda and Tomei [14]) that for the Toda lattice equations on $sl(n)$ the only stable equilibrium is that where the eigenvalues of L are sorted into increasing order along the diagonal of L . We can prove the Lie algebraic generalization of this result using our double bracket

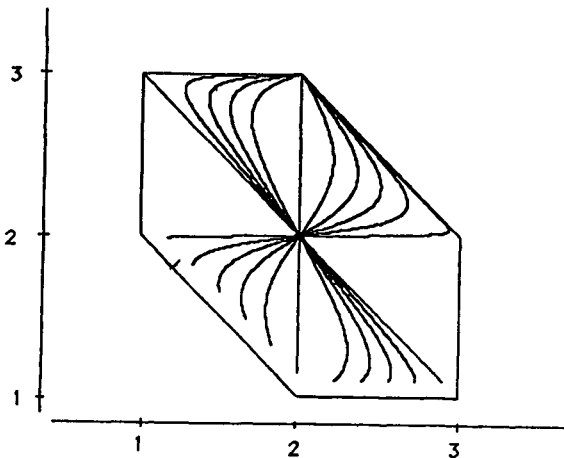


Fig. 1. The projection of the Toda trajectories under the image of the momentum map for A_2

formulation, but also a more general result on sorting which applies to the (non-Toda) double bracket flow on a compact Lie algebra. This result states essentially that the only stable equilibrium of the double bracket flow is that where L is diagonal and the diagonal entries are the eigenvalues of L sorted commensurately with N . More precisely, our result (which is a Lie algebraic generalization of the Hermitian result in [10]) is

Theorem 4.2. *Consider, as in Theorem 4.1, the (co)adjoint orbit \mathcal{O} of G_u in \mathcal{G}_u and the vector field $\dot{L}(t) = [L(t), [L(t), N]]$ on \mathcal{O} . Assume N is regular. The only stable equilibrium, L_0 , of this vector field, lying necessarily in $\mathcal{O} \cap \mathcal{A}$, where \mathcal{A} is the Cartan subalgebra of \mathcal{G}_u containing N , is the (unique) equilibrium which lies in the same Weyl chamber as $-N$.*

Proof. We prove the result by examining the linearization of the vector field $\dot{L}(t) = [L(t), [L(t), N]]$ at L_0 .

Taking the differential of the vector field at $L = L_0$ gives

$$\begin{aligned} \delta([L, [L, N]]) &= [\delta L, [L_0, N]] + [L_0, [\delta L, N]] \\ &= [L_0, [\delta L, N]], \end{aligned} \quad (4.1)$$

since L_0 and N commute, L_0 being an equilibrium point.

Hence the linearized vector field is given by

$$\delta \dot{L} = [L_0, [\delta L, N]]. \quad (4.2)$$

As in Sect. 2, let $\{h_j, e_\alpha | j=1, \dots, l, \alpha \in \Delta\}$ be a Chevalley basis of \mathcal{G} , the complexification of the Lie algebra \mathcal{G}_u . \mathcal{A} is a maximal abelian subalgebra of \mathcal{G}_u , $\mathcal{H} = \mathcal{A} \oplus i\mathcal{A}$ is the Cartan subalgebra of \mathcal{G} and $h_j \in \mathcal{A}$ for $j=1, \dots, l$. $\alpha_1, \dots, \alpha_l$ denote the simple roots. Then a basis for \mathcal{G}_u is $\{ih_j, x_\alpha := e_\alpha - e_{-\alpha}, y_\alpha := i(e_\alpha + e_{-\alpha}) | j=1, \dots, l, \alpha \text{ a positive root}\}$. Let $(,)$ denote the inner product induced by the Killing form on roots and let $\langle \alpha, \beta \rangle = 2(\alpha, \beta)/(\beta, \beta) \in \mathbb{Z}$. We have

$$[ih_j, x_\alpha] = \langle \alpha, \alpha_j \rangle y_\alpha \quad \text{and} \quad [ih_j, y_\alpha] = -\langle \alpha, \alpha_j \rangle x_\alpha. \quad (4.3)$$

Therefore, if

$$\begin{aligned} N &= \sum_{j=1}^l x_j ih_j, \quad L_0 = \sum_{k=1}^l y_k ih_k \\ \delta L &= \sum_{p=1}^l b_p ih_p + \sum_{\alpha > 0} (c_\alpha x_\alpha + d_\alpha y_\alpha) \end{aligned}$$

for $x_j, y_k, b_p, c_\alpha, d_\alpha \in \mathbb{R}$, we get

$$\begin{aligned} [\delta L, N] &= -[N, \delta L] = -\sum_{\alpha > 0, j} (x_j c_\alpha [ih_j, x_\alpha] + x_j d_\alpha [ih_j, y_\alpha]) \\ &= -\sum_{\alpha > 0, j} (x_j c_\alpha \langle \alpha, \alpha_j \rangle y_\alpha - x_j d_\alpha \langle \alpha, \alpha_j \rangle x_\alpha), \end{aligned} \quad (4.4)$$

and hence

$$\begin{aligned} [L_0, [\delta L, N]] &= -\left[\sum_{k=1}^l y_k ih_k, \sum_{\alpha > 0} \sum_{j=1}^l (x_j c_\alpha \langle \alpha, \alpha_j \rangle y_\alpha - x_j d_\alpha \langle \alpha, \alpha_j \rangle x_\alpha) \right] \\ &= -\sum_{\substack{\alpha > 0 \\ j, k}} y_k x_j c_\alpha \langle \alpha, \alpha_j \rangle [ih_k, y_\alpha] + \sum_{\substack{\alpha > 0 \\ j, k}} y_k x_j d_\alpha \langle \alpha, \alpha_j \rangle [ih_k, x_\alpha] \end{aligned}$$

$$\begin{aligned}
&= \sum_{\substack{\alpha > 0 \\ j,k}} (y_k x_j c_\alpha \langle \alpha, \alpha_j \rangle \langle \alpha, \alpha_k \rangle x_\alpha + y_k x_j d_\alpha \langle \alpha, \alpha_j \rangle \langle \alpha, \alpha_k \rangle y_\alpha) \\
&= \sum_{\alpha > 0} \left\langle \alpha, \sum_{j=1}^l x_j \alpha_j \right\rangle \left\langle \alpha, \sum_{k=1}^l y_k \alpha_k \right\rangle (c_\alpha x_\alpha + d_\alpha y_\alpha). \tag{4.5}
\end{aligned}$$

Therefore, the linearized system (4.2) decouples into

$$\begin{cases} \dot{b}_p = 0, & p = 1, \dots, l \\ \dot{c}_\alpha = c_\alpha \left\langle \alpha, \sum_{j=1}^l x_j \alpha_j \right\rangle \left\langle \alpha, \sum_{k=1}^l y_k \alpha_k \right\rangle, & \alpha > 0 \\ \dot{d}_\alpha = d_\alpha \left\langle \alpha, \sum_{j=1}^l x_j \alpha_j \right\rangle \left\langle \alpha, \sum_{k=1}^l y_k \alpha_k \right\rangle, & \alpha > 0. \end{cases} \tag{4.6}$$

Now, since the Weyl group acts simply transitively on the Weyl chambers, we may assume that N lies in the positive Weyl chamber, i.e., $\left\langle \alpha, \sum_{j=1}^l x_j \alpha_j \right\rangle > 0$. Therefore, the origin will be a stable equilibrium for the decoupled linear system (4.6) iff $\left\langle \alpha, \sum_{k=1}^l y_k \alpha_k \right\rangle \leq 0$, i.e. iff L_0 lies in the reflection through the origin of the positive Weyl chamber, that is, the Weyl chamber of $-N$. \square

Remarks. 1) If N is not regular, the conclusion of the theorem does not hold any more. Indeed, if for some $\alpha > 0$, $\left\langle \alpha, \sum_{j=1}^l x_j \alpha_j \right\rangle = 0$, i.e., N lies on a wall, then in (4.6) the sign of $\left\langle \alpha, \sum_{k=1}^l y_k \alpha_k \right\rangle$ cannot be controlled.

2) If $\mathcal{G}_u = su(l+1)$, the off-diagonal part of (4.6) reduces to

$$\delta \dot{L}_{ij} = (\lambda_{\pi(j)} - \lambda_{\pi(i)})(n_j - n_i) \delta L_{ij},$$

where $N = \text{diag}(in_1, \dots, in_{l+1})$, $L_0 = \text{diag}(i\lambda_{\pi(1)}, \dots, i\lambda_{\pi(l+1)})$ are traceless diagonal matrices with purely imaginary entries, and π is a permutation of $\{1, 2, \dots, l+1\}$. In other words, in the limit the flow of $\dot{L}(t) = [L(t), [L(t), N]]$ sorts the eigenvalues of L such that they appear on the imaginary axis in opposite order to the eigenvalues of N . This result in the Hermitian context was obtained by Brockett [10]; there the real eigenvalues of N and the limit of L appear in the same order on the real axis.

Requiring all the coefficients in (4.6) to be positive characterizes the unique equilibrium which is a source. We get:

Corollary 4.3. *In the hypotheses of Theorem 4.2, the unique equilibrium of $\dot{L}(t) = [L(t), [L(t), N]]$ which is a source lies in the same Weyl chamber as N .*

Both Theorem 4.2 and Corollary 4.3 generalize to the following.

Theorem 4.4. *In the hypotheses and notations of Theorem 4.2, let L_s be the unique equilibrium which is a source and let $\sigma(L_s)$ be another equilibrium, where σ is an element in the Weyl group. Then the dimension of the stable manifold of $\sigma(L_s)$ equals twice the length of σ .*

Proof. From the system (4.6), we see that the dimension of the stable subspace of the linearized system at the equilibrium $\sigma(L_s) = \sum_{k=1}^l y_k i h_k$, equals twice the number of $\alpha > 0$ such that $\left(\alpha, \sum_{k=1}^l y_k \alpha_k\right) < 0$. However, $\left(\alpha, \sum_{k=1}^l y_k \alpha_k\right) = \left(\sigma\alpha, \sigma\left(\sum_{k=1}^l y_k \alpha_k\right)\right)$ and $\sigma\left(\sum_{k=1}^l y_k \alpha_k\right)$ is in the positive Weyl chamber by construction and choice of ordering. Since the only roots which have positive inner product with an element in the positive Weyl chamber are the positive roots, we need to determine all positive roots α for which $\sigma\alpha < 0$. This number, however, is the length of σ (see Humphreys [24] p. 52). \square

We remark that Theorem 4.4 can also be proved by examining the gradient flow $\kappa(L, N)$ with respect to the Kähler metric on \mathcal{O} . This is done in Atiyah [3], where the dimension of the stable manifold is shown to coincide with the dimension of Bruhat cell association with the equilibrium. This result is equivalent to ours, since the dimensions of the stable and unstable manifolds at the equilibria of a gradient vector field are independent of the metric used. This may be seen as follows. The linearization of the vector field ∇f at a given equilibrium is $g^{-1}H$ where H is the Hessian of f at the equilibrium and g is the local matrix representation of the metric. The eigenvalues of $g^{-1}H$ are equal to those of $g^{-1/2}Hg^{-1/2} = (g^{-1/2})^T H g^{-1/2}$, which is a quadratic form having the same rank and signature as H .

We can thus see that the gradient flow (1.3) has a number of remarkable asymptotic properties. When restricted to Jacobi elements, we have seen these equations are actually Hamiltonian. Although there are other invariant subspaces for these equations, as we have shown, we know of no other cases where they are Hamiltonian. We note however that a key point which allowed us to show the nonperiodic Toda lattice equations are gradient on their isospectral set is the noncompactness of the level sets. We suspect that other integrable systems with noncompact level sets may exhibit gradient-like behavior or indeed be gradient flow on their level sets. We are currently investigating this idea.

Acknowledgements. We would like to thank H. Flaschka for his valuable comments during the course of this work and J. E. Marsden for several useful conversations. We would also like to thank J. Lagarias for some illuminating discussions on gradient flows and integrable systems, and the referee, whose suggestions greatly improved the exposition.

References

1. Abraham, R. A., Marsden, J. E.: Foundations of mechanics. Benjamin Cummings. 1978
2. Adler, M.: On a trace functional for pseudo-differential operators and the symplectic structure of Korteweg–de Vries type equations. Invent. Math. **50**, 219–248 (1979)
3. Atiyah, M. F.: Convexity and commuting Hamiltonians. Bull. Lond. Math. Soc. **16**, 1–15 (1982)
4. Bloch, A. M.: Estimation, principal components and Hamiltonian systems. Syst. Control Letters **6**, 103–108 (1985)
5. Bloch, A. M.: Steepest descent, linear programming and Hamiltonian flows. Contemp. Math. AMS **114**, 77–88 (1990)
6. Bloch, A. M., Flaschka, H., Ratiu, T.: A convexity theorem for isospectral sets of Jacobi matrices in a compact Lie algebra. Duke Math. J. **61** (1), 41–66 (1990)

7. Bloch, A. M., Brockett, R. W., Ratiu, T.: A new formulation of the generalized Toda lattice equations and their fixed point analysis via the moment map, *Bull. AMS* **23** (2), 447–456 (1990)
8. Bremner, M. R., Moody, R. V., Patera, J.: Tables of dominant weight multiplicities for representations of simple Lie algebras. Marcel Dekker
9. Brockett, R. W.: Least squares matching problems. *Linear algebra and its applications*, **122/123/124**, 761–777 (1989)
10. Brockett, R. W.: Dynamical systems that sort lists and solve linear programming problems. In: *Proc. 27th IEEE Conf. on decision and control*, IEEE, 1988, pp. 799–803, subsequently published in *Linear Algebra and its Applications* **146**, 79–91 (1991)
11. Byrnes, C. I., Willems, J. C.: Least squares estimation, linear programming and momentum maps. *IMA J*
12. Chu, M. T.: The generalized Toda lattice, the QR-algorithm and the centre manifold theory. *SIAM J. Alg. Disc. Math.* **5**, 187–201 (1984)
13. Davis, M. S.: Some aspherical manifolds. *Duke Math. J.* **5**, 105–139 (1987)
14. Deift, P., Nanda, T., Tomei, C.: Differential equations for the symmetric eigenvalue problem. *SIAM J. Numerical Anal.* **20**, 1–22 (1983)
15. Flaschka, H.: The Toda lattice. *Phys. Rev. B* **9**, 1924–1925 (1976)
16. Fried, D.: The cohomology of an isospectral flow. *Proc. AMS* **98**, 363–368 (1986)
17. Guillemin, V., Sternberg, S.: Convexity properties of the moment mapping. *Invent. Math.* **67**, 491–513 (1982)
18. Guillemin, V., Sternberg, S.: On the method of Symes for integrating systems of the Toda type. *Lett. Math. Phys.* **7**, 113–115 (1983)
19. Horn, A.: Doubly stochastic matrices and the diagonal of a rotation matrix. *Amer. J. Math.* **76**, 620–630 (1956)
20. Kostant, B.: The principal three-dimensional subgroup and the Betti numbers of a complex simple Lie group. *Am. J. Math.* **81**, 973–1032 (1959)
21. Kostant, B.: On Convexity, the Weyl group and the Iwasawa decomposition, *Ann. Sci. Ec. Norm. Sup.* **6**, 413–455 (1973)
22. Kostant, B.: The solution to a generalized Toda lattice and representation theory. *Adv. Math.* **34**, 195–338 (1979)
23. Kostant, B.: Quantization and representation theory. In: *Representation theory of Lie groups*, London mathematical society. Lecture note series **34**, pp. 91–150. Cambridge University Press 1979
24. Humphries, J. E.: *Introduction to Lie algebras and representation theory*. Berlin, Heidelberg, New York: Springer 1972
25. Lagarias, J.: Monotonicity properties of the generalized Toda flow and QR flow. *SIAM J. Matrix Anal. App.* (to appear)
26. MacDonald, I. G.: *Algebraic structure of Lie groups*, in representation theory of Lie groups, London mathematical society. Lecture note series **34**, pp 91–150. Cambridge University Press 1979
27. Moser, J.: Finitely many mass points on the line under the influence of an exponential potential. *Batelles Recontres*. Springer notes in physics, pp. 417–497. Berlin, Heidelberg, New York: Springer 1974
28. Sattinger, D. H., Weaver, D. L.: *Lie groups and Lie algebras with applications to physics, geometry and mechanics*. Berlin, Heidelberg, New York: Springer 1986
29. Symes, W. W.: Hamiltonian group actions and integrable systems. *Physica D* **1**, 339–376 (1980)
30. Symes, W. W.: Systems of Toda type, inverse spectral problems and representation theory. *Invent. Math.* **59**, 13–51 (1982)
31. Symes, W. W.: The QR algorithm and scattering for the nonperiodic Toda lattice. *Physica D* **4**, 275–280 (1982)
32. Toda, M.: Studies of a non-linear lattice. *Phys. Rep. Phys. Lett.* **8**, 1–125 (1975)
33. Tomei, C.: The topology of isospectral manifolds of tridiagonal matrices. *Duke Math. J.* **51**, 981–996 (1984)
34. van Moerbeke, P.: The spectrum of Jacobi matrices. *Invent. Math.* **37**, 45–81 (1976)