Smooth and Discrete Integrable Systems and Optimal Control

In honor of Darryl Holm's 60th Birthday

Anthony Bloch

Work with Peter Crouch, Arieh Iserles, Jerry Marsden, Tudor Ratiu and Amit Sanyal and Darryl Holm

- Toda
- Symmetric rigid body equations – smooth and discrete
- Flows on Stiefel manfolds – Jacobi flow on ellipsoid
- Flows on Quadratic groups

- Fluid Flows.
- Symmetric/Symplectic Flows and their Lie Poisson Structure
- Optimal Control.

Rigid Body Equations:

$$\dot{M} = [M, \Omega], \qquad M = \Lambda\Omega + \Omega\Lambda$$

Symmetric Rigid Body Equations:

$$\dot{Q} = Q\Omega \qquad \dot{P} = P\Omega$$

Toda Flow:

$$\dot{X} = [X, \Pi_S X]$$

Double Bracket Flow:

$$\dot{X} = [X, [X, N]]$$

– gradient but special case yields Toda.
(See B, Brockett and Ratiu)
Generalized Double Bracket Flow:

$$\dot{X} = [X, [X, G(X - N)]]$$

– in particular $G(X - N) = (X - N)^k$.
(See B and Iserles)
Infinite (dispersionless) flow:

$$\dot{x} = \{x, \{x, n\}, \ x = x(z, \theta).$$

(B. Brockett, Flashcka and Ratiu)

Double Double Bracket Flow: Geodesic Flows on Grassmannians:

$$\dot{X} = [X, [X, P] \ \dot{P} = [P, [X, P]].$$

Flow on the symmetric matrices/symplectic groups
(B, Brockett and Crouch)

$$\dot{X} = [X^2, N] = [X, [XN + NX]$$

(B, Brinznescu,Iserles, Marsden, Ratiu)

Matrix form of nonperiodic tridiagonal Toda:

$$\frac{d}{dt}L = [B, L] = BL - LB, \tag{0.1}$$

$$L = \begin{pmatrix} b_1 & a_1 & 0 & \cdots & 0 \\ a_1 & b_2 & a_2 & \cdots & 0 \\ & & \ddots & & \\ & & & b_{n-1} & a_{n-1} \\ 0 & & & a_{n-1} & b_n \end{pmatrix}$$

$$B = \begin{pmatrix} 0 & a_1 & 0 & \cdots & 0 \\ -a_1 & 0 & a_2 & \cdots & 0 \\ & & \ddots & & \\ & & & 0 & a_{n-1} \\ 0 & & & -a_{n-1} & 0 \end{pmatrix}$$

The double bracket flow is a gradient flow on an adjoint orbit $\mathcal{O}$ endowed with the "standard" or "normal" metric:

• When the matrix L in the double bracket flow is tridiagonal and the matrix N is the diagonal matrix $\mathrm{diag}(1,2,\ldots n)$, the double bracket flow is both gradient and Hamiltonian on a level set of its integrals – the Toda lattice flow.

• Level set noncompact and diffeomorphic to a product of lines, unlike many Hamiltonian systems where the level set of the integrals is diffeomorphic to a torus.

Flow can be mapped into interior of Schur-Horn polytope, equilibria at the vertices.

Early key work on this: Moser, Symes, Deift, Nanda and Tomei.

Related flows: full Toda flows:

$$\dot{L} = [L, \pi_S L]$$

.

See work of Deift, Li, Nanda, Tomei; Ercolani, Flaschka, Singer.

PDE on Diff(Annulus)

$$\dot{x} = \{x, \{x, n\}\}$$

Special case: $x(z, \theta) = u(z) + 2v(z)\cos\theta$.
Tridiagonal dispersionless Toda:

$$u_t = 4vv_z \; v_t = vu_z$$

.

Rearrangments of functions – infinite dimensional Schur-Horn.

More work on double brackets: recent work of Holm et. al.

1 The n-dimensional Rigid Body.

• Here review the classical rigid body equations in in n dimensions.

Use the following pairing on $\mathfrak{so}(n)$, the Lie algebra of the n-dimensional proper rotation group $\mathrm{SO}(n)$:

$$\langle \xi, \eta \rangle = -\frac{1}{2}\operatorname{trace}(\xi\eta).$$

Use this inner product to identify $\mathfrak{so}(n)^* \; \mathfrak{so}(n)$.

• Recall from Manakov [1976] and Ratiu [1980] that the left invariant generalized rigid body equations on $\mathrm{SO}(n)$ may be written as

$$\begin{aligned} \dot{Q} &= Q\Omega \\ \dot{M} &= [M, \Omega]\,, \end{aligned} \tag{RBn}$$

where $Q \in \mathrm{SO}(n)$ denotes the configuration space variable (the attitude of the body), $\Omega = Q^{-1}\dot{Q} \in \mathfrak{so}(n)$ is the body angular velocity, and the body angular momentum is

$$M := J(\Omega) = \Lambda\Omega + \Omega\Lambda \in \mathfrak{so}(n)\,.$$

- Here $J : \mathfrak{so}(n) \to \mathfrak{so}(n)$ is the symmetric pos def operator defined by

$$J(\Omega) = \Lambda\Omega + \Omega\Lambda,$$

where Λ is a diagonal matrix sat $\Lambda_i + \Lambda_j > 0$ for all $i \neq j$.

There is a similar formalism for any semisimple Lie group.

Right Invariant System. The system (RBn) has a right invariant counterpart. This right invariant system is given as follows:

$$\dot{Q}_r = \Omega_r Q_r; \qquad \dot{M}_r = [\Omega_r, M_r] \tag{RightRBn}$$

where in this case $\Omega_r = \dot{Q}_r Q_r^{-1}$ and $M_r = J(\Omega_r)$ where J has the same form as above.

Relating the Left and the Right Rigid Body Systems.

Proposition 1.1. *If $(Q(t), M(t))$ satisfies (RBn) then the pair $(Q_r(t), M_r(t))$, where $Q_r(t) = Q(t)^T$ and $M_r(t) = -M(t)$ satisfies (RightRBn). There is a similar converse statement.*

2 The Symmetric Rigid Body Equations.

The System (SRBn). By definition, ***the left invariant symmetric rigid body system*** (SRBn) is given by the first order equations

$$\begin{aligned} \dot{Q} &= Q\Omega \\ \dot{P} &= P\Omega \end{aligned} \tag{SRBn}$$

where Ω is regarded as a function of Q and P via the equations

$$\Omega := J^{-1}(M) \in \mathfrak{so}(n) \quad \text{and} \quad M := Q^T P - P^T Q.$$

Proposition 2.1. *If (Q, P) is a solution of (SRBn), then (Q, M) where $M = J(\Omega)$ and $\Omega = Q^{-1}\dot{Q}$ satisfies the rigid body equations (RBn).*

Proof. Differentiating $M = Q^T P - P^T Q$ and using the equations (SRBn) gives the second of the equations (RBn). ■

• Local Equivalence of the Rigid Body and the Symmetric Rigid Body Equations.

Above saw that solutions of the symmetric rigid body system can be mapped to solutions of the rigid body system. Now consider the converse question:

Suppose have a solution (Q, M) of the standard left invariant rigid body equations. Sseek to solve for P in

$$M = Q^T P - P^T Q. \tag{2.1}$$

Definition 2.2. *Let C denote the set of (Q, P) that map to M's with operator norm equal to 2 and let S denote the set of (Q, P) that map to M's with operator norm strictly less than 2. Also denote by S_M the set of points $(Q, M) \in T^* \mathrm{SO}(n)$ with $\|M\|_{\mathrm{op}} \leq 2$.*

Proposition 2.3. *For $\|M\|_{\mathrm{op}} < 2$, the equation(2.1) has the solution*

$$P = Q \left(e^{\sinh^{-1} M/2} \right) \tag{2.2}$$

The System (RightSRBn). By definition, the ***symmetric representation of the rigid body equations in right invariant form*** on $\mathrm{SO}(n) \times \mathrm{SO}(n)$ are given by the first order equations

$$\dot{Q}_r = \Omega_r Q_r; \qquad \dot{P}_r = \Omega_r P_r \qquad \text{(RightSRBn)}$$

where $\Omega_r := J^{-1}(M_r) \in \mathfrak{so}(n)$ and where $M_r = P_r Q_r^T - Q_r P_r^T$.

It is easy to check that that this system is right invariant on $\mathrm{SO}(n) \times \mathrm{SO}(n)$.

Proposition 2.4. *If (Q_r, P_r) is a solution of (RightSRBn), then (Q_r, M_r), where $M_r = J(\Omega_r)$ and $\Omega_r = \dot{Q}_r Q_r^{-1}$, satisfies the right rigid body equations (RightRBn).*

The Hamiltonian Form of (SRBn).

Recall that the classical rigid body equations are Hamiltonian on $T^* \mathrm{SO}(n)$ with respect to the canonical symplectic structure on the cotangent bundle of $\mathrm{SO}(n)$.

In symmetric case have:

Proposition 2.5. *Consider the Hamiltonian system on the symplectic vector space* $\mathfrak{gl}(n) \times \mathfrak{gl}(n)$ *with the symplectic structure*

$$\Omega_{\mathfrak{gl}(n)}(\xi_1, \eta_1, \xi_2, \eta_2) = \frac{1}{2} \operatorname{trace}(\eta_2^T \xi_1 - \eta_1^T \xi_2) \tag{2.3}$$

and Hamiltonian

$$H(\xi, \eta) = -\frac{1}{8} \operatorname{trace}\left[\left(J^{-1}(\xi^T \eta - \eta^T \xi)\right)\left(\xi^T \eta - \eta^T \xi\right)\right]. \tag{2.4}$$

The corresponding Hamiltonian system leaves $\mathrm{SO}(n) \times \mathrm{SO}(n)$ *invariant and induces on it, the symmetric rigid body flow.*

Note that the above Hamiltonian is equivalent to

$$H = \frac{1}{4} \left\langle J^{-1} M, M \right\rangle.$$

3 Optimal Control formulation of Rigid Body

Definition 3.1. *Let $T > 0$, $Q_0, Q_T \in \mathrm{SO}(n)$ be given and fixed. Let the rigid body optimal control problem be given by*

$$\min_{U \in \mathfrak{so}(n)} \frac{1}{4} \int_0^T \langle U, J(U) \rangle dt \tag{3.1}$$

subject to the constraint on U that there be a curve $Q(t) \in \mathrm{SO}(n)$ such that

$$\dot{Q} = QU \qquad Q(0) = Q_0, \qquad Q(T) = Q_T. \tag{3.2}$$

Proposition 3.2. *The rigid body optimal control problem (3.1) has optimal evolution equations (SRBn) where P is the costate vector given by the maximum principle.*

The optimal controls in this case are given by

$$U = J^{-1}(Q^T P - P^T Q). \tag{3.3}$$

The proof involves writing the Hamiltonian of the maximum principle as

$$H = \langle P, QU \rangle + \frac{1}{4} \langle U, J(U) \rangle . \tag{3.4}$$

Merging the Left and Right Problems.

Definition 3.3. *Let* $\mathfrak{u}(n)$ *denote the Lie algebra of the unitary group* $\mathrm{U}(n)$.

Let Q *be a* $p \times q$ *complex matrix and let* $U \in \mathfrak{u}(p)$ *and* $V \in \mathfrak{u}(q)$. *Let* J_U *and* J_V *be constant symmetric positive definite operators on the space of complex* $p \times p$ *and* $q \times q$ *matrices respectively and let* $\langle \cdot, \cdot \rangle$ *denote the trace inner product* $\langle A, B \rangle = \frac{1}{2}\operatorname{trace}(A^\dagger B)$, *where* $A^\dagger$ *is the adjoint; that is, the transpose conjugate.*

Let $T > 0$, Q_0, Q_T *be given and fixed. Define the optimal control problem over* $\mathfrak{u}(p) \times \mathfrak{u}(q)$

$$\min_{U,V} \frac{1}{4} \int \{\langle U, J_U U \rangle + \langle V, J_V V \rangle\} dt \tag{3.5}$$

subject to the constraint that there exists a curve $Q(t)$ *such that*

$$\dot{Q} = UQ - QV, \qquad Q(0) = Q_0, \qquad Q(T) = Q_T. \tag{3.6}$$

This problem was motivated by an optimal control problem on adjoint orbits of compact Lie groups as discussed by Brockett.

Theorem 3.4. *The optimal control problem 3.3 has optimal controls given by*

$$U = J_U^{-1}(PQ^\dagger - QP^\dagger); \qquad V = J_V^{-1}(P^\dagger Q - Q^\dagger P)\,. \tag{3.7}$$

and the optimal evolution of the states Q and costates P is given by

$$\begin{aligned} \dot{Q} &= J_U^{-1}(PQ^\dagger - QP^\dagger)Q - QJ_V^{-1}(P^\dagger Q - Q^\dagger P) \\ \dot{P} &= J_U^{-1}(PQ^\dagger - QP^\dagger)P - PJ_V^{-1}(P^\dagger Q - Q^\dagger P). \end{aligned} \tag{3.8}$$

Corollary 3.5. *The equations (3.8) are given by the coupled double bracket equations*

$$\dot{\hat{Q}} = [\hat{Q}, \hat{J}^{-1}[\hat{P}, \hat{Q}]]; \qquad \dot{\hat{P}} = [\hat{P}, \hat{J}^{-1}[\hat{P}, \hat{Q}]]\,. \tag{3.9}$$

where $\hat{J}$ is the operator $\mathrm{diag}(J_U, J_V)$,

$$\hat{Q} = \begin{bmatrix} 0 & Q \\ -Q^\dagger & 0 \end{bmatrix} \in \mathfrak{u}(p+q), \tag{3.10}$$

Q is a complex $p \times q$ matrix of full rank, $Q^\dagger$ is its adjoint, and similarly for P.

4 Discrete Variational Problems

This general method is closely related to the development of variational integrators for the integration of mechanical systems, as in Kane, Marsden, Ortiz and West [2000]. See also Iserles, McLachlan, and Zanna [1999] and Budd and Iserles [1999].

Key notion: ***discrete Lagrangian***, which is a map $L_d : Q \times Q \to \mathbb{R}$. The important point here is that the velocity phase space TQ of Lagrangian mechanics has been replaced by $Q \times Q$.

In the discrete setting, the action integral of Lagrangian mechanics is replaced by an action sum

$$S_d = \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}) \tag{4.1}$$

where $q_k \in Q$, the sum is over discrete time, and the equations are obtained by a discrete action principle which minimizes the discrete action given fixed endpoints q_0 and q_N.

Taking the extremum over $q_1, \cdots, q_{N-1}$ gives the ***discrete Euler-Lagrange equations***

$$D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}) = 0, \tag{4.2}$$

for $k = 1, \cdots, N-1$.

We can rewrite this as follows

$$D_2 L_d + D_1 L_d \circ \Phi = 0, \tag{4.3}$$

where $\Phi : Q \times Q \to Q \times Q$ is defined implicitly by $\Phi(q_{k-1}, q_k) = (q_k, q_{k+1})$.

5 Moser–Veselov Discretization

Recall now the Moser–Veselov [1991] discrete rigid body equations. This system will be called DRBn.

See also Deift, Li and Tomei [1992].

Discretize the configuration matrix and let $Q_k \in \mathrm{SO}(n)$ denote the rigid body configuration at time k, let $\Omega_k \in \mathrm{SO}(n)$ denote the discrete rigid body angular velocity at time k, let I denote the diagonal moment of inertia matrix, and let M_k denote the rigid body angular momentum at time k.

These quantities are related by the Moser-Veselov equations

$$\Omega_k = Q_k^T Q_{k-1} \tag{5.1}$$

$$M_k = \Omega_k^T \Lambda - \Lambda \Omega_k \tag{5.2}$$

$$M_{k+1} = \Omega_k M_k \Omega_k^T . \tag{5.3}$$

(DRBn)

The Moser-Veslov equations (5.1)-(5.3) can in fact be obtained by a discrete variational principle (see Moser and Veselov [1991]) of the form described above: one considers the stationary points of the functional

$$S = \sum_k \operatorname{trace}(Q_k I Q_{k+1}) \tag{5.4}$$

on sequences of orthogonal $n \times n$ matrices.

See also Marsden, Pekarsky and Shkoller [1999].

The Discrete Symmetric Rigid Body.

We now define the symmetric discrete rigid body equations as follows:

$$\begin{aligned} Q_{k+1} &= Q_k U_k \\ P_{k+1} &= P_k U_k \,, \end{aligned} \tag{SDRBn}$$

where U_k is defined by

$$U_k \Lambda - \Lambda U_k^T = Q_k^T P_k - P_k^T Q_k \,. \tag{5.5}$$

Using these equations, we have the algorithm $(Q_k, P_k) \mapsto (Q_{k+1}, P_{k+1})$ defined by: compute U_k from (5.5), compute Q_{k+1} and P_{k+1} using (SDRBn). We note that the update map for Q and P is done in parallel here.

Have:

Proposition 5.1. *The symmetric discrete rigid body equations (SDRBn) on S are equivalent to the Moser-Veselov equations (5.1)– (5.3) (DRBn) on the set S_M where S and S_M are defined in Proposition 2.2.*

Note that $m_k = P_k Q_k^T - Q_k P_k^T$ then $m_k = Q_k M_k Q_k^T$ and is conserved spatial momentum.

Discrete Optimal Control

Definition 5.2. *Let* Λ *be a positive definite diagonal matrix. Let* $\overline{Q}_0, \overline{Q}_N \in \mathrm{SO}(n)$ *be given and fixed. Let*

$$\hat{V} = \sum_{k=1}^{N} \mathrm{trace}(\Lambda U_k). \tag{5.6}$$

Define the optimal control problem

$$\min_{U_k} \hat{V} = \min_{U_k} \sum_{k=1}^{N} \mathrm{trace}(\Lambda U_k) \tag{5.7}$$

subject to dynamics and initial and final data

$$Q_{k+1} = Q_k U_k, \qquad Q_0 = \overline{Q}_0, \qquad Q_N = \overline{Q}_N \tag{5.8}$$

for $Q_k, U_k \in \mathrm{SO}(n)$.

Theorem 5.3. *A solution of the optimal control problem (5.2) satisfies the optimal evolution equations (SDRBn)*

$$Q_{k+1} = Q_k U_k; \qquad P_{k+1} = P_k U_k \,, \tag{5.9}$$

where P_k *is the discrete covector in the discrete maximum principle and* U_k *is defined by*

$$U_k \Lambda - \Lambda U_k^T = Q_k^T P_k - P_k^T Q_k \,. \tag{5.10}$$

The Symmetric Rigid Body Equations with Parameter

• Key observation: can write the generalized rigid body equations as Lax equations with parameter:

$$\frac{d}{dt}(M + \lambda\Lambda^2) = [M + \lambda\Lambda^2, \Omega + \lambda\Lambda], \tag{5.11}$$

Coefficients of λ in the traces of the powers of $M + \lambda\Lambda^2$ then yield the right number of independent integrals in involution to prove integrability of the flow on a generic adjoint orbit of $\mathrm{SO}(n)$ (identified with the corresponding coadjoint orbit).

• Moser and Veselov [1991] show that there is a corresponding formulation of the discrete rigid body equations with parameter.

Possible in fact also to write the full symmetric rigid body equations with parameter:

$$\begin{aligned} \dot{Q}_\lambda &= Q_\lambda(\Omega + \lambda\Lambda) \\ \dot{P}_\lambda &= P_\lambda(\Omega + \lambda\Lambda). \end{aligned} \tag{5.12}$$

6 Variational Problems on Stiefel Manifolds

Also give the extremal flows obtained in the limiting cases of the sphere/ellipsoid ($n = 1$), and the N dimensional rigid body ($n = N$). Extremal flows in these cases are well-known and integrable.

The Stiefel manifold $V(n, N) \subset \mathbb{R}^{nN}$ consists of orthogonal n frames in N dimensional real Euclidean space,

$$V(n, N) = \{Q \in \mathbb{R}^{nN}; \quad QQ^{\mathrm{T}} = I_n\}.$$

Introduce the pairing in $\mathbb{R}^{rs}$ given by

$$\langle A, B\rangle = \mathrm{Tr}(A^{\mathrm{T}}B), \tag{6.1}$$

where $\mathrm{Tr}(\cdot)$ denotes trace of a matrix and the left invariant metric on $\mathbb{R}^{nN}$ given by

$$\langle\langle W_1, W_2\rangle\rangle = \langle W_1\Lambda, W_2\rangle = \langle W_1, W_2\Lambda\rangle, \tag{6.2}$$

where Λ is a positive definite $N \times N$ diagonal matrix.

Consider the variational problem given by:

$$\min_{Q(\cdot)} \int_0^T \frac{1}{2}\langle\langle \dot{Q}, \dot{Q}\rangle\rangle dt \tag{6.3}$$

subject to: $QQ^{\mathrm{T}} = I_n$, $Q \in \mathbb{R}^{nN}$, $1 \leq n \leq N$, $Q(0) = Q_0$, $Q(T) = Q_T$, I_n denotes the $n \times n$ identity matrix. This is a variational problem defined on the Stiefel manifold $V(n, N)$. The dimension of this manifold is given by

$$\mathrm{Dim}\ V(n, N) = nN - \frac{n(n+1)}{2} = n(N-n) + \frac{n(n-1)}{2}.$$

Or:

$$\min_{U(\cdot)} \int_0^T \frac{1}{2}\langle\langle QU, QU\rangle\rangle dt \tag{6.4}$$

subject to: $\dot{Q} = QU$; $QQ^{\mathrm{T}} = I_n$, $Q(0) = Q_0$, $Q(T) = Q_T$ where $U \in \mathfrak{so}(N)$. Note that the quantity to be minimized is invariant with respect to the left action of $SO(n)$ on $V(n, N)$ since the metric (6.2) is left invariant.

The Rigid Body equations

For the special case when $n = N$, $V(N,N) \equiv SO(N)$ and the extremal trajectories of the optimal control problem (6.4) give the N-dimensional rigid body equations.

Geodesic flow on the ellipsoid

For the other extreme case, when $n = 1$, we obtain the equations for the geodesic flow on the sphere $V(1,N) \equiv \mathbb{S}^{N-1}$ with $Q = q^{\mathrm{T}}$, $q^{\mathrm{T}}q = 1$. This can be also be regarded as the geodesic flow on the ellipsoid

$$\bar{q}^{\mathrm{T}}\Lambda^{-1}\bar{q} = 1,$$

where $q = \Lambda^{-1/2}\bar{q}$. The costate variable $P = p^{\mathrm{T}}$ is used to enforce the constraint $\dot{q} = -Uq$ for the (6.4) when $n = 1$. The extremal solutions to this problem are

$$\dot{q} = -Uq, \quad \dot{p} = -Up + Aq, \tag{6.5}$$

where $A = qq^{\mathrm{T}}U\Lambda U - U\Lambda U qq^{\mathrm{T}}$.

The body momentum is obtained as

$$M = qp^{\mathrm{T}} - pq^{\mathrm{T}}, \tag{6.6}$$

in terms of the solution (q, p). Equations (6.5) can then be expressed in terms of the body momentum as

$$\dot{q} = -Uq, \quad \dot{M} = [M, U] - A. \tag{6.7}$$

The Lagrangian (variational) formulation for this problem gives us the equations for the geodesic flow on the sphere. To obtain these equations, we take reduced variations on $V(1, N) = \mathbb{S}^{N-1}$.

We get the Lagrangian (variational) equations for the geodesic flow on the sphere ($\mathbb{S}^{N-1}$) as

$$\ddot{q} = -\frac{\dot{q}^{\mathrm{T}}\dot{q}}{q^{\mathrm{T}}\Lambda^{-1}q}\Lambda^{-1}q. \tag{6.8}$$

Integrability of these extremal flows were proven by Jacobi with relation to Neumann problem of motion on sphere with quadratic potential, as shown by Knorrer (1982). Contemporary version of integrability of the geodesic flow on an ellipsoid was demonstrated by Moser (1980) using Theorem of Chasles and geometry of quadrics.

Obtain a symmetric form and discretization.

7 Quadratic Matrix Lie Groups

We consider quadratic matrix groups of the form

$$G := \left\{ g \in \mathbb{R}^{n\times n} \mid g^{\mathsf{T}} J g = J \right\}, \tag{7.1}$$

where g^{T} is the transpose of the $n \times n$ matrix g, $J^2 = \alpha I_n$ and $J^{\mathsf{T}} = \alpha J$ for $\alpha = \pm 1$.

This class of groups includes standard classical groups of interest including the symplectic group and $\mathrm{O}(p, q)$.

This class of matrix groups gives matrix representations of linear transformations on $\mathbb{R}^n$ that leave the following symmetric, bilinear form invariant:

$$f(x, y) = x^{\mathsf{T}} J y, \quad x, y \in \mathbb{R}^n.$$

Observation The Lie algebra of the group G is given by

$$\mathfrak{g} = \left\{ X \in \mathbb{R}^{n\times n} \mid X^{\mathsf{T}} J + J X = 0 \right\}.$$

If $g \in G$ then $g^{\mathsf{T}} \in G$ and $g - g^{-1} \in \mathfrak{g}$.

Let $\Sigma : \mathfrak{g} \to \mathfrak{g}$ be a fixed symmetric positive definite operator with respect to the inner product given by (6.1). Consider the optimal control problem on G given by

$$\min \int_0^T \frac{1}{4}\langle U, \Sigma(U)\rangle \mathrm{d}t \tag{7.2}$$

subject to $\dot{Q} = QU$ where $U \in \mathfrak{g}$, and where the minimum is taken over all curves $Q(t) \in G$ with $t \in [0, T]$ and with fixed endpoints $Q(0) = Q_0$ and $Q(T) = Q_T$.

The Hamiltonian for the optimal control problem (7.2) is then defined as

$$\begin{aligned} H(P, Q, U) &= \langle P, QU\rangle - \frac{1}{4}\langle U, \Sigma(U)\rangle \\ &= \langle Q^{\mathsf{T}} P, U\rangle - \frac{1}{4}\langle U, \Sigma(U)\rangle. \end{aligned} \tag{7.3}$$

Proposition 7.1. *The necessary conditions for optimality of a solution to the optimal control problem (7.2) with costate $P \in \mathbb{R}^{n\times n}$ yield the following Hamilton's equations*

$$\dot{Q} = QU, \quad \dot{P} = -PU^{\mathsf{T}}. \tag{7.4}$$

Lemma 7.2. *The extremal controls for the optimal control problem (7.2) when $P \in G$ are given by*

$$U_{ext} = \Sigma^{-1}\left(Q^{\mathsf{T}}P - (Q^{\mathsf{T}}P)^{-1}\right). \tag{7.5}$$

The space $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$ is a symplectic manifold with the canonical symplectic form

$$\Omega_{\text{can}}((X_1, Y_1), (X_2, Y_2)) = \langle Y_2, X_1\rangle - \langle Y_1, X_2\rangle. \tag{7.6}$$

Proposition 7.3. *The extremal flow (7.4) generated by the optimal control problem (7.2) which evolves on the canonical symplectic manifold* $(\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}, \Omega_{\text{can}})$ *as a Hamiltonian flow, naturally restricts to a flow on* $G \times G$.

Let $M = Q^\mathsf{T} P - (Q^\mathsf{T} P)^{-1}$, then $M \in \mathfrak{g}$ if $P \in G$ in which case

$$H(P, Q, U_{ext}) = \frac{1}{4}\langle M, \Sigma^{-1}(M)\rangle, \tag{7.7}$$

and the extremal control can be expressed as

$$U_{ext} = \Sigma^{-1}(M) \in \mathfrak{g}. \tag{7.8}$$

Extremal flow in terms of an involution Consider the Lie algebra automorphism of $\mathfrak{g}$ and $\mathfrak{gl}(n)$, given by

$$\widehat{\sigma} : \mathfrak{g} \to \mathfrak{g}; \quad \widehat{\sigma}(A) = -A^{\mathsf{T}}. \tag{7.9}$$

Can show:

Theorem 7.4. *The "generalized Euler" equations for the optimal control problem (7.2) are given by*

$$\dot{Q} = QU, \;\; \dot{M} = [M, \widehat{\sigma}(U)], \;\; U = \Sigma^{-1}(M). \tag{7.10}$$

To pass between the two formulations we consider the map

$$\Phi : G \times G \to G \times \mathfrak{g}, \;\; (Q, P) \;\mapsto\; (Q, M) \tag{7.11}$$

where $M = \sigma(Q^{-1})P - P^{-1}\sigma(Q)$.

The inverse of the map Φ, where defined, is obtained simply by setting

$$P = \sigma(Q) \exp\left(\sinh^{-1} \frac{M}{2}\right), \tag{7.12}$$

Note that $\sinh(\cdot)$ does indeed restrict to a map from $\mathfrak{g}$ to $\mathfrak{g}$ since if $X \in \mathfrak{g}$, $\exp(X) \in G$, and hence $\exp(X) - \exp(-X) \in \mathfrak{g}$ by our earlier observation.

Can show:

Theorem 7.5. *The set $S \subset G \times G \subset \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$ given by*

$$S \triangleq \left\{(Q, P) \in G \times G \ \mid \ m = P\sigma(Q^{-1}) - \sigma(Q)P^{-1}, \ \|m\| < 2\right\}, \tag{7.13}$$

is a symplectic submanifold of $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$.

Discrete Optimal Control Problem

Let the matrix Λ satisfing $\Lambda^{\mathsf{T}} J = J\Lambda$, be such that $\Lambda + \Lambda^{\mathsf{T}}$ is positive definite. Let $\overline{Q}_0$, $\overline{Q}_N \in G$ be given fixed endpoints. We define the optimal control problem

$$\min_{U_k} \sum_{k=1}^{N} \langle \Delta, U_k \rangle, \qquad \Delta = \frac{1}{2}(\Lambda + \Lambda^{\mathsf{T}}), \tag{7.14}$$

subject to

$$Q_{k+1} = Q_k U_k, \qquad Q_0 = \overline{Q}_0, \;\; Q_N = \overline{Q}_N. \tag{7.15}$$

Therefore $U_k = Q_k^{-1} Q_{k+1} \in G$, and Δ is positive definite satisfying the condition $\Delta^{\mathsf{T}} J = \Delta J = J \Delta$.

Theorem 7.6. *A solution of the discrete optimal control problem (7.14) is given by a sequence of matrices (Q_k, P_k) in $G \times G$ satisfying the optimal evolution equations*

$$Q_{k+1} = Q_k U_k, \qquad P_{k+1} = P_k \sigma(U_k), \tag{7.16}$$

where $\sigma : \mathrm{GL}(n) \to \mathrm{GL}(n)$ is the involution defined above, and U_k is defined by

$$U_k \Delta - \Delta U_k^{-1} = P_k^{\mathsf{T}} Q_k - \left(P_k^{\mathsf{T}} Q_k\right)^{-1}. \tag{7.17}$$

8 Fluid Flows and Optimal Control

First introduce the usual dynamics for inviscid, incompressible fluid flow, impulse density and the vorticity dynamics. The basic equations we consider are:

$$\frac{\partial v}{\partial t} + (v \cdot \operatorname{grad})v = -\operatorname{grad} p; \quad \operatorname{div} v = 0 \tag{8.1}$$
$$x \in \Omega; \qquad v = v(x,t), \quad p = p(x,t).$$

We assume, for simplicty only that the flow is in all of space or in a periodic box so we do not need to deal with boundary conditions. This is not an essential restriction.

Here, v is the fluid velocity and p is the pressure. We introduce the impulse density z,

$$z = v + \operatorname{grad} k. \tag{8.2}$$

where k is an arbitrary scalar field, $k = k(x,t)$.

Take the time derivative of (8.2) to get

$$\frac{\partial z}{\partial t} - v \times \operatorname{curl} z = \operatorname{grad} \Lambda, \quad \operatorname{div} v = 0 \tag{8.3}$$

where

$$\Lambda = \frac{\partial k}{\partial t} - p - \frac{1}{2} v \cdot v;$$

Λ is called the gauge. Any choice of gauge is possible, but to be concrete, we consider the "geometric gauge" $\Lambda = -v \cdot z$.

With this choice

$$\frac{\partial z}{\partial t} + (v \cdot \operatorname{grad}) z + (\operatorname{grad} v)^T z = 0, \quad \operatorname{div} v = 0$$

and k is now fixed by the equation

$$\frac{dk}{dt} = p - \frac{1}{2} v \cdot v.$$

Lemma 8.1. $w = \operatorname{curl} z = \operatorname{curl} v$ *satisfies the vorticity equation:*

$$\frac{\partial w}{\partial t} + [v, w] = 0 \tag{8.4}$$

We denote the Lagrange or material variables by X_i and the Euler or spatial variables by x_i, and set

$$x_i = \phi_i(X, t), \quad 1 \leq i \leq 3.$$

We assume $\phi : \Omega \to \Omega$ is a volume preserving diffeomorphism, with Jacobian equal to unity, $|\phi_*| = 1$.

Total vorticity equations:

$$\frac{\partial \phi}{\partial t} = v \circ \phi; \qquad \frac{\partial w}{\partial t} = [w, v] : \operatorname{div} v = 0. \tag{8.5}$$

Compare these equations with the right invariant Euler equations for the rigid body:

$$\dot{Q} = \Omega Q; \qquad \dot{M} = [\Omega, M] \tag{8.6}$$

$$[\Omega, M] = \Omega M - M\Omega \begin{pmatrix} = [M, \Omega] \text{ interpreted} \\ \text{as vector fields} \end{pmatrix}.$$

Optimal Control formulation The problem can be posed as:

$$\min_{v(\cdot)} \frac{1}{2} \int_0^T \langle v, v \rangle dt$$

subject to:

$$\operatorname{div} v = 0; \qquad \frac{\partial \phi}{\partial t} = v \circ \phi \tag{8.7}$$

and

$$\phi(X, 0) = \phi_0(X), \qquad \phi(X, T) = \phi_T(X) \text{ fixed},$$

and, for flow in all of space, suitable conditions at infinity.

Goal here is to analyze Hamilton principle for fluid mechanics from the point of view of the Pontryagin maximum principle.

Thus solve this problem by introducing Lagrange multipliers and the cost

$$J(v, \phi, \pi, k) = \int_0^T \left(\left\langle \pi, v \circ \phi - \frac{\partial \phi}{\partial t} \right\rangle - \frac{1}{2} \langle v, v \rangle + \langle k, \operatorname{div} v \rangle \right) dt$$

The problem (8.7) may be recast as: $\min J$, subject to $\operatorname{div} v = 0$, $\frac{\partial \phi}{\partial t} = v \circ \phi$, and boundary conditions.

Theorem 8.2. *The extremals of problem (8.7) are given by*

$$\begin{aligned} \frac{\partial \pi}{\partial t} =& - (v_* \circ \phi)^T \pi, \quad \frac{\partial \phi}{\partial t} = v \circ \phi \\ v =& \pi \circ \phi^{-1} - \operatorname{grad} k, \quad \operatorname{div} v = 0. \end{aligned} \tag{8.8}$$

Now set

$$\begin{aligned} H(\pi, \phi) &= \frac{1}{2} \langle \operatorname{curl} \pi \circ \phi^{-1}, \psi \rangle \\ &= \frac{1}{2} \langle \pi \circ \phi^{-1}, v \rangle \\ &= \frac{1}{2} \langle \omega, A\omega \rangle \\ &= \frac{1}{2} \langle \operatorname{curl} \pi \circ \phi^{-1}, A \operatorname{curl} \pi \circ \phi^{-1} \rangle. \end{aligned} \tag{8.9}$$

Theorem 8.3. *For this Hamiltonian*

$$\frac{\delta H}{\delta \pi}(\pi, \phi) = v \circ \phi; \qquad \frac{\partial H}{\partial \phi}(\pi, \phi) = (v_* \circ \phi)^T \pi.$$

Thus the extremal equations (8.8) may be written as

$$\frac{\partial \pi}{\partial t} = -\frac{\delta H}{\delta \phi}; \qquad \frac{\partial \phi}{\partial t} = \frac{\delta H}{\delta \pi}. \tag{8.10}$$

These equations are canonical with respect to the natural symplectic form on $L_2(\mathbb{R}^3 : \mathbb{R}^3) \times L_2(\mathbb{R}^3 : \mathbb{R}^3)$

$$\omega((X_1, Y_1), (X_2, Y_2)) = \int_{\mathbb{R}^3} (Y_2 \cdot X_1 - X_2 \cdot Y_1) dx$$

9 Flows on Symmetric Matrices and the Symplectic Group

Consider here analysis of the set of ordinary differential equations

$$\dot{X} = [X^2, N], \tag{9.1}$$

where $X \in \mathrm{Sym}(n)$, the linear space of $n \times n$ symmetric matrices, $\dot{X}$ denotes the time derivative, $N \in \mathfrak{so}(n)$, the space of skew symmetric $n \times n$ matrices, is given, and where initial conditions $X(0) = X_0 \in \mathrm{Sym}(n)$ are also given.

It is easy to check that $[X^2, N] \in \mathrm{Sym}(n)$, so that if the initial condition is in $\mathrm{Sym}(n)$, then $X(t) \in \mathrm{Sym}(n)$ for all t.

Also, because of the straightforward identity $\left[X^2, N\right] = [X, XN + NX]$, this equation may be rewritten in the Lax form

$$\dot{X} = [X, XN + NX], \tag{9.2}$$

again with initial conditions $X(0) = X_0 \in \mathrm{Sym}(n)$.

10 The Lie Algebra

Can regard N as a Poisson tensor on $\mathbb{R}^n$ by defining the bracket of two functions f, g as

$$\{f, g\}_N = (\nabla f)^T N \nabla g. \tag{10.1}$$

The Hamiltonian vector field associated with a function h is given by

$$.X_h(z) = N\nabla h(z), \tag{10.2}$$

For each $X \in \mathrm{Sym}(n)$ define the quadratic Hamiltonian Q_X by

$$Q_X(z) := \frac{1}{2} z^T X z, \quad z \in \mathbb{R}^n.$$

Let $\mathcal{Q} := \{Q_X \mid X \in \mathrm{Sym}(n)\}$ be the vector space of all such functions.

Follows that the Hamiltonian vector field of Q_X has the form

$$X_{Q_X}(z) = NXz. \tag{10.3}$$

Poisson bracket:

Lemma 10.1. *For* $X, Y \in \mathrm{Sym}(n)$*, we have*

$$\{Q_X, Q_Y\}_N = Q_{[X,Y]_N}, \tag{10.4}$$

where $[X, Y]_N = XNY - YNX \in \mathrm{Sym}(n)$*. In addition,* $\mathrm{Sym}(n)$ *is a Lie algebra relative to the Lie bracket* $[\cdot, \cdot]_N$*. Therefore,* $Q : X \in (\mathrm{Sym}(n), [\cdot, \cdot]_N) \mapsto Q_X \in (\mathcal{Q}, \{\cdot, \cdot\}_N)$ *is a Lie algebra isomorphism.*

Know:

$$[X_f, X_g] = -X_{\{f,g\}}. \tag{10.5}$$

If we take $f = Q_X$ and $g = Q_Y$, with $X_f = NX$ and $X_g = NY$, and recall that the Jacobi-Lie bracket of *linear* vector fields is the negative of the commutator of the associated matrices, then we have

Proposition 10.2. *Equations (10.4) and (10.5) imply*

$$N[X, Y]_N = [NX, NY]\,. \tag{10.6}$$

Letting $\mathcal{LH}$ denote the Lie algebra of linear Hamiltonian vector fields on $\mathbb{R}^n$ relative to the commutator bracket of matrices, (10.6) states that the map

$$X \in (\mathrm{Sym}(n), [\cdot,\cdot]_N) \mapsto NX \in (\mathcal{LH}, [\cdot,\cdot])$$

is a homomorphism of Lie algebras.

Have:

Proposition 10.3. *Let $N \in \mathfrak{so}(n)$. The map $Q : X \in (\mathrm{Sym}(n), [\cdot,\cdot]_N) \mapsto Q_X \in (\mathcal{Q}, \{\cdot,\cdot\}_N)$ is a Lie algebra isomorphism. The map $X \in (\mathrm{Sym}(n), [\cdot,\cdot]_N) \mapsto NX \in (\mathcal{LH}, [\cdot,\cdot])$ is a Lie algebra homomorphism and if N is invertible it induces an isomorphism of $(\mathrm{Sym}(n), [\cdot,\cdot]_N)$ with $\mathfrak{sp}(n,\mathbb{R})$.*

Euler-Poincaré Form

Identify $\mathrm{Sym}(n)$ with its dual using the the positive definite inner product

$$\langle\langle X, Y \rangle\rangle := \mathrm{trace}\,(XY)\,, \quad \text{for} \quad X, Y \in \mathrm{Sym}(n). \tag{10.7}$$

Remark. The inner product $\langle\langle X, Y \rangle\rangle$ is not ad invariant relative to the N-bracket, but another one, namely $\kappa_N(X, Y) := \mathrm{trace}(NXNY)$ is invariant, as is easy to check.

Define the Lagrangian $l : \mathrm{Sym}(n) \to \mathbb{R}$ on the Lie algebra $(\mathrm{Sym}(n), [\cdot, \cdot]_N)$ by

$$l(X) = \frac{1}{2}\,\mathrm{trace}\left(X^2\right) = \frac{1}{2}\,\mathrm{trace}\left(XX^T\right) =: \frac{1}{2}\,\langle\langle X, X \rangle\rangle\,. \tag{10.8}$$

Proposition 10.4. *The equations*

$$\dot{X} = [X^2, N] \tag{10.9}$$

are the Euler-Poincaré equations corresponding to the Lagrangian (10.8) *on the Lie algebra* $(\mathrm{Sym}(n), [\cdot, \cdot]_N)$.

Noninvertible case Let $2p = \operatorname{rank} N$ and $d := n - 2p$. Then $\bar{N} := N|_{\operatorname{im} N} : \operatorname{im} N \to \operatorname{im} N$ defines a nondegenerate skew symmetric bilinear form and, by the previous proposition, $(\operatorname{Sym}(2p), [\cdot,\cdot]_{\bar{N}})$ is isomorphic as a Lie algebra to $(\mathfrak{sp}(\mathbb{R}^{2p}, \bar{N}^{-1}), [\cdot,\cdot])$.

Proposition 10.5. *Can find a map*

$$\Psi : ((\operatorname{Sym}(2p) \circledS \mathcal{M}_{(2p)\times d}) \oplus \operatorname{Sym}(d), [\cdot,\cdot]^C) \to (\operatorname{Sym}(n), [\cdot,\cdot]_N)$$

given by

$$\Psi(S, A, B) := \begin{bmatrix} S & A \\ A^T & B \end{bmatrix} \tag{10.10}$$

which is a Lie algebra isomorphism.

Poisson structure

Identifying Sym(n) with its dual using the inner product (10.7) endows Sym(n) with the the (left, or minus) Lie Poisson bracket

$$\{f, g\}_N (X) = -\operatorname{trace}\Big[X\Big(\nabla f(X) N \nabla g(X) - \nabla g(X) N \nabla f(X)\Big)\Big], \tag{10.11}$$

where ∇f is the gradient of f relative to the inner product $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ on Sym(n). It is easy to check that the equations $\dot{X} = \left[X^2, N\right]$ are Hamiltonian relative to the function l defined in (10.8) and the Lie-Poisson bracket (10.11).

Later on we shall also need the *frozen* Poisson bracket

$$\{f, g\}_{FN} (X) = -\operatorname{trace}\Big(\nabla f(X) N \nabla g(X) - \nabla g(X) N \nabla f(X)\Big). \tag{10.12}$$

It is a general fact that the Poisson structures (10.11) and (10.12) are *compatible* in the sense that their sum is a Poisson structure.

Proposition 10.6. *Let $n = 2p + d$, where $2p = \operatorname{rank} N$. The generic leaves of the Lie-Poisson bracket $\{\cdot, \cdot\}_N$ are $2p(p + d)$-dimensional.*

Proposition 10.7. *All leaves of the frozen Poisson bracket $\{\cdot, \cdot\}_{FN}$ are*

(i) *$2p(p+d)$-dimensional if N is generic, that is, all its non-zero eigenvalues are distinct, and*

(ii) *$p(p + 1 + 2d)$-dimensional if all non-zero eigenvalue pairs of N are equal.*

Proposition 10.8. *Denote the value at* $X \in \mathrm{Sym}(n)$ *of the Poisson tensors corresponding to the Lie-Poisson* (10.11) *and frozen* (10.12) *brackets by* B_X *and* C_X*, respectively. Then for any* $Y \in \mathrm{Sym}(n)$ *we have*

$$B_X(Y) = XYN - NYX \tag{10.13}$$

$$C_X(Y) = YN - NY. \tag{10.14}$$

Casimir Functions.

Proposition 10.9. *Let the skew symmetric matrix N have rank $2p$ and size $n := 2p + d$. Choose an orthonormal basis of $\mathbb{R}^{2p+d}$ in which N is written as*

$$N = \begin{bmatrix} 0 & V & 0 \\ -V & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

where V is a real diagonal matrix whose entries are $v_1, \ldots, v_p$.

(i) *If $v_i \neq v_j$ for all $i \neq j$, the $p + d(d+1)/2$ Casimir functions for the frozen Poisson structure are given by*

$$C_F^i(X) = \mathrm{trace}(E_i X), \quad i = 1, \ldots, p + \frac{1}{2}d(d+1),$$

where E_i is any of the matrices

$$\begin{bmatrix} S_{kk} & 0 & 0 \\ 0 & S_{kk} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & S_{ab} \end{bmatrix}.$$

Here S_{kk} is the $p \times p$ matrix all of whose entries are zero except the diagonal (k,k) entry which is one and S_{ab} is the $d \times d$ symmetric matrix having all entries equal to zero except for the (a,b) and (b,a) entries that are equal to one.

(ii) *If $v_i = v_j$ for all $i, j = 1, \ldots, p$, the $p^2 + d(d+1)/2$ Casimir functions for the frozen Poisson structure are given by*

$$C_F^i(X) = \operatorname{trace}(E_i X), \quad i = 1, \ldots, p^2 + \frac{1}{2}d(d+1),$$

where E_i is any of the matrices

$$\begin{bmatrix} S_{kl} & 0 & 0 \\ 0 & S_{kl} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & A_{kl} & 0 \\ -A_{kl} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & S_{ab} \end{bmatrix}.$$

Here S_{kl} is the $p \times p$ symmetric matrix having all entries equal to zero except for the (k,l) and (l,k) entries that are equal to one and A_{kl} is the $p \times p$ skew symmteric matrix with all entries equal to zero except for the (k,l) entry which is 1 and the (l,k) entry which is -1.

(iii) *Denote*

$$\bar{N} = \begin{bmatrix} 0 & V \\ -V & 0 \end{bmatrix} \quad \textit{and} \quad \hat{N} = \begin{bmatrix} \bar{N}^{-1} & 0 \\ 0 & 0 \end{bmatrix}.$$

The $p + d(d+1)/2$ *Casimir functions for the Lie-Poisson bracket* $\{\cdot,\cdot\}_N$ *are given by*

$$C^k(X) = \frac{1}{2k}\,\mathrm{trace}\left[\left(X\hat{N}\right)^{2k}\right], \quad \textit{for} \quad k = 1,\ldots,p$$

and

$$C^k(X) = \mathrm{trace}(XE_k), \quad \textit{for} \quad k = p+1,\ldots,p+\frac{1}{2}d(d+1)\,,$$

where E_k *is any matrix of the form*

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & S_{ab} \end{bmatrix}.$$

In the special case when N *is full rank the Casimirs are just*

$$C^k(X) = \frac{1}{2k}\,\mathrm{trace}\left[\left(XN^{-1}\right)^{2k}\right], \quad \textit{for} \quad k = 1,\ldots,p,$$

Mischenko-Fomenko

Can show that our equation is not of the sectional operator type. However the system may be mapped to a Mischenko-Fomenko type system in the case N is invertible with distinct eigenvalues.

The Mischenko-Fomenko Construction. Consider a semisimple complex or real split Lie algebra $\mathfrak{g}$ with Killing form $\langle \cdot, \cdot \rangle$. Let $\mathfrak{h}$ be a Cartan subalgebra, let $a, b \in \mathfrak{h}$ and a be regular (i.e. its value on every root is non-zero). Define the ***sectional operators*** $C_{a,b,D} : \mathfrak{g} \to \mathfrak{g}$ by $C_{a,b,D}(\xi) := \mathrm{ad}_a^{-1} \mathrm{ad}_b(\xi_1) + D(\xi_2)$ where $\xi = \xi_1 + \xi_2$, $\xi_2 \in \mathfrak{h}$, $\xi_1 \in \mathfrak{h}^\perp$ (the perpendicular is taken relative to the Killing form and thus $\mathfrak{h}^\perp$ is the direct sum of all the root spaces), and $D : \mathfrak{h} \to \mathfrak{h}$ is an arbitrary invertible symmetric operator on $\mathfrak{h}$. Then $C_{a,b,D} : \mathfrak{g} \to \mathfrak{g}$ is an invertible symmetric operator (relative to the Killing form) satisfying the condition

$$[C_{a,b,D}(\xi), a] = [\xi, b] \tag{10.15}$$

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

The equations of motion are

$$\dot{\xi} = [\xi, C_{a,b,D}(\xi)]. \tag{10.16}$$

For N invertible we can show can map the system to one of MF type:

$$\dot{Z} = [Z, NZN] \tag{10.17}$$

11 Lax Pairs with Parameter

To prove that system (9.1) is integrable for any choice of N, we will compute its flow invariants.

Due to isospectral representation (9.2), we already know that the eigenvalues of X, or alternatively, the quantities trace X^k for $k = 1, 2, \ldots, n-1$, are invariants.

Rewrite the system as a Lax pair with a parameter. One can do this in a fashion similar to that for the generalized rigid body equations.

Theorem 11.1. *Let λ be a real parameter. The system (9.2) is equivalent to the following Lax pair system*

$$\frac{d}{dt}(X + \lambda N) = \left[X + \lambda N, NX + XN + \lambda N^2\right] \tag{11.1}$$

Recall Manakov:

$$\frac{d}{dt}(M + \lambda\Lambda^2) = [M + \lambda\Lambda^2, \Omega + \lambda\Lambda]. \tag{11.2}$$

For the generalized rigid body the nontrivial coefficients of $\lambda^i, 0 < i < k$ in the traces of the powers of $M + \lambda\Lambda^2$ then yield the right number of independent integrals in involution to prove integrability of the flow on a generic adjoint orbit of $SO(n)$ (identified with the corresponding coadjoint orbit). The case $i = 0$ needs to be eliminated, because these are Casimir functions.

Similarly, in our case, the nontrivial coefficients of $\lambda^i, 0 \leq i \leq k$, in

$$h_k^\lambda(X) := \frac{1}{k}\operatorname{trace}(X + \lambda N)^k, \qquad k = 1, 2, \ldots, n-1 \tag{11.3}$$

yield the conserved quantities.

We find the nontrivial invariants

$$\operatorname{trace} \sum_{|i|=k-2r} \sum_{|j|=2r} X^{i_1} N^{j_1} X^{i_2} \cdots X^{i_s} N^{j_s} \tag{11.4}$$

for $i_q, j_q = 0, \ldots, k-1$, $r = 1, \ldots, \left[\frac{k-1}{2}\right]$, where $[p]$ denotes the integer part of $p \in \mathbb{R}$. Altogether, this results in

$$\left[\frac{n}{2}\right]\left[\frac{n+1}{2}\right]$$

invariants as an easy inductive argument shows.

Are these integrals the right candidates to prove complete integrability of the system $\dot{X} = [X^2, N]$?

- If N is invertible, then $n = 2p$ and hence

$$\left[\frac{n}{2}\right]\left[\frac{n+1}{2}\right] = \left[\frac{2p}{2}\right]\left[\frac{2p+1}{2}\right] = p^2 = \frac{1}{2}\left(2p^2 + p - p\right)$$
$$= \frac{1}{2}\left(\dim \mathfrak{sp}(2p, \mathbb{R}) - \operatorname{rank} \mathfrak{sp}(2p, \mathbb{R})\right)$$

which is half the dimension of the generic adjoint orbit in $\mathfrak{sp}(2p, \mathbb{R})$. Therefore, these conserved quantities are the right candidates to prove that this system is integrable on the generic coadjoint orbit of $\mathrm{Sym}(n)$.

- If N is non-invertible (which is equivalent to $d \neq 0$), then $n = 2p + d$ and hence

$$\begin{aligned}\left[\frac{n}{2}\right]\left[\frac{n+1}{2}\right] &= \left[\frac{2p+d}{2}\right]\left[\frac{2p+d+1}{2}\right] \\ &= \left(p + \left[\frac{d}{2}\right]\right)\left(p + \left[\frac{d+1}{2}\right]\right) \\ &= p^2 + p\left(\left[\frac{d}{2}\right] + \left[\frac{d+1}{2}\right]\right) + \left[\frac{d}{2}\right]\left[\frac{d+1}{2}\right] \\ &= p^2 + pd + \left[\frac{d}{2}\right]\left[\frac{d+1}{2}\right].\end{aligned}$$

The right number of integrals is $p(p+d)$ according to Proposition 10.6, so this calculation seems to indicate that there are additional integrals. The situation is not so simple since there are redundancies due to the degeneracy of N. Note, however, that if $d = 1$, then we do get the right number of integrals.

12 Integrability

This section shows that the Hamiltonian system (9.1) is integrable in the case $n = 2p$.

Bihamiltonian structure. We begin with the following observation.

Proposition 12.1. *The system $\dot{X} = X^2N - NX^2$ is Hamiltonian with respect to the bracket $\{f, g\}_N$ defined in (10.11) using the Hamiltonian $h_2(X) := \frac{1}{2}\operatorname{trace}(X^2)$ and is also Hamiltonian with respect to the compatible bracket $\{f, g\}_{FN}$ defined in (10.12) using the Hamiltonian $h_3(X) := \frac{1}{3}\operatorname{trace}(X^3)$.*

Involution. We prove that the $\left[\frac{n}{2}\right]\left[\frac{n+1}{2}\right]$ integrals given in (11.4), namely

$$h_{k,2r}(X) := \operatorname{trace} \sum_{|i|=k-2r} \sum_{|j|=2r} X^{i_1}N^{j_1}X^{i_2}\cdots X^{i_s}N^{j_s},$$

where $i_q, j_q = 0, \ldots, k-1$, $r = 1, \ldots, \left[\frac{k-1}{2}\right]$, $k = 1, \ldots, n-1$, are in involution. Denote by $h_{k,k-r}$ the coefficient of λ^{k-r} in $\frac{1}{k}\operatorname{trace}\left(X + \lambda N\right)^k$ so that we have

$$h_k^\lambda(X) = \frac{1}{k}\operatorname{trace}\left(X + \lambda N\right)^k = \sum_{r=0}^{k} \lambda^{k-r} h_{k,k-r}(X)\,. \tag{12.1}$$

As explained before, not all of these coefficients should be counted: roughly half of them vanish and the last one, namely, $h_{k,k}$, is the constant N^k. Consistent with our notation for the Hamiltonians, we set $h_k = h_{k,0}$.

Firstly we need the gradients of the functions h_k^λ.

Lemma 12.2. *The gradients ∇h_k^λ are given by*

$$\nabla h_k^\lambda(X) = \frac{1}{2}(X + \lambda N)^{k-1} + \frac{1}{2}(X - \lambda N)^{k-1}. \tag{12.2}$$

Proposition 12.3.

$$B_X(\nabla h_k^\lambda(X)) = C_X(\nabla h_{k+1}^\lambda(X)) \tag{12.3}$$

Proposition 12.4. *The functions $h_{k,k-r}$ satisfy the recursion relation*

$$B_X(\nabla h_{k,k-r}(X)) = C_X(\nabla h_{k+1,k-r}(X)) \tag{12.4}$$

Using the recursion relations involution follows immediately.

Proposition 12.5. *The invariants $h_{k,k-r}$ are in involution with respect to both Poisson brackets $\{f,g\}_N$ and $\{f,g\}_{FN}$.*

Proof. The definition of the Poisson tensors B_X and C_X and the recursion relation (12.4) give

$$\begin{aligned}
\{h_{k,k-r}, h_{l,l-q}\}_N &= \langle\langle \nabla h_{k,k-r}(X), B_X(\nabla h_{l+1,l-q}(X))\rangle\rangle \\
&= \langle\langle \nabla h_{k,k-r}(X), C_X(\nabla h_{l+1,l-q}(X))\rangle\rangle \\
&= \{h_{k,k-r}, h_{l+1,l-q}\}_{FN} = -\{h_{l+1,l-q}, h_{k,k-r}\}_{FN} \\
&= -\langle\langle \nabla h_{l+1,l-q}(X), C_X(\nabla h_{k,k-r}(X))\rangle\rangle \\
&= -\langle\langle \nabla h_{l+1,l-q}(X), B_X(\nabla h_{k-1,k-r}(X))\rangle\rangle \\
&= -\{h_{l+1,l-q}, h_{k-1,k-r}\}_N = \{h_{k-1,k-r}, h_{l+1,l-q}\}_N
\end{aligned}$$

for any $k,l = 1,\ldots,n-1$, $r = 1,\ldots,k$ and $q = 0,\ldots,l-1$.

Repeated application of this relation eventually leads to Hamiltonians $h_{k,k-r}$ where either $k-r$ is a power that does not exist for k, in which case the Hamiltonian is zero, or one is led to $h_{0,0}$ which is constant. This shows that $\{h_{k,k-r}, h_{l,l-q}\}_N = 0$ for any pair of indices.

In a similar way one shows that $\{h_{k,k-r}, h_{l,l-q}\}_{FN} = 0$. ∎

Independence

Theorem 12.6. *For generic N the integrals $h_{k,2r}$ given by equation (11.4) are independent.*

Hence, since we have involution and independence we have proved the following.

Theorem 12.7. *For N invertible with distinct eigenvalues the system* (9.1) *is completely integrable.*

Corollary 12.8. *For N odd with distinct eigenvalues and nullity one, the system* (9.1) *is completely integrable.*

It is also of interest to analyze linearization on the Jacobi variety of the curve

$$\det(zI - \lambda N - X) = 0$$

– use work of Adler/van Moerbeke. Griffiths.