3.1

MSRI
Mathematical Sciences Research Institute

17 Gauss Way Berkeley, CA 94720-5070 p: 510.642.0143 f: 510.642.8609 www.msri.org

NOTETAKER CHECKLIST FORM

(Complete one for each talk.)

Name: ORI KATZ Email/Phone: ORIKATZ.OK@gmail.com

Speaker's Name: Anthony Bloch

Talk Title: The Clebsch representation in optimal control, integrable systems & discrete dynamics

Date: 10/10/18 Time: 9:15 am / pm (circle one)

Please summarize the lecture in 5 or fewer sentences: Bloch discussed recent work on a geometric approach to certain optimal control problems & the relationship of solutions to some classical integrable dynamical systems, including the rigid body equations, geodesic flows on the ellipsoid & Toda lattice flows. Discussed relations to Hamiltonian structure, discrete dynamics & symplectic integration.

CHECK LIST

(This is **NOT** optional, we will **not** pay for **incomplete** forms)

- ☒ Introduce yourself to the speaker prior to the talk. Tell them that you will be the note taker, and that you will need to make copies of their notes and materials, if any.
- ☒ Obtain ALL presentation materials from speaker. This can be done before the talk is to begin or after the talk; please make arrangements with the speaker as to when you can do this. You may scan and send materials as a .pdf to yourself using the scanner on the 3rd floor.
  - **Computer Presentations**: Obtain a copy of their presentation
  - **Overhead**: Obtain a copy or use the originals and scan them
  - **Blackboard**: Take blackboard notes in black or blue **PEN**. We will **NOT** accept notes in pencil or in colored ink other than black or blue.
  - **Handouts**: Obtain copies of and scan all handouts
- ☐ For each talk, all materials must be saved in a single .pdf and named according to the naming convention on the "Materials Received" check list. To do this, compile all materials for a specific talk into one stack with this completed sheet on top and insert face up into the tray on the top of the scanner. Proceed to scan and email the file to yourself. Do this for the materials from each talk.
- ☐ When you have emailed all files to yourself, please save and re-name each file according to the naming convention listed below the talk title on the "Materials Received" check list. (*YYYY.MM.DD.TIME.SpeakerLastName*)
- ☐ Email the re-named files to notes@msri.org with the workshop name and your name in the subject line.

The Clebsch Representation in Optimal Control, integrable systems, and discrete dynamics

Anthony Bloch
Work with Francois Gay-Balmaz, Tudor Ratiu

(See also earlier work with Brockett, Crouch, Marsden, Nordkvist, Sanyal)

- **Symmetric rigid body equations – smooth and discrete**

- **Optimal Control and the Clebsch problem**

- **Flows on Stiefel manfolds, the Neumann problem, Jacobi flow on ellipsoid**

- **Integrable Flows on symmetric matrices, Toda flows**

- **Moser and the Geometry of Quadrics**

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: general and tridiagonal

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

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

(with Brockett, Ratiu, Flaschka)
Flow on the symmetric matrices/symplectic groups

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

(w. Iserles, Marden, Ratiu)

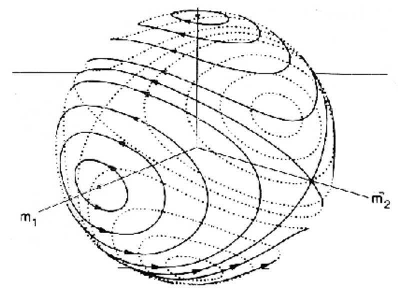

Figure 0.1: Rigid Body Phase Portrait

1 The n-dimensional Rigid Body.

• Here review the classical rigid body equations in in n dimensions.
Euler equations for the rigid body in three dimensions:

$$I\dot{\Omega} = I\Omega \times \Omega \tag{1.1}$$

–dynamics in the body frame. I here is the moment of inertia matrix and Ω the vector of angular velocities in the body frame. Also need kinematics: $\dot{\theta} = \theta\hat{\Omega}$ where hat maps vectors in $\mathbb{R}^3$ to matrices in the special orthogonal group, a Lie algebra homomorphism.

The dynamics evolve on the momentum sphere: $||\mathbf{M}^2|| = M_1^2 + M_2^2 + M_3^2 = c$ **where** $M_i = I_i\Omega_i$**.**

Simpler than using the Euler angles. This is an Euler-Poincaré equation.

Hamiltonian form, a Lie-Poisson equation:

$$\dot{\mathbf{M}} = \mathbf{M} \times \Omega = \{\mathbf{M}, H(\mathbf{M})\}$$

where $H = 1/2(\mathbf{M} \cdot \Omega)$**. This is integrable system.**

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

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

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 (8.6) gives the second of the equations (RBn). ■

Proposition 2.2. *For a solution of the left invariant rigid body equations (RBn) obtained by means of Proposition 2.1, the spatial angular momentum is given by* $m = PQ^T - QP^T$ *and hence* m *is conserved along the rigid body flow.*

• 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 2.3. *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 2.4. *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**

$$\dot{Q}_r = \Omega_r Q_r; \qquad \dot{P}_r = \Omega_r P_r \tag{RightSRBn}$$

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 2.5. *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 2.6. *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 Control and Optimal Control

An *affine nonlinear control systems*, has the form

$$\dot{x} = f(x,u) = f(x) + \sum_{i=1}^{m} g_i(x)u_i\,, \tag{3.1}$$

where f and the g_i, $i = 1,\ldots,m$, are smooth vector fields on a smooth manifold M and u_i are admissable controls. The vector field f is usually called the *drift vector field*, and the g_i are called the *control vector fields*. When $f = 0$ we say the system is kinematic.

Typical optimal control problem: given a smooth cost function $g(x,u)$

$$\min_{u(\cdot)} \int_0^T g(x,u)dt\,, \tag{3.2}$$

subject to the following conditions:

(i) **a differential equation constraint $\dot{x} = f(x,u)$, and a state space constraint $x \in M$, and a constraint on the controls $u \in \Omega \subset \mathbb{R}^k$;**

(ii) **the endpoint conditions: $x(0) = x_0$ and $x(T) = x_T$.**

• Pontryagin maximum principle

Necessary conditions for optimality: introduce a parametrized Hamiltonian function on T^*M:

$$\hat{H}(x,p,u) = \langle p, f(x,u)\rangle - p_0 g(x,u), \tag{3.3}$$

where $p_0 \geq 0$ is a fixed positive constant, and $p \in T^*M$. Note that p_0 is the multiplier of the cost function and that $\hat{H}$ is linear in p.

We denote by $t \mapsto u^*(t)$ a curve that satisfies the following relationship along a trajectory $t \mapsto (x(t), p(t))$ in T^*M:

$$H(x(t), p(t), u^*(t)) = \max_{u \in \Omega} \hat{H}(x(t), p(t), u). \tag{3.4}$$

Then if u^* is defined implicitly as a function of x and p by equation (3.4), we can define H^* by

$$H^*(x(t), p(t), t) = H(x(t), p(t), u^*(t)). \tag{3.5}$$

The time-varying Hamiltonian function H^* defines a time-varying Hamiltonian vector field X_{H^*} on T^*M with respect to the canonical symplectic structure on T^*M.

Integrability: recall Arnold-Lioville theorem: for a system on a $2n$-dimensional sympletic manifold M a Hamiltonian system with Hamiltonian H is integrable if there exist n almost everywhere independent integrals on M which are involution – commute under the Poisson bracket – with H.

4 Optimal Control formulation of Rigid Body

Definition 4.1. *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*

$$\min_{U \in \mathfrak{so}(n)} \frac{1}{4} \int_0^T \langle U, J(U) \rangle dt \tag{4.1}$$

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

$$\dot{Q} = QU \qquad Q(0) = Q_0, \qquad Q(T) = Q_T. \tag{4.2}$$

Proposition 4.2. *The rigid body optimal control problem (4.1) has optimal evolution equations (8.6) where P is the costate vector given by the maximum principle.*

The optimal controls in this case are given by

$$U = J^{-1}(Q^T P - P^T Q). \tag{4.3}$$

The proof involves writing the Hamiltonian of the maximum principle as

$$H = \langle P, QU \rangle + \frac{1}{4} \langle U, J(U) \rangle, \tag{4.4}$$

where the costate vector P is a multiplier enforcing the dynamics, and then maximizing with respect to U in the standard fashion (see, for example, Brockett [1973]).

5 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

$$S_d = \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}) \tag{5.1}$$

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*

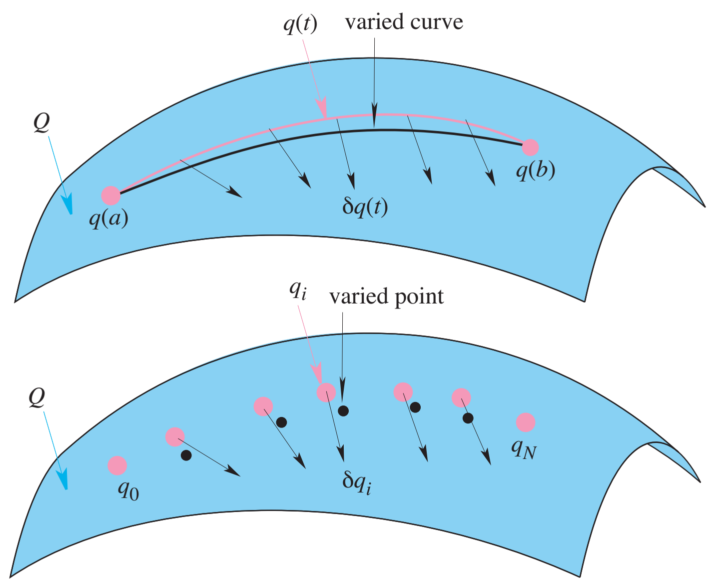

Figure 5.1: The discrete variational principle.

$$D_2L_d(q_{k-1},q_k)+D_1L_d(q_k,q_{k+1})=0, \tag{5.2}$$

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

We can rewrite this as follows

$$D_2L_d + D_1L_d \circ \Phi = 0, \tag{5.3}$$

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

6 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

$$\begin{aligned} \Omega_k &= Q_k^T Q_{k-1} && (6.1)\\ M_k &= \Omega_k^T \Lambda - \Lambda \Omega_k && (6.2)\\ M_{k+1} &= \Omega_k M_k \Omega_k^T. && (6.3) \end{aligned} \qquad \text{(DRBn)}$$

The Moser-Veslov equations (6.1)-(6.3) 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

$$S = \sum_k \operatorname{trace}(Q_k I Q_{k+1}) \tag{6.4}$$

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**

$$
U_k \Lambda - \Lambda U_k^T = Q_k^T P_k - P_k^T Q_k \,. \tag{6.5}
$$

Using these equations, we have the algorithm $(Q_k, P_k) \mapsto (Q_{k+1}, P_{k+1})$ **defined by: compute** U_k **from (11.2), compute** Q_{k+1} **and** P_{k+1} **using (11.1). We note that the update map for** Q **and** P **is done in parallel here.**

Have:

Proposition 6.1. *The symmetric discrete rigid body equations (11.1) on* S *are equivalent to the Moser-Veselov equations (6.1)– (6.3) (DRBn) on the set* S_M *where* S *and* S_M *are defined in Proposition 2.3.*

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

$$\hat{V} = \sum_{k=1}^{N} \operatorname{trace}(\Lambda U_k). \tag{6.6}$$

Define the optimal control problem

$$\min_{U_k} \hat{V} = \min_{U_k} \sum_{k=1}^{N} \operatorname{trace}(\Lambda U_k) \tag{6.7}$$

subject to dynamics and initial and final data

$$Q_{k+1} = Q_k U_k, \qquad Q_0 = \overline{Q}_0, \qquad Q_N = \overline{Q}_N \tag{6.8}$$

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

Theorem 6.3. *A solution of the optimal control problem (6.2) satisfies the optimal evolution equations (11.1)*

$$Q_{k+1} = Q_k U_k; \qquad P_{k+1} = P_k U_k\,, \tag{6.9}$$

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

$$U_k\Lambda - \Lambda U_k^T = Q_k^T P_k - P_k^T Q_k \,. \tag{6.10}$$

7 The Clebsch Optimal control problem

We recall from Gay-Balmaz and Ratiu some facts concerning the Clebsch optimal control problem. Let $\Phi : G \times Q \to Q$ be an action of a Lie group G on a manifold Q. We denote by $qg := \Phi(q, g)$ the action of $g \in G$ on $q \in Q$. Given $u \in \mathfrak{g}$, where $\mathfrak{g}$ is the Lie algebra of G, we denote by $u_Q \in \mathfrak{X}(Q)$ the infinitesimal generator of the action.

Recall that u_Q is the vector field on Q defined at $q \in Q$ by $u_Q(q) := \frac{d}{dt}\big|_{t=0} q\exp(tu)$, where $\exp : \mathfrak{g} \to G$ **is the exponential map.**

Given a cost function $\ell : \mathfrak{g} \times Q \to \mathbb{R}$, also called here a Lagrangian, *Clebsch optimal control problem* is

$$\min_{u(t)} \int_0^T \ell(u(t), q(t))dt \tag{7.1}$$

subject to the following conditions:

(A) $\dot{q}(t) = u(t)_Q(q(t))$**;**

(B) $q(0) = q_0$ **and** $q(T) = q_T$**.**

The *momentum map* for the cotangent lifted G-action on T^*Q is the map $\mathbf{J}: T^*Q \to \mathfrak{g}^*$, defined by $\langle \mathbf{J}(\alpha_q), u\rangle := \langle \alpha_q, u_Q(q)\rangle$, for any $\alpha_q \in T^*Q$, $u \in \mathfrak{g}$. This map is equivariant relative to the cotangent lifted G-action on T^*Q and the coadjoint G-action on $\mathfrak{g}^*$.

If $\alpha, \beta \in T^*_qQ$, the *vertical lift of β relative to α* is defined by

$$\mathrm{Ver}_\alpha \beta := \left.\frac{d}{ds}\right|_{s=0} (\alpha + s\beta) \in T_\alpha(T^*Q).$$

Theorem 7.1. *Assume that $u \in \mathfrak{g} \mapsto \frac{\delta \ell}{\delta u} \in \mathfrak{g}^*$ is a diffeomorphism and that G act on the left (resp. on the right). Then*

- *An extremal solution of the Clebsch optimal control problem (7.1) is a solution of*

$$\frac{\delta \ell}{\delta u} = \mathbf{J}(\alpha), \quad \dot{\alpha} = u_{T^*Q}(\alpha) + \mathrm{Ver}_\alpha \frac{\delta \ell}{\delta q}. \tag{7.2}$$

- *These equations are Hamiltonian on T^*Q relative to the Hamiltonian H given by*

$$H(\alpha_q) = h(\mathbf{J}(\alpha_q), q),$$

  where $h: \mathfrak{g}^* \times Q \to \mathbb{R}$ is the Hamiltonian obtained from ℓ by Legendre transformation on $\mathfrak{g}$.

- *These equations imply (a generalization of) the Euler-Poincaré equations for the control u, given by*

$$\frac{d}{dt}\frac{\delta\ell}{\delta u} = -\operatorname{ad}_u^*\frac{\delta\ell}{\delta u} + \mathbf{J}\left(\frac{\delta\ell}{\delta q}\right), \quad \textit{resp.} \quad \frac{d}{dt}\frac{\delta\ell}{\delta u} = \operatorname{ad}_u^*\frac{\delta\ell}{\delta u} + \mathbf{J}\left(\frac{\delta\ell}{\delta q}\right). \tag{7.3}$$

The case of a representation: When Q is a dual vector space V^* on which G acts by a left representation, $a \in V^* \mapsto ga \in V^*$, the momentum map recovers the usual expression

$$\mathbf{J}\left(\frac{\delta \ell}{\delta a}\right) = -\frac{\delta \ell}{\delta a} \diamond a$$

appearing in semidirect product theory. Here the diamond operator $\diamond : V \times V^* \to \mathfrak{g}$ is defined by $\langle v \diamond a, \xi \rangle = -\langle \xi a, v \rangle$, where ξa denotes the infinitesimal action of the Lie algebra element $\xi \in \mathfrak{g}$ on $a \in V^*$. We now reformulate the previous in this case.

Theorem 7.2. *Assume that $u \in \mathfrak{g} \mapsto \frac{\delta \ell}{\delta u} \in \mathfrak{g}^*$ is a diffeomorphism and that V^* is a left (resp. right) representation space of G. Then*

- *An extremal solution of the Clebsch optimal control problem (7.1) is a solution of*
$$\frac{\delta \ell}{\delta u} = -p \diamond a, \quad \dot{a} = ua, \quad \dot{p} = up + \frac{\delta \ell}{\delta a}, \quad \textit{resp.} \qquad \dot{a} = au, \quad \dot{p} = pu + \frac{\delta \ell}{\delta a} \tag{7.4}$$
- *These equations are Hamiltonian on T^*V^* relative to the Hamiltonian H given by*
$$H(a,p) = h(-p \diamond a, a),$$
where $h : \mathfrak{g}^* \times V^* \to \mathbb{R}$ is the Hamiltonian obtained from ℓ by Legendre transformation on $\mathfrak{g}$.
- *These equations imply the Euler-Poincaré equations for semidirect product for the control u, given by*
$$\frac{d}{dt}\frac{\delta \ell}{\delta u} = -\operatorname{ad}^*_u \frac{\delta \ell}{\delta u} - \frac{\delta \ell}{\delta a} \diamond a, \quad \textit{resp.} \quad \frac{d}{dt}\frac{\delta \ell}{\delta u} = \operatorname{ad}^*_u \frac{\delta \ell}{\delta u} - \frac{\delta \ell}{\delta a} \diamond a. \tag{7.5}$$

Equivalently, the Lie-Poisson equations on the dual of the semidirect product Lie algebra $\mathfrak{g}\circledS V$ *hold:*

$$\begin{cases} \dfrac{d}{dt}\mu = -\operatorname{ad}^*_{\frac{\delta h}{\delta \mu}}\mu + \dfrac{\delta h}{\delta a}\diamond a \\ \dfrac{d}{dt}a = \dfrac{\delta h}{\delta \mu}a \end{cases}, \quad \textit{resp.} \quad \begin{cases} \dfrac{d}{dt}\mu = \operatorname{ad}^*_{\frac{\delta h}{\delta \mu}}\mu + \dfrac{\delta h}{\delta a}\diamond a \\ \dfrac{d}{dt}a = \dfrac{\delta h}{\delta \mu}a \end{cases} \tag{7.6}$$

Restriction to G-orbits:

Note that, due to condition (A), the solution $q(t)$ of the Clebsch optimal control problem (7.1) necessarily preserves the G-orbit $\mathscr{O}$ of the initial condition q_0. Therefore, we always assume that q_0 and and q_T belong to the same G-orbit, in order to have a well posed problem. As a consequence, the Clebsch optimal control problem (7.1) on $\mathfrak{g} \times Q$ with $q_0, q_T \in \mathscr{O}$ has the same solutions as the *restricted Clebsch optimal control problem* on $\mathfrak{g} \times \mathscr{O}$ given by

$$\min_{u(t)} \int_0^T \ell^{\mathscr{O}}(u(t), q(t))dt \tag{7.7}$$

subject to the following conditions:

(A) $\dot{q}(t) = u(t)_{\mathscr{O}}(q(t))$;

(B) $q(0) = q_0$ **and** $q(T) = q_T$.

In (7.7), the cost function $\ell^{\mathscr{O}} : \mathfrak{g} \times \mathscr{O} \to \mathbb{R}$ is defined by $\ell^{\mathscr{O}}(u,q) = \ell(u, i(q))$, where $i : \mathscr{O} \hookrightarrow Q$ is the inclusion, and $u_{\mathscr{O}}$ denotes the infinitesimal generator of the G-action on $\mathscr{O}$. We have the relation $Ti(u_{\mathscr{O}}(q)) = u_Q(i(q))$, for all $q \in \mathscr{O}$.

Quadratic cost functions and the normal metric:

Study the Clebsch optimal control problem in the special case where its cost function is given by the kinetic energy of a given inner product on the Lie algebra. We then show that the extremals are geodesics relative to an induced Riemannian metric on orbits. Let γ be the inner product on $\mathfrak{g}$ and consider

$$\ell(u,q) = \frac{1}{2}\gamma(u,u). \tag{7.8}$$

Defining the flat operator $\mathfrak{g} \ni u \mapsto u^\flat \in \mathfrak{g}^*$ by $u^\flat := \gamma(u,\_)$, we have the functional derivatives

$$\frac{\delta \ell}{\delta u} = u^\flat \quad \textbf{and} \quad \frac{\delta \ell}{\delta q} = 0.$$

The stationarity conditions (7.2) and the Euler-Poincaré equations (7.3) read

$$\dot{\alpha}_q = u_{T^*Q}(\alpha), \qquad u^\flat = \mathbf{J}(\alpha_q), \qquad \textbf{and} \qquad \frac{d}{dt}u^\flat = \mathrm{ad}^*_u u^\flat.$$

The Hamiltonian.:

Since the Lagrangian ℓ is hyperregular we can consider its associated Hamiltonian

$$h(\mu,q)=\frac{1}{2}\gamma(\mu^\sharp,\mu^\sharp),$$

where the sharp operator $\mathfrak{g}^*\ni\mu\mapsto\mu^\sharp\in\mathfrak{g}$ is defined as the inverse of the flat operator. The Hamiltonian $H:T^*Q\to\mathbb{R}$ is thus

$$H(\alpha_q)=\frac{1}{2}\gamma\left(\mathbf{J}(\alpha_q)^\sharp,\mathbf{J}(\alpha_q)^\sharp\right)=:\frac{1}{2}\kappa(q)(\alpha_q,\alpha_q),$$

where we defined the symmetric positive 2-contravariant tensor κ on Q by

$$\kappa(q)\left(\alpha_q,\beta_q\right):=\gamma\left(\mathbf{J}(\alpha_q)^\sharp,\mathbf{J}(\beta_q)^\sharp\right),\quad \textbf{for all } \alpha_q,\beta_q\in T^*Q.$$

We shall show that the tensor κ, and hence the Hamiltonian H, are closely related to a particular Riemannian metric on the G-orbits, called the normal metric.

The normal metric on orbits:

We now recall the definition of the normal metric on G-orbits. Given $q \in Q$, let $\mathfrak{g}_q := \{\xi \in \mathfrak{g} \mid \xi_Q(q) = 0\}$ denote the isotropy Lie algebra of q. Using the inner product γ on $\mathfrak{g}$, orthogonally decompose $\mathfrak{g} = \mathfrak{g}_q \oplus \mathfrak{g}_q^{\perp}$, and denote by $u = u_q + u^q$ the associated splitting of $u \in \mathfrak{g}$ in this direct sum. With these notations, the *normal metric* on a G-orbit $\mathscr{O}$ is defined by

$$\gamma_{\mathscr{O}}(q)(u_Q(q), v_Q(q)) := \gamma(u^q, v^q), \text{ for all } q \in Q \text{ and } u, v \in \mathfrak{g}. \tag{7.9}$$

Theorem 7.3. *Let G be a Lie group acting on the right on the smooth manifold Q and let γ be an inner product on $\mathfrak{g}$. Define the following symmetric positive 2-contravariant tensor on Q:*

$$\kappa(q)(\alpha_q, \beta_q) := \gamma\left(\mathbf{J}(\alpha_q)^\sharp, \mathbf{J}(\beta_q)^\sharp\right), \quad \alpha_q, \beta_q \in T^*Q. \tag{7.10}$$

Then:

- κ *is non-degenerate if and only if the G-action on Q is infinitesimally transitive.*
- κ *induces a well-defined co-metric $\kappa_{\mathscr{O}}$ on each G-orbit $\mathscr{O}$ of Q, through the following relation*

$$\kappa_{\mathscr{O}}(q)\left(T^*i\left(\alpha_{i(q)}\right), T^*i\left(\beta_{i(q)}\right)\right) = \kappa(i(q))\left(\alpha_{i(q)}, \beta_{i(q)}\right), \tag{7.11}$$

for $q \in \mathscr{O}$ *and* $\alpha_{i(q)}, \beta_{i(q)} \in T^*_{i(q)}Q$. *The co-metric* $\kappa_{\mathscr{O}}$ *is explicitly given by*

$$\kappa_{\mathscr{O}}(q)(\alpha_q, \beta_q) = \gamma\left(\mathbf{J}^{\mathscr{O}}(\alpha_q)^{\sharp}, \mathbf{J}^{\mathscr{O}}(\beta_q)^{\sharp}\right), \quad \textit{for} \quad \alpha_q, \beta_q \in T^*\mathscr{O}. \tag{7.12}$$

- $\kappa_{\mathscr{O}}$ *is the co-metric associated to the normal metric on* $\mathscr{O}$, *i.e.,*

$$\kappa(q)(\alpha_q, \beta_q) = \gamma_{\mathscr{O}}(q)(\alpha_q^{\sharp}, \beta_q^{\sharp}), \ \textit{for all} \ \ q \in \mathscr{O} \ \ \textit{and all} \ \ \alpha_q, \beta_q \in T_q^*\mathscr{O}, \tag{7.13}$$

where $T_q^*\mathscr{O} \ni \alpha_q \mapsto \alpha_q^{\sharp} \in T_q\mathscr{O}$ *is the sharp operator associated to* $\gamma_{\mathscr{O}}$.

As we have seen we can restrict the Clebsch optimal control problem to the G-orbit $\mathscr{O}$ containing q_0. In this case, by using Theorem 7.3, the collective Hamiltonian turns out to be the kinetic energy of the normal metric, i.e.,

$$H^{\mathscr{O}}(\alpha_q) = \frac{1}{2}\kappa_{\mathscr{O}}(q)(\alpha_q, \alpha_q) = \frac{1}{2}\gamma_{\mathscr{O}}(\alpha_q^\sharp, \alpha_q^\sharp). \tag{7.14}$$

We thus obtain the following instance of Theorem 7.1 which allows to interpret the solution $q(t)$ of the Clebsch optimal control problem for (7.8) as geodesics on G-obits.

Corollary 7.4 (Clebsch optimal control and geodesics of the normal metric)**.** *Let the Lie group G act on the right on Q, let γ be an inner product, suppose $q_0, q_T \in \mathscr{O}$, and consider the cost function $\ell(u,q) = \frac{1}{2}\gamma(u,u)$. Then:*

- *If $t \mapsto (u(t), q(t)) \in \mathfrak{g} \times \mathscr{O}$ is an extremal solution of the Clebsch optimal control problem (7.1), then there is a curve $t \mapsto \alpha(t) \in T^*\mathscr{O}$ covering $q(t)$, such that the following equations holds:*

$$u^\flat = \mathbf{J}(\alpha), \quad \dot{\alpha} = u_{T^*Q}(\alpha). \tag{7.15}$$

- *Equations (7.15) imply the Euler-Poincaré equations for the control u*

$$\frac{d}{dt}u^\flat = \mathrm{ad}^*_u u^\flat. \tag{7.16}$$

- *The second equation in (7.15), in which the first equation is used, is Hamiltonian on $T^*\mathscr{O}$ for the Hamiltonian (7.14). Therefore, $q(t)$ is a geodesic on $\mathscr{O}$ with respect to the normal metric $\gamma_{\mathscr{O}}$.*

The previous discussion can be easily adapted to the case with a potential, i.e.,

$$\ell(u,q) = \frac{1}{2}\gamma(u,u) - \mathscr{V}(q).$$

Equations (7.15) and (7.16) then become

$$u^\flat = \mathbf{J}(\alpha), \quad \dot{\alpha} = u_{T^*Q}(\alpha) - \mathrm{Ver}_\alpha \frac{\delta \mathscr{V}}{\delta q} \quad \textbf{and} \quad \frac{d}{dt}u^\flat = \mathrm{ad}^*_u u^\flat - \mathbf{J}\left(\frac{\delta \mathscr{V}}{\delta q}\right). \tag{7.17}$$

The Hamiltonian $H^{\mathscr{O}} : T^*\mathscr{O} \to \mathbb{R}$ takes the standard kinetic plus potential form

$$H^{\mathscr{O}}(\alpha_q) = \frac{1}{2}\gamma_{\mathscr{O}}(\alpha_q^\sharp, \alpha_q^\sharp) + \mathscr{V}(q).$$

Optimal control associated to geodesics:

Suppose that (Q,g) is a Riemannian manifold and consider the minimization of the Riemannian distance

$$\min \int_0^T \frac{1}{2}\|\dot{q}(t)\|^2 dt \tag{7.18}$$

subject to the condition $q(0)=q_0$ and $q(T)=q_T$. Suppose that there is a *transitive* action of the Lie group G on Q. Then this minimization problem can be reformulated as a Clebsch optimal control problem, namely,

$$\min_{u(t)} \int_0^T \frac{1}{2}\|u_Q(q)\|^2 dt \tag{7.19}$$

subject to the following conditions:

(A) $\dot{q}(t) = u(t)_Q(q(t))$**;**

(B) $q(0) = q_0$ **and** $q(T) = q_T$**.**

We can thus write the cost function as

$$\ell(u,q) = \frac{1}{2}\|u_Q(q)\|^2 = \frac{1}{2}\left\langle \mathbb{I}(q)u, u\right\rangle ,$$

where for each $q \in Q$, $\mathbb{I}(q)$ is the *locked inertia tensor* $\mathbb{I}(q) : \mathfrak{g} \to \mathfrak{g}^*$ defined by

$$\langle \mathbb{I}(q)u, v \rangle := g(q)(u_Q(q), v_Q(q)),$$

for any $u, v \in \mathfrak{g}$. The functional derivatives are

$$\frac{\delta \ell}{\delta u} = \mathbb{I}(q)u \in \mathfrak{g}_q^\circ \quad \textbf{and} \quad \frac{\delta \ell}{\delta q} = g(u_Q(q), \nabla u_Q(q)) = \frac{1}{2} \langle \mathbf{d}\mathbb{I}(q)(\cdot)u, u \rangle \in T_q^*Q,$$

where ∇ is the covariant derivative corresponding to the Riemannian metric. We note that $\ker(\mathbb{I}(q)) = \mathfrak{g}_q$ and $\operatorname{im}(\mathbb{I}(q)) = \mathfrak{g}_q^\circ$, therefore, ℓ is hyperregular if and only if the action is infinitesimally free, i.e., $\mathfrak{g}_q = \{0\}$.

In the hyperregular case, we obtain the Hamiltonian $h : \mathfrak{g}^* \times Q \to \mathbb{R}$, given by

$$h(\mu, q) = \frac{1}{2} \left\langle \mu, \mathbb{I}(q)^{-1} \mu \right\rangle, \tag{7.20}$$

and the Hamiltonian $H : T^*Q \to \mathbb{R}$ reads

$$H(\alpha_q) = h(\mathbf{J}(\alpha_q), q) = \frac{1}{2} \left\langle \mathbf{J}(\alpha_q), \mathbb{I}(q)^{-1} \mathbf{J}(\alpha_q) \right\rangle.$$

Can extend to the nonregular case.

8 Optimal control on Stiefel manifolds

An optimal control problem on Stiefel manifolds is introduced and studied in Bloch, Crouch, Sanyal, as a generalization of the geodesic flow on the sphere (case $n = 1$) and the motion of the free N-dimensional rigid body (case $n = N$). In Gay-Balmaz and Ratiu this problem was generalized to arbitrary Lagrangians and formulated as a Clebsch optimal control problem of the form (7.1).

Here we show that the Clebsch optimal control problem on Stiefel manifolds offers a unified point of view for the formulation of several integrable systems. These systems turn out to be associated to two classes of cost functions. From this setting, we also deduce a geodesic interpretation of the solution of some of these integrable systems.

Stiefel manifolds:

For $n \leq N$, define the *Stiefel manifold* $V_n(\mathbb{R}^N)$ to be the set of orthonormal n-frames in $\mathbb{R}^N$ (i.e., an ordered set of n orthonormal vectors). So, $V_n(\mathbb{R}^N)$ is the set of linear isometric embeddings of $\mathbb{R}^n$ into $\mathbb{R}^N$. Let S^{N-1} denoted the unit sphere in $\mathbb{R}^N$. Since $V_n(\mathbb{R}^N) \subset \left(S^{N-1}\right)^n$ is closed, it follows that $V_n(\mathbb{R}^N)$ is compact. Collect the n vectors of an orthonormal frame in $\mathbb{R}^N$ as columns of a $N \times n$ matrix $Q \in V_n(\mathbb{R}^N)$. If $\mathrm{Mat}(N \times n)$ denotes the vector space of matrices having N rows and n columns, then the Stiefel manifold can be described as

$$V_n(\mathbb{R}^N) = \{Q \in \mathrm{Mat}(N \times n) \mid Q^\mathsf{T} Q = I_n\}, \tag{8.1}$$

where I_n is the $n \times n$ identity matrix. The dimension of $V_n(\mathbb{R}^N)$ is $Nn - (n+1)n/2$.

The characterization (8.1) of $V_n(\mathbb{R}^N)$ immediately shows that if $n = 1$, then $V_1(\mathbb{R}^N) = S^{N-1}$ and if $n = N$, then $V_N(\mathbb{R}^N) = O(N)$, the group of orthogonal isomorphisms of $\mathbb{R}^N$.

The tangent space at $Q \in V_n(\mathbb{R}^N)$ to the Stiefel manifold $V_n(\mathbb{R}^N)$ is given by

$$T_Q V_n(\mathbb{R}^N) = \{V \in \mathrm{Mat}(N \times n) \mid V^\mathsf{T} Q + Q^\mathsf{T} V = 0\}. \tag{8.2}$$

We identify $T^* V_n(\mathbb{R}^N)$ with $TV_n(\mathbb{R}^N)$ using the pairing $T_Q^* V_n(\mathbb{R}^N) \times T_Q V_n(\mathbb{R}^N) \ni (P_Q, V_Q) \mapsto \mathrm{Trace}\left(P_Q^\mathsf{T} V_Q\right) \in \mathbb{R}$ for every $Q \in V_n(\mathbb{R}^N)$.

Remark 8.1. *It is also known that*

$$V_n(\mathbb{R}^N) = SO(N)/SO(N-n) \to SO(N)/(SO(n) \times SO(N-n)) =: \widetilde{Gr}_n(\mathbb{R}^N)$$

is a principal $SO(n)$-bundle, where $\widetilde{Gr}_n(\mathbb{R}^N)$ is the Grassmannian of oriented n-planes in $\mathbb{R}^N$. The notation $Gr_n(\mathbb{R}^N)$ is reserved for the Grassmannian of n-planes in $\mathbb{R}^N$ (regardless of orientation).

Note that for $N > 1$, $V_1(\mathbb{R}^N) = S^{N-1} = \widetilde{Gr}_1(\mathbb{R}^N)$, while $Gr_1(\mathbb{R}^N) = \mathbb{RP}^{N-1}$. $\Diamond$

8.1 Clebsch optimal control on Stiefel manifolds

From now on, we assume that $n<N$. We consider the *right* $SO(N)$-action on $V_n(\mathbb{R}^N)$ given by $Q\mapsto R^{-1}Q$ for $R\in SO(N)$. The infinitesimal generator of this action is $U_{V_n(\mathbb{R}^N)}(Q)=-UQ\in T_QV_n(\mathbb{R}^N)$, $U\in\mathfrak{so}(N)$.

Given $Q_0,Q_T\in V_n(\mathbb{R}^N)$, the Clebsch optimal control problem (7.1) reads

$$\min_{U(t)}\int_0^T \ell(U(t),Q(t))dt \tag{8.3}$$

subject to the following conditions:

(A) $\dot{Q}(t)=-U(t)Q(t)$**;**

(B) $Q(0)=Q_0$ **and** $Q(T)=Q_T$**.**

We identify the dual $\mathfrak{so}(N)^*$ with itself using the non-degenerate pairing $\mathfrak{so}(N)\times\mathfrak{so}(N)\ni(U_1,U_2)\mapsto\mathrm{Trace}(U_1^\mathsf{T}U_2)\in\mathbb{R}$. The cotangent bundle momentum map $\mathbf{J}:T^*V_n(\mathbb{R}^N)\to\mathfrak{so}(N)^*$ is easily verified to be

$$\mathbf{J}(Q,P)=\frac{1}{2}\left(QP^\mathsf{T}-PQ^\mathsf{T}\right).$$

The optimal control is thus given by $\delta\ell/\delta U = \left(QP^\mathsf{T} - PQ^\mathsf{T}\right)/2$ (see (7.2)). The cotangent lifted action on $T^*V_n(\mathbb{R}^N)$ reads $(Q,P) \mapsto (R^\mathsf{T}Q, R^\mathsf{T}P)$ and hence Hamilton's equations (7.2) become

$$\dot{Q} = -UQ, \quad \dot{P} = -UP + \frac{\delta\ell}{\delta Q}, \tag{8.4}$$

in this particular case. Recall that here $\delta\ell/\delta Q \in T^*_Q V_n(\mathbb{R}^N)$ denotes the functional derivative of ℓ relative to the above defined pairing. The optimal control U, given algebraically by $\delta\ell/\delta U = \left(QP^\mathsf{T} - PQ^\mathsf{T}\right)/2$, is necessarily the solution of the Euler-Poincaré equation (7.3) given in this particular case by

$$\frac{d}{dt}\frac{\delta\ell}{\delta U} = \left[\frac{\delta\ell}{\delta U}, U\right] + \frac{1}{2}\left(Q\left(\frac{\delta\ell}{\delta Q}\right)^\mathsf{T} - \frac{\delta\ell}{\delta Q}Q^\mathsf{T}\right). \tag{8.5}$$

Example 1: N-dimensional free rigid body:

We consider as a cost function the free rigid body Lagrangian $\ell(U) = \frac{1}{2}\langle U, J(U)\rangle$, where $J(U) = \Lambda U + U\Lambda$, $\Lambda = \mathrm{diag}(\Lambda_1, \ldots \Lambda_N)$, $\Lambda_i + \Lambda_j > 0$ for $i \neq j$. The corresponding Clebsch optimal control falls into the setting discussed earlier.

Since

$$M := \frac{\delta \ell}{\delta U} = J(U), \quad \frac{\delta \ell}{\delta Q} = 0,$$

equations (8.4) and (8.5) become

$$\dot{Q} = -UQ, \qquad \dot{P} = -UP \tag{8.6}$$

and

$$\dot{M} = [M, U], \quad \textbf{where} \quad M = J(U) = \frac{1}{2}\left(QP^{\mathsf{T}} - PQ^{\mathsf{T}}\right).$$

From Corollary 7.4, the solution $Q(t)$ is a geodesic on the $SO(N)$-orbit of $Q(0)$ in $V_n(\mathbb{R}^N)$, relative to the normal metric induced by the inner product $\gamma(U,V) := \langle U, J(V)\rangle$ on this orbit.

The free rigid body equations $\dot{M} = [M,U]$, $M = J(U) = \Lambda U + U\Lambda$ on $\mathfrak{so}(N)$, and indeed their generalization on any semisimple Lie algebra, are integrable, see Mischenko and Fomenko (1976). A key observation in this regard, pointed out for the first time in Manakov (1976), was that one can write the generalized rigid body equations as a Lax equation with parameter:

$$\frac{d}{dt}(M + \lambda\Lambda^2) = [M + \lambda\Lambda^2, U + \lambda\Lambda]. \tag{8.7}$$

The nontrivial 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 the generic coadjoint orbits of $SO(n)$ (see also Ratiu 1980).

Equation (8.7) is of the form $\dot{L} = [L,B]$ with L expressed in terms of the canonical variables as $L(Q,P) = \frac{1}{2}\left(QP^\mathsf{T} - PQ^\mathsf{T}\right) + \Lambda^2\lambda$.

Example 1A: symmetric representation of the N-dimensional free rigid body. Consider the special case $n = N-1$, i.e. $V_{N-1}(\mathbb{R}^N) = SO(N)$. Note that if the initial condition $P(0) \in SO(N)$, then the solution $(Q(t), P(t))$ of (8.6) preserves $SO(N) \times SO(N)$. Since in this case the formulation (8.6) of the free rigid body equation is symmetric in Q and P, it is called the *symmetric representation of the rigid body on* $SO(N) \times SO(N)$. As before, if (Q, P) is a solution of (8.6), then (Q, M), where $M = J(U)$ and $U = -\dot{Q}Q^{-1}$, satisfies the rigid body equations $\dot{Q} = -UQ$, $\dot{M} = [M, U]$ (see Bloch, Crouch, Marsden and Ratiu for a study of this system and its discretization).

Example 1B: $n = 1$**, the rank 2 free rigid body.** We compute equations (8.4) and (8.5) for the case $n = 1$, i.e., $V_1(\mathbb{R}^N) = S^{N-1}$. For $(\mathbf{q},\mathbf{p}) \in T^*S^{N-1}$ we get

$$\dot{\mathbf{q}} = -U\mathbf{q}, \qquad \dot{\mathbf{p}} = -U\mathbf{p}$$

and

$$\dot{M} = [M,U], \quad \text{where} \quad M = J(U) = \frac{1}{2}\left(\mathbf{q}\otimes\mathbf{p} - \mathbf{p}\otimes\mathbf{q}\right). \tag{8.8}$$

Note that, generically, M has rank 2. Associated to the Manakov equation (8.7), Moser introduces the Lax pair matrices L and B given by

$$\begin{aligned} L(\mathbf{q},\mathbf{p}) &= \Lambda^2 + a\mathbf{q}\otimes\mathbf{q} + b\mathbf{q}\otimes\mathbf{p} + c\mathbf{p}\otimes\mathbf{q} + d\mathbf{p}\otimes\mathbf{p}, \\ B(\mathbf{q},\mathbf{p}) &= J^{-1}(\mathbf{q}\otimes\mathbf{p} - \mathbf{p}\otimes\mathbf{q}) + \lambda\Lambda, \end{aligned} \tag{8.9}$$

with $a = d = 0$, $b = -c = \frac{1}{2\lambda}$. For these values of the parameters, the expression of the matrix L in (8.9) is reminiscent of the expression of the momentum map (8.8) arising from the Clebsch optimal control formulation. We have

$$J^{-1}(\mathbf{q}\otimes\mathbf{p} - \mathbf{p}\otimes\mathbf{q}) = \frac{q_i p_j - q_j p_i}{2(\Lambda_i + \Lambda_j)}.$$

Recall that the equations for the rank 2 free rigid body arise from an optimal control problem on S^{N-1}, rather than on the orthogonal group: minimize $\frac{1}{2}\int_0^T \langle U, J(U)\rangle\, dt$, where U is a skew symmetric control, subject to $\dot{\mathbf{q}} = -U\mathbf{q}$ as in (7.1).

From the result of Corollary 7.4, the curve $\mathbf{q}(t) \in S^{N-1}$ (there is only one orbit for $n = 1$) is a geodesic on S^{N-1} relative to the normal metric induced from the inner product $\gamma(U,V) := \langle U, J(V)\rangle$.

Example 2:
We consider as a cost function the expression

$$\ell(U,Q) = \frac{1}{2}\langle \Lambda UQ, UQ\rangle - \mathscr{V}(Q), \tag{8.10}$$

where Λ is a given symmetric positive definite $N \times N$ matrix and $\mathscr{V} \in C^\infty(V_n(\mathbb{R}^N))$. The case $\mathscr{V} = 0$ is the geodesic problem. The first term in (8.10) is the kinetic energy associated to the Riemannian metric g on $V_n(\mathbb{R}^N)$ defined by

$$g_Q(V,W) = \langle \Lambda V, W\rangle = \mathrm{Tr}(V^\top \Lambda W) \qquad V,W \in T_Q V_n(\mathbb{R}^N).$$

In each of the examples mentioned below, the Clebsch optimal control formulation allows us to efficiently derive the explicit form of geodesic equations; see (8.14), (8.18). This approach also yields a natural setting for generalizing certain integrable systems from the sphere to the Stiefel manifold, such as the C. Neumann problem.

Hamilton's equations (7.2) become in this case

$$\dot{Q} = -UQ, \qquad \dot{P} = -UP + [QQ^\mathsf{T}, U\Lambda U]Q - \frac{\delta \mathscr{V}}{\delta Q}. \tag{8.11}$$

The corresponding Euler-Poincaré equations (8.5) are

$$\dot{M} = [M, U] + \frac{1}{2}[U\Lambda U, QQ^\mathsf{T}] - \frac{1}{2}\left(Q \left(\frac{\delta \mathscr{V}}{\delta Q} \right)^\mathsf{T} - \frac{\delta \mathscr{V}}{\delta Q} Q^\mathsf{T} \right), \tag{8.12}$$

where $M = \frac{\delta \ell}{\delta U} = \frac{1}{2}\left(QQ^\mathsf{T} U\Lambda + \Lambda U QQ^\mathsf{T}\right) = \frac{1}{2}\left(QP^\mathsf{T} - PQ^\mathsf{T}\right)$**.**

Example 2A: $n = 1$, $\mathscr{V} = 0$, geodesics on the ellipsoid. Let us consider the case $n = 1$, i.e., $V_1(\mathbb{R}^N) = S^{N-1}$. The geodesic flow on S^{N-1} for the metric $g(q)(u,v) := \langle u, \Lambda v\rangle$, for $q \in S^{N-1}$, $u, v \in T_q S^{N-1}$ is equivalent to the geodesic flow on the ellipsoid $\bar{q}^\mathsf{T} \Lambda^{-1} \bar{q} = 1$, with $q = \Lambda^{-1/2}\bar{q}$. Equations (8.11) and (8.12) yield

$$\dot{q} = -Uq, \quad \dot{p} = -Up + [qq^\mathsf{T}, U\Lambda U]q, \quad \dot{M} = [M,U] + \frac{1}{2}[U\Lambda U, qq^\mathsf{T}] \tag{8.13}$$

where $M = \delta\ell/\delta U = \frac{1}{2}\left(qq^\mathsf{T} U\Lambda + \Lambda U qq^\mathsf{T}\right) = \frac{1}{2}\left(qp^\mathsf{T} - pq^\mathsf{T}\right)$.

We now deduce from (8.13) the geodesic equations for the ellipsoid (see Theorem 7.1, (7.2)). Using the equality $M = \frac{1}{2}\left(qq^\mathsf{T} U\Lambda + \Lambda U qq^\mathsf{T}\right)$, we get

$$\dot{M}\Lambda^{-1}q = \frac{1}{2}\left(qq^\mathsf{T}U^2q + \Lambda\dot{U}q(q^\mathsf{T}\Lambda^{-1}q) - \Lambda U^2 q(q^\mathsf{T}\Lambda^{-1}q) + \Lambda Uq(q^\mathsf{T}U\Lambda^{-1}q)\right)$$

from where we solve for $\dot{U}q$, which inserted in $\ddot{q} = -\dot{U}q + U^2 q$ yields

$$\ddot{q} = \left(-2\Lambda^{-1}\dot{M}\Lambda^{-1}q + \Lambda^{-1}qq^\mathsf{T}U^2q + Uqq^\mathsf{T}U\Lambda^{-1}q\right)\left(q^\mathsf{T}\Lambda^{-1}q\right)^{-1}.$$

Now, we replace in this formula $\dot{M}$ by its expression in (8.13) and we get the geodesic equations

$$\ddot{q} = -\frac{|\dot{q}|^2}{q^{\mathsf{T}}\Lambda^{-1}q}\Lambda^{-1}q. \tag{8.14}$$

The geodesic equations on the triaxial ellipsoid were solved by Jacobi. The complete solution is found in his course notes.

Remark 8.2. *As a particular case, the geodesic equations on the sphere ($\Lambda = I_N$), are $\ddot{q} = -|\dot{q}|^2 q$.* $\Diamond$

Example 2B: $\mathscr{V}=0$**, geodesics on the Stiefel manifolds.** When $\mathscr{V}=0$, (8.11) and (8.12) yield

$$\dot{Q}=-UQ, \quad \dot{P}=-UP+[QQ^{\mathsf{T}},U\Lambda U]Q, \quad \dot{M}=[M,U]+\frac{1}{2}[U\Lambda U,QQ^{\mathsf{T}}], \tag{8.15}$$

where $M=\delta\ell/\delta U=\frac{1}{2}\left(QQ^{\mathsf{T}}U\Lambda+\Lambda UQQ^{\mathsf{T}}\right)=\frac{1}{2}\left(QP^{\mathsf{T}}-PQ^{\mathsf{T}}\right)$.

We now deduce from (8.15) the geodesic equations for the Stiefel manifolds (see Theorem 7.1, (7.2)). A direct computation yields

$$\begin{aligned}\dot{M}\Lambda^{-1}Q=\frac{1}{2}\big(&-UQQ^{\mathsf{T}}UQ+QQ^{\mathsf{T}}U^2Q+L(\dot{U}Q)-\Lambda U^2Q(Q^{\mathsf{T}}\Lambda^{-1}Q)\\&+\Lambda UQ(Q^{\mathsf{T}}U\Lambda^{-1}Q)\big),\end{aligned} \tag{8.16}$$

where the linear operator L on the vector space of $N\times n$ matrices is defined by

$$L(X):=QQ^{\mathsf{T}}X+\Lambda XQ^{\mathsf{T}}\Lambda^{-1}Q. \tag{8.17}$$

Note that if $\Lambda=I_N$, then $L(X)=(I_N+QQ^{\mathsf{T}})X$.

We study the properties of the operator $L : \mathrm{Mat}(N\times n) \to \mathrm{Mat}(N\times n)$, where $\mathrm{Mat}(N\times n)$ denotes the real vector space of matrices with N rows and n columns. Recall that $\mathrm{Mat}(N\times n)$ has the natural inner product $\langle\langle A,B\rangle\rangle := \mathrm{Tr}\left(A^\mathsf{T}B\right)$. A direct computation shows that L is a linear symmetric operator relative to the inner product:

$$\langle\langle L(X),Y\rangle\rangle = \langle\langle X,L(Y)\rangle\rangle = \mathrm{Tr}\left(X^\mathsf{T}QQ^\mathsf{T}Y\right) + \mathrm{Tr}\left(X^\mathsf{T}\Lambda Y Q^\mathsf{T}\Lambda^{-1}Q\right).$$

In particular,

$$\langle\langle L(X),X\rangle\rangle = \left\langle\!\!\left\langle Q^\mathsf{T}X, Q^\mathsf{T}X\right\rangle\!\!\right\rangle + \mathrm{Tr}\left(X^\mathsf{T}\Lambda X Q^\mathsf{T}\Lambda^{-1}Q\right).$$

Note that $Q^\mathsf{T}\Lambda^{-1}Q$ is a symmetric positive definite matrix because Λ is a symmetric positive definite matrix and $Q\in V_n(\mathbb{R}^N)$. Therefore, there is a symmetric positive definite $n\times n$ matrix R such that $R^2 = Q^\mathsf{T}\Lambda^{-1}Q$. Hence the previous expression becomes

$$\langle\langle L(X),X\rangle\rangle = \left\langle\!\!\left\langle Q^\mathsf{T}X, Q^\mathsf{T}X\right\rangle\!\!\right\rangle + \mathrm{Tr}\left((XR)^\mathsf{T}\Lambda(XR)\right)$$

and we note that each summand is ≥ 0. Hence $\langle\langle L(X),X\rangle\rangle = 0 \Rightarrow \mathrm{Tr}\left((XR)^\mathsf{T}\Lambda(XR)\right) = 0$. Since Λ is positive definite, we conclude that $XR = 0$ which implies that $X = 0$ because R is invertible. We conclude that $L : \mathrm{Mat}(N\times n) \to \mathrm{Mat}(N\times n)$ is a symmetric positive definite operator and hence invertible.

Returning to (8.16), we isolate $\left(\dot{U}Q\right)$, replace in this formula $\dot{M}$ by (8.15), and we get

$$L\left(\ddot{Q}\right) = L\left(-\dot{U}Q + U^2Q\right) = 2QQ^\mathsf{T}U^2Q \overset{(8.15)}{=} -2Q\dot{Q}^\mathsf{T}\dot{Q}, \tag{8.18}$$

which are the geodesic equations on the Stiefel manifold.

Remark 8.3. *When $n = 1$, (8.18) coincide with (8.14). Indeed, in this case (8.18) becomes*

$$q\left(q^\mathsf{T}\ddot{q}\right) + \Lambda\ddot{q}\left(q^\mathsf{T}\Lambda^{-1}q\right) = -2q\dot{q}^\mathsf{T}\dot{q}, \quad q \in S^{N-1}.$$

Since $q^\mathsf{T}q = 1$ we have $q\ddot{q}^\mathsf{T} + \dot{q}^\mathsf{T}\dot{q} = 0$, which then implies the geodesic equations on the ellipsoid (8.14). $\Diamond$

Example 2C: $n=1$, $\Lambda = I_N$, $\mathscr{V}(q) = \frac{1}{2}Aq\cdot q$, $A := \mathrm{diag}(a_1,\ldots,a_N)$, **the C. Neumann problem. We now study the motion of a point on the sphere** S^{N-1} **under the influence of the quadratic potential** $\frac{1}{2}Aq\cdot q$**. For** $N=3$ **the associated Hamilton equations were shown to be completely integrable by Carl Neumann ; for general** N **and a study of various geometric and dynamic aspects of this problem see Uhlenbeck, , Moser, Adler and Van Moerbeke and Ratiu.**

Since $\dot{q} = -Uq$**, the Lagrangian of this system is**

$$\ell(U,q) = \frac{1}{2}\dot{q}^{\mathsf{T}}\dot{q} - \frac{1}{2}q^{\mathsf{T}}Aq = -\frac{1}{2}q^{\mathsf{T}}\left(U^2+A\right)q \tag{8.19}$$

and hence

$$\frac{\delta\ell}{\delta U} = \frac{1}{2}\left(qq^{\mathsf{T}}U + Uqq^{\mathsf{T}}\right), \qquad \frac{\delta\ell}{\delta q} = -(U^2+A)q + q(q^{\mathsf{T}}(U^2+A)q).$$

Since $M := \frac{\delta\ell}{\delta U}$**, (8.5) implies**

$$\dot{M} = [M,U] + \frac{1}{2}[U^2+A, qq^{\mathsf{T}}]. \tag{8.20}$$

On the other hand, using the definition of M, we get $\dot{M}q = \frac{1}{2}(qq^\mathsf{T}U^2q + \dot{U}q - U^2q)$ which yields the equations of motion for the Neumann system

$$\ddot{q} = -2\dot{M}q + qq^\mathsf{T}U^2q \overset{(8.20)}{=} -Aq + \left(Aq\cdot q - |\dot{q}|^2\right)q. \tag{8.21}$$

Example 2D: $\Lambda = I_N$, $\mathscr{V}(Q) = \frac{1}{2}\langle\langle AQ, Q\rangle\rangle$, $A := \operatorname{diag}(a_1, \ldots, a_N)$, the C. Neumann problem on Stiefel manifolds. We now consider the motion of a point on the Stiefel manifold $V_n(\mathbb{R}^N)$ under the influence of the quadratic potential $\mathscr{V}(Q) = \frac{1}{2}\langle\langle AQ, Q\rangle\rangle$, where we can assume, without loss of generality, that $A = \operatorname{diag}(a_1, \ldots, a_N)$. We work in the generic case when $a_i \neq 0$ for all $i = 1, \ldots, N$.

Since $\dot{Q} = -UQ$, the Lagrangian of this system is

$$\ell(U, Q) = \frac{1}{2}\operatorname{Tr}\left(\dot{Q}^{\mathsf{T}}\dot{Q}\right) - \frac{1}{2}\operatorname{Tr}\left(Q^{\mathsf{T}}AQ\right) = -\frac{1}{2}\operatorname{Tr}\left(Q^{\mathsf{T}}\left(U^2 + A\right)Q\right) \tag{8.22}$$

and hence

$$\frac{\delta\ell}{\delta U} = \frac{1}{2}\left(QQ^{\mathsf{T}}U + UQQ^{\mathsf{T}}\right), \qquad \frac{\delta\ell}{\delta Q} = -(U^2 + A)Q + Q(Q^{\mathsf{T}}(U^2 + A)Q).$$

Since $M := \frac{\delta\ell}{\delta U}$, (8.5) implies

$$\dot{M} = [M, U] + \frac{1}{2}[U^2 + A, QQ^{\mathsf{T}}]. \tag{8.23}$$

On the other hand, using the definition of M, we get

$$\dot{M}Q = \frac{1}{2}\left(QQ^{\mathsf{T}}U^2Q + L(\dot{U}Q) - U^2Q\right), \tag{8.24}$$

where $L(X) := (I_N + QQ^{\mathsf{T}})X$, for $X \in \mathrm{Mat}(N \times n)$ (see (8.17)). Using (8.24), (8.23), and $2M = QQ^{\mathsf{T}}U + UQQ^{\mathsf{T}}$, we get $L\left(\ddot{Q}\right) = L\left(-\dot{U}Q + U^2Q\right) = -2Q\dot{Q}^{\mathsf{T}}\dot{Q} - AQ + QQ^{\mathsf{T}}AQ$, which yield the equations of motion for the Neumann system on $V_n(\mathbb{R}^N)$

$$\ddot{Q} = (I_N + QQ^{\mathsf{T}})^{-1}\left(-2Q\dot{Q}^{\mathsf{T}}\dot{Q} - AQ + QQ^{\mathsf{T}}AQ\right). \tag{8.25}$$

These equations for $A = 0$ coincide with (8.18) and for $n = 1$ with (8.21).

9 Clebsch optimal control formulation for the flows on symmetric matrices

Given $N \in \mathfrak{so}(n)$, the system (Bloch Iserles (2006), Bloch, Brinzenesco, Iserles, Marsden and Ratiu (2009) is the ordinary differential equation on the space $\mathrm{sym}(n)$ of $n \times n$ symmetric matrices given by

$$\dot{X} = [X^2, N], \quad X(t) \in \mathrm{sym}(n). \tag{9.1}$$

Assume that N is invertible, $n = 2k$, and consider the symplectic group

$$Sp(2k, N^{-1}) := \left\{ Q \in GL(2k, \mathbb{R}) \mid Q^\mathsf{T} N^{-1} Q = N^{-1} \right\} \tag{9.2}$$

with Lie algebra $\mathfrak{sp}(2k, N^{-1}) = \{U \in \mathfrak{gl}(2k) \mid U^\mathsf{T} N^{-1} + N^{-1} U = 0\}$. The system (9.1) can be written as the Euler-Poincaré equation on $\mathfrak{sp}(2k, N^{-1})$ for the Lagrangian

$$\ell(U) = \frac{1}{2} \mathrm{Tr}((N^{-1}U)^2). \tag{9.3}$$

Indeed, using the identification $\mathfrak{sp}(2k, N^{-1})^* := \operatorname{sym}(2k)$ with duality pairing $\langle\langle X, U \rangle\rangle = \operatorname{Tr}(XN^{-1}U)$ for $X \in \operatorname{sym}(2k)$ and $U \in \mathfrak{sp}(2k, N^{-1})$, we have $\delta\ell/\delta U = N^{-1}U$ and $\operatorname{ad}^*_U X = XN^{-1}UN - UX$, so the Euler-Poincaré equation $\frac{d}{dt}\frac{\delta\ell}{\delta U} = \operatorname{ad}^*_U \frac{\delta\ell}{\delta U}$ becomes $N^{-1}\dot{U} = N^{-1}UN^{-1}UN - UN^{-1}U$. Setting $X := N^{-1}U$, we recover (9.1). As a consequence, (9.1) describes left invariant geodesics on the Lie group (9.2).

When N is not invertible, then (9.1) describes left invariant geodesics on the Jacobi group and its generalizations; see Gay-Balmaz and Tronci.

9.1 Clebsch optimal control formulation

Assume that N is invertible and consider the right action of the group $\mathrm{Sp}(2k,N^{-1})$ by multiplication on $\mathrm{GL}(2k,\mathbb{R})$. Consider the cost function $\ell : \mathfrak{sp}(2k,N^{-1}) \to \mathbb{R}$ given in (9.3). The associated Clebsch optimal control problem is

$$\min \int_0^T \ell(U)dt, \quad \textbf{subject to} \quad \dot{Q} = QU,\ Q(0) = Q_0,\ Q(T) = Q_T.$$

Conditions (7.2) read

$$\frac{\delta \ell}{\delta U} = \mathbf{J}(Q,P) = \frac{1}{2}(P^\top QN + (QN)^\top P), \quad \dot{Q} = QU, \quad \dot{P} = -PU^\top, \tag{9.4}$$

with respect to the duality pairing $\langle P,V\rangle := \mathrm{Tr}(P^\top V)$, for $V \in T_Q GL(2k,\mathbb{R})$ and $P \in T^*GL(2k,\mathbb{R})$.

By Theorem 7.1, if Q, P satisfy the last two equations in (9.4), then $X = \frac{\delta \ell}{\delta U}$ verifies the equations (9.1). Let's check this directly. We compute

$$
\begin{aligned}
2\dot{X} = 2N^{-1}\dot{U} &= \frac{d}{dt}\left(P^{\mathsf{T}}QN - NQ^{\mathsf{T}}P\right) \\
&\overset{(9.4)}{=} -UP^{\mathsf{T}}QN + P^{\mathsf{T}}QUN - NU^{\mathsf{T}}Q^{\mathsf{T}}P + NQ^{\mathsf{T}}PU^{\mathsf{T}} \\
&= -U\left(P^{\mathsf{T}}QN - NQ^{\mathsf{T}}P\right) + \left(P^{\mathsf{T}}QN - NQ^{\mathsf{T}}P\right)N^{-1}UN \\
&= 2\left(XN^{-1}UN - UX\right) = 2\left[X^2, N\right]
\end{aligned}
$$

since $U = NX$, as stated.

This approach generalizes to the right action of $\mathrm{Sp}(2k, N^{-1})$ on $\mathfrak{gl}(2k, \mathbb{R})$ or, more generally, on the space $\mathrm{Mat}(n \times 2k)$ of rectangular $n \times 2k$ matrices.

Note that (9.1) is equivalent to the following Lax equation with parameter

$$
\frac{d}{dt}(X + \lambda N) = \left[X + \lambda N, NX + XN + \lambda N^2\right]. \tag{9.5}
$$

In this case, the Lax equation with parameter $\dot{L} = [L, B]$ has $L(Q,P) := P^\top QN - NQ^\top P + N\lambda$. For example, if $n = 1$, i.e., $\mathbf{q} \in \mathbb{R}^{2k}$ (seen as a row), then we have

$$\frac{\delta \ell}{\delta U} = \frac{1}{2}(\mathbf{p} \otimes \mathbf{q}N + \mathbf{q}N \otimes \mathbf{p})$$

9.2 Symmetric representation of the flow on symmetric matrices

Since $U \in \mathfrak{sp}(2k, N^{-1})$, the last two equations in system (9.4) are equivalent to

$$\dot{Q} = QU, \qquad \dot{P}N^{-1} = (PN^{-1})U,$$

which shows that if $U \in \mathfrak{sp}(2k, N^{-1})$ and the initial conditions $(Q(0), P(0)N^{-1}) \in Sp(2k, N^{-1}) \times Sp(2k, N^{-1})$, then $(Q(t), P(t)N^{-1}) \in Sp(2k, N^{-1}) \times Sp(2k, N^{-1})$.

Since $\delta\ell/\delta U = X = N^{-1}U$, the Hamiltonian $h : \mathrm{sym}(2k) \to \mathbb{R}$ has the expression

$$h(X) := \langle\langle X, U \rangle\rangle - \ell(U) = \frac{1}{2}\mathrm{Tr}(X^2).$$

Therefore, using (9.4), the Hamiltonian $H : T^*\mathfrak{gl}(2k, \mathbb{R}) \to \mathbb{R}$ is

$$H(Q, P) := h(\mathbf{J}(Q, P)) = \frac{1}{8}\mathrm{Tr}\left(\left(P^\top QN - NQ^\top P\right)^2\right). \tag{9.6}$$

By Theorem 7.1, we get the following result.

Proposition 9.1. *Consider the canonical Hamiltonian system on $T^*\mathfrak{gl}(2k,\mathbb{R})$ with the symplectic structure*

$$\Omega_{\text{can}}((Q_1,P_1),(Q_2,P_2)) = \operatorname{Tr}(P_2^\top Q_1 - P_1^\top Q_2) \tag{9.7}$$

and Hamiltonian (9.6). Then its solutions are mapped by $\mathbf{J} : T^*\mathfrak{gl}(2k,\mathbb{R}) \to \operatorname{sym}(2k)$ to integral curves of the system (9.1). The flow generated by (9.6) preserves the submanifold $\left\{(Q,P) \in \mathfrak{gl}(2k,\mathbb{R}) \times \mathfrak{gl}(2k,\mathbb{R}) \mid Q, PN^{-1} \in Sp(2k,N^{-1})\right\}$.

10 Clebsch optimal control formulation for the finite Toda lattice

Consider a complex semisimple Lie algebra $\mathfrak{g}^{\mathbb{C}}$, its split normal real form $\mathfrak{g}$, and the decomposition $\mathfrak{g} = \mathfrak{b}_- \oplus \mathfrak{k}$, where $\mathfrak{k}$ is the compact normal Lie algebra and $\mathfrak{b}_-$ a Borel Lie subalgebra – see Bloch, Gay-Balmaz and Ratiu.

Let us quickly recall how the full Toda equation can be viewed as the Euler-Poincaré equation on the Lie algebra $\mathfrak{b}_-$ for the Lagrangian $\ell(U) = \frac{1}{2}\kappa(U,U)$, with κ the Killing form. If we identify the dual Lie algebra as $(\mathfrak{b}_-)^* = \mathfrak{k}^\perp$ by using κ, we have $\delta\ell/\delta U = \pi_{\mathfrak{k}^\perp}(U)$ and $\mathrm{ad}^*_U \mu = -\pi_{\mathfrak{k}^\perp}([U,\mu])$, so the Euler-Poincaré equation reads

$$\pi_{\mathfrak{k}^\perp}(\dot{U}) = -\pi_{\mathfrak{k}^\perp}([U, \pi_{\mathfrak{k}^\perp}(U)]) \tag{10.1}$$

Note that $(\pi_{\mathfrak{b}_-})|_{\mathfrak{k}^\perp} : \mathfrak{k}^\perp \to \mathfrak{b}_-$ is an isomorphism with inverse $(\pi_{\mathfrak{k}^\perp})|_{\mathfrak{b}_-} : \mathfrak{b}_- \to \mathfrak{k}^\perp$.

We can rewrite the right hand side as

$$\begin{aligned}\pi_{\mathfrak{k}^\perp}\left([U,\pi_{\mathfrak{k}^\perp}(U)]\right) &= \pi_{\mathfrak{k}^\perp}\left([\pi_{\mathfrak{b}_-}\pi_{\mathfrak{k}^\perp}(U),\pi_{\mathfrak{k}^\perp}(U)]\right)\\ &= \pi_{\mathfrak{k}^\perp}\left([\pi_{\mathfrak{k}^\perp}(U)-\pi_{\mathfrak{k}}\pi_{\mathfrak{k}^\perp}(U),\pi_{\mathfrak{k}^\perp}(U)]\right)\\ &= -\pi_{\mathfrak{k}^\perp}\left([\pi_{\mathfrak{k}}\pi_{\mathfrak{k}^\perp}(U),\pi_{\mathfrak{k}^\perp}(U)]\right)\\ &= -\left[\pi_{\mathfrak{k}}\pi_{\mathfrak{k}^\perp}(U),\pi_{\mathfrak{k}^\perp}(U)\right],\end{aligned}$$

so defining $\mu := \pi_{\mathfrak{k}^\perp}(U)$, by using the isomorphism $(\pi_{\mathfrak{b}_-})|_{\mathfrak{k}^\perp} : \mathfrak{k}^\perp \to \mathfrak{b}_-$, we can rewrite (10.1) as

$$\dot{\mu} = [\pi_{\mathfrak{k}}(\mu),\mu],$$

which is the full Toda equation.

10.1 Clebsch optimal control formulation for A_r-Toda lattice

We first study to the A_r-Toda system. In this case, B_- is the group of lower triangular $(r+1)\times(r+1)$ matrices with determinant 1 and strictly positive diagonal elements; $\mathfrak{b}_-$ is the Lie algebra of lower triangular traceless matrices; $\mathfrak{k}$ is the Lie algebra of skew-symmetric matrices; $\mathfrak{b}_-^\perp$ consists of strictly lower triangular matrices; and $\mathfrak{k}^\perp$ consists of symmetric traceless matrices. Given $U \in \mathfrak{g} = \mathfrak{sl}(r+1,\mathbb{R})$, we have

$$\pi_{\mathfrak{b}_-}(U) = U_- + U_+^\mathsf{T} + U_0, \quad \pi_{\mathfrak{k}}(U) = U_+ - U_+^\mathsf{T},$$

where the indices $\pm$ and 0 on the matrices denote the strictly upper, lower, and diagonal part, respectively. For $X \in \mathfrak{g}^* = \mathfrak{sl}(r+1,\mathbb{R})$, we have

$$\pi_{\mathfrak{b}_-^\perp}(X) = X_- - X_+^\mathsf{T}, \quad \pi_{\mathfrak{k}^\perp}(X) = X_+^\mathsf{T} + X_0 + X_+.$$

We consider the action of B_- by multiplication on the right on $SL(r+1,\mathbb{R})$ and use the duality pairing between $TSL(r+1,\mathbb{R})$ and $T^*SL(r+1,\mathbb{R})$ given by the bi-invariant extension of the Killing form. For $P,V \in T_QSL(r+1,\mathbb{R})$, we have $\langle P,V\rangle := \operatorname{Tr}(Q^{-1}PQ^{-1}V)$. With respect to this pairing, the momentum map is

$$\mathbf{J} : TSL(r+1,\mathbb{R}) \to \mathfrak{k}^\perp, \quad \mathbf{J}(Q,P) = \pi_{\mathfrak{k}^\perp}(Q^{-1}P).$$

The associated Clebsch optimal control problem, with cost function $\ell(U) = \frac{1}{2}\kappa(U,U)$, yields (see (7.2)),

$$\dot{Q} = QU, \quad \dot{P} = PU, \quad \pi_{\mathfrak{k}^\perp}(U) = \pi_{\mathfrak{k}^\perp}(Q^{-1}P). \tag{10.2}$$

The first two equations represent the *symmetric representation* of the A_r- Toda lattice. In particular, the solution curve $(Q(t),P(t))$ preserves $B_- \times B_-$ similarly to the rigid body case.

10.2 Clebsch optimal control formulation for the Toda lattice associated to an arbitrary Dynkin diagram

For the general Toda system, we let B_- act on the right on G (the connected Lie group underlying the split normal real form). We identify T^*G with TG by using the bi-invariant duality pairing $\langle\,,\rangle_\kappa$ induced by κ.

In this case, the momentum map is given by

$$\mathbf{J} : T^*G = TG \to \mathfrak{k}^\perp, \quad \mathbf{J}(\alpha_Q) = \pi_{\mathfrak{k}^\perp}(TL_{Q^{-1}}\alpha_Q),$$

where we have $\alpha_Q \in T_QG = T_Q^*G$. Indeed,

$$\begin{aligned}\kappa(\mathbf{J}(\alpha_Q), U) &= \langle \alpha_Q, TL_Q U\rangle_\kappa = \left\langle TL_{Q^{-1}}\alpha_Q, U\right\rangle_\kappa = \kappa(TL_{Q^{-1}}\alpha_Q, U) \\ &= \kappa(\pi_{\mathfrak{k}^\perp}(TL_{Q^{-1}}\alpha_Q), U).\end{aligned}$$

For the non exceptional cases at least, the formulas can be written more explicitly since G is given by matrix groups: for A_r, B_r, C_r, D_r we have:

$$G = SL(r-1,\mathbb{R}), \quad G = SO(r+1,r), \quad G = Sp(2r,\mathbb{R}), \quad G = SO(r,r)$$

and the Killing form is given by a multiple of the trace: $\kappa(X,U) = c\,\mathrm{Tr}(XU)$. In this case, the momentum map reads $\mathbf{J}(Q,P) = \pi_{\mathfrak{k}^\perp}(Q^{-1}P)$.

The associated Clebsch optimal control problem with cost function $\ell(U) = \frac{1}{2}\kappa(U,U)$ yields the same equations as in (10.2), understood now in the general sense of A_r, B_r, C_r, D_r. The first two equations being the symmetric representation of the Toda equations. From these conditions, one directly obtains:

$$\begin{aligned}\frac{d}{dt}\pi_{\mathfrak{k}^\perp}(Q^{-1}P) &= \pi_{\mathfrak{k}^\perp}(-Q^{-1}\dot{Q}Q^{-1}P + Q^{-1}\dot{P}) = \pi_{\mathfrak{k}^\perp}(-UQ^{-1}P + Q^{-1}PU) \\ &= -\pi_{\mathfrak{k}^\perp}([U, Q^{-1}P]) = -\pi_{\mathfrak{k}^\perp}([U, \pi_{\mathfrak{k}^\perp}(Q^{-1}P) + \pi_{\mathfrak{b}^\perp}(Q^{-1}P)]) \\ &= -\pi_{\mathfrak{k}^\perp}([U, \pi_{\mathfrak{k}^\perp}(Q^{-1}P)] = -\pi_{\mathfrak{k}^\perp}([U, \pi_{\mathfrak{k}^\perp}(U)],\end{aligned}$$

which is the full Toda equation in Euler-Poincaré form (10.1).

As earlier, the solution curve $(Q(t), P(t))$ preserves the set $B_- \times B_-$.

11 Discrete Models

The symmetric representation of the discrete rigid body:

The Clebsch approach leads to a natural symmetric representation of the discrete rigid body equations of Moser and Veselov. We now define the symmetric representation of the discrete rigid body equations as follows:

$$Q_{k+1} = -U_k Q_k; \qquad P_{k+1} = -U_k P_k\,, \tag{11.1}$$

where $U_k \in SO(N)$ **is defined by**

$$\Lambda U_k - U_k^\top \Lambda = Q_k P_k^\top - P_k Q_k^\top\,. \tag{11.2}$$

We will write this as

$$J_D U_k = Q_k P_k^\top - P_k^\top Q_k^\top\,, \tag{11.3}$$

where $J_D : SO(N) \to \mathfrak{so}(N)$ **(the discrete version of** J**) is defined by** $J_D U = \Lambda U - U^\top \Lambda$**. Notice that the derivative of** J_D **at the identity is** J **and hence, since** J **is invertible,** J_D **is a diffeomorphism from a neighborhood of the identity in** $SO(N)$ **to a neighborhood of** 0 **in** $\mathfrak{so}(N)$**.**

Using these equations, we have the algorithm $(Q_k, P_k) \mapsto (Q_{k+1}, P_{k+1})$ defined by: compute U_k from (11.2), compute Q_{k+1} and P_{k+1} using (11.1). Note that the update map for Q and P is done in parallel.

11.1 The Discrete Variational Problem in the Stiefel case

The discrete variational problem on the Stiefel manifold is given by

$$\min_{Q_k} \sum_k \frac{1}{2} \langle \Lambda Q_{k+1}, Q_k \rangle, \tag{11.4}$$

subject to $Q_k^\top Q_k = I_n$**, i.e.,** $Q_k \in V_n(\mathbb{R}^N)$**. The extremal trajectories for this discrete variational problem are given by**

$$\Lambda Q_{k+1} + \Lambda Q_{k-1} = Q_k B_k, \;\; k \in \mathbb{Z}, \tag{11.5}$$

where $B_k = B_k^\top$ **is a (symmetric) Lagrange multiplier matrix for the symmetric constraint** $Q_k^\top Q_k = I_n$**. Let us define** $U_k := -Q_k Q_{k-1}^\top$ **which implies that**

$$Q_k = -U_k Q_{k-1},$$

Then the following proposition gives the discrete extremal trajectories in terms of U_k **and the discrete body momentum** $M_k := \Lambda U_k - U_k^\top \Lambda$**.**

Proposition 11.1. *The extremal trajectories of the discrete variational problem (11.8) on the Stiefel manifold* $V_n(\mathbb{R}^N)$ *in terms of* (M_k, U_k) *are given by:*

$$M_{k+1} = U_k M_k U_k^\top + A_k, \tag{11.6}$$

where

$$A_k := U_k \Lambda \big(I_N - U_k U_k^\top\big) - \big(I_N - U_k U_k^\top\big) \Lambda U_k^\top. \tag{11.7}$$

11.2 Discrete variational problem for the flows on symmetric matrices:

The natural optimization problem for the flow on symmetric matrices is

$$\min_{U_k} \sum_k \frac{1}{2} \langle N^{-1} U_k, N^{-1} U_k \rangle, \tag{11.8}$$

subject to $Q_{k+1} = Q_k U_k$**.**

Here, as in the smooth case

$$\left\{ Q_k \in GL(2k, \mathbb{R}) \mid Q_k^\top N^{-1} Q_k = N^{-1} \right\} . \tag{11.9}$$

Thus we have

$$Q_k^\top N^{-1} Q_{k+1} = N^{-1} U_k \tag{11.10}$$

and hence the optimization problem may be reformulated as

$$\min_{Q_k} \sum_k \frac{1}{2} \langle Q_k^\top N^{-1} Q_{k+1}, Q_k^\top N^{-1} Q_{k+1} \rangle, \tag{11.11}$$

subject to

$$Q_k^\top N^{-1} Q_k = N^{-1} . \tag{11.12}$$

Choosing a skew symmetric matrix B_k of Lagrange multipliers we see that the relevant equations take the form

$$N^{-1}Q_{k+1}Q_{k+1}N^{-1}Q_k + N^{-1}Q_{k-1}Q_{k-1}N^{-1}Q_k + N^{-1}Q_kB_k = 0. \tag{11.13}$$

This gives a natural analogue of the Moser Veselov equations.

There is much more on the connection with Moser...

The Clebsch representation in optimal control, integrable systems and discrete dynamics - Talk by Anthony Bloch

Lecture notes - Ori S. Katz

October 10, 2018

Abstract

In this talk we discuss recent work on a geometric approach to certain optimal control problems and the relationship of the solutions of these problems to some classical integrable dynamical systems. These systems include the rigid body equations, geodesic flows on the ellipsoid and the Toda lattice flows. We discuss the Hamiltonian structure of these systems and relate our work to some classical work of Moser. We also discuss the link to discrete dynamics and symplectic integration. The work is joint with Francois Gay-Balmaz and Tudor Ratiu.

1 Lecture notes

Links to the talk of Melvin Leok from yesterday.

Symmetric rigid body equations - alternative formulation of the rigid body equations, related to discrete dynamics.

Rigid body equations can be written as a Lax pair form $\dot{M} = [M, \Omega]$. Λ - matrix of the angular momentum.

Toda flow - arises from study of particles on a line. Gradient form - $\dot{X} = [X, [X, N]]$.

n-dimensional rigid body: I - inertia matrix, Ω - matrix of angular velocities. Evolves on the momentum sphere. Has three equilibria, two stable and one unstable, around rotation directions.

Dynamics can be written as an Euler-Poincare form, or in Hamiltonian form.

In n dimensions, look at the rigid body equations on $SO(n)$. M is a general skew-symmetric matrix, Ω is the angular velocity, and the symmetric operator J relates the two. This is an integrable system that can be solved in various ways.

This can be generalized to any semi-simple Lie group. That is the so-called left-invariant system.

It can also be formulated as a right-invariant system by simple transformations and a sign change.

Symmetric rigid body equations: Related to optimal control problems and discrete dynamics. First we will derive them “out of thin air”.

Q - configuration variable, P - momentum variable that can be thought of as a dual variable to Q.

M is skew symmetric (by construction), so instead of the original rigid body equations we have these square equations.

Proposition 2.1 is proved by simple computation - substitute these equations into the derivative of the classical rigid body equations.

Are they exactly equivalent? They are not exactly equivalent; the equations are topologically different, but they are very close. Solutions from one set of equations can be mapped to the other (in both directions). On any level set of momenta, they give the exact same dynamics.

How is this related to optimal control and discrete dynamics?

Optimal control - reminder: Given an affine control system, many times people consider kinematic systems where the drift vector field $f(x)$ is zero. The controls may be bounded or un-bounded, and the main difference between control systems and ODE’s is that in control systems the endpoints are pre-determined.

Pontryagin maximum principle - H^* is the optimal Hamiltonian producing the optimal dynamics.

Can ask, is the system solvable? Can one explicitly find the controls? Generally, the equations cannot be solved. Here, we look at the solvable case, solvable in the classical Hamiltonian way - Arnold-Liouville integrable.

Optimal control formulation of rigid body (page 16) - $\langle U, J(U)\rangle$ is the cost function. P is the costate vector that enforces the kinematics, it is the constraint. In fact, the symmetric rigid body equations presented before were originally found by solving this optimal control problem.

This is related to discrete variational problems.

Page 19 - the discrete Euler-Lagrange equations.

They can be rewritten with Φ, a type of forward map.

Discrete rigid body equations of dimension n = DRBn.

Here, Ω_k are group elements in $SO(n)$.

The discrete symmetric rigid body - cubic set of discrete equations. Looks like the original symmetric rigid body equations. Locally, these equations are exactly the same.

This can also be intrinsically derived from the discrete optimal control problem.

Generalization - the Clebsch optimal control problem. Generating a cost function l, the control problem is solved by a momentum map method resulting in a Hamiltonian formulation.

Observation: Dynamics come from an infinitesimal generator, so the problem can be reformulated on a G-orbit.

(page 35) There is a very nice metric that comes out of this - the normal metric. It is related to the Toda lattice. It is complementary to the symplectic flow; it is the metric to be used in order to write down the gradient flow - the dual flow.

(page 41) Stiefel manifolds - the framework (for non-square matrices) includes the rigid body equations and other interesting problems.

(page 45) This is a generalization of the rigid body equations. For the rigid body equations, the last term disappears (example 1).

The free rigid body system is integrable. How to find the corresponding integrals of motion? By rewriting the equations as a Lax pair equation set (Mischenko and Fomenko paper).

This Clebsch optimal control on Steifel manifolds framework produces a rich family of integrable problems, from geodesic flow on ellipsoids to the rigid body system.