Smooth and Discrete Integrable Systems and Optimal Control

Anthony Bloch
(Work with Peter Crouch and Nikolay Nordkvist)
Also: Jerry Marsden, Tudor Ratiu and Amit Sanyal, Darryl Holm
See also work of Gay-Balmaz/Ratiu

- Embedded and Associated Optimal Control Problems

- Symmetric rigid body equations – smooth and discrete

- Flows on Stiefel manfolds – Jacobi flow on ellipsoid

- Flows on Quadratic groups

Rigid Body Equations:

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

Symmetric Rigid Body Equations:

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

Coupled double bracket flows:

$$\dot{x} = [x, [x, p]], \tag{0.1}$$

$$\dot{p} = [p, [x, p]], \tag{0.2}$$

(with Brockett and Crouch)

• Continuous Embedded Optimal Control Problems

Let Q denote a (finite dimensional) manifold and let $\mathfrak{X}(Q)$ denote the space of smooth vector fields on Q. Consider $D \subset \mathfrak{X}(Q)$ and $\mathcal{D} \subset TQ$ defined at every $q \in Q$ by

$$\mathcal{D}(q) = \text{span}\{X(q),\ X \in D\}.$$

This is an example of a generalized distribution; if the rank of $\mathcal{D}$ is constant on Q it is a distribution in the classical sense. Consider the control problem

$$\dot{q} = \sum_{i=1}^{m} X_i(q) u_i, \tag{0.3}$$

where $q \in Q$,$X_i \in \mathfrak{X}(Q)$, $u = (u_1, \ldots, u_m) \in U$ with $U \subset \mathcal{R}^m$ an open neighborhood of $0 \in \mathcal{R}^m$. Let $N \subset Q$ denote an immersed submanifold of Q which is invariant under the flow of (0.3). If N is a maximal integral manifold of an integrable generalized distribution $\mathcal{D}$ then N is an immersed submanifold of Q which is invariant by construction.

Let $\ell : Q \times U \to \mathcal{R}$ be a cost function. For each choice of invariant submanifold $N \subset Q$ we introduce the concept of an embedded optimal control problem as follows:

Embedded optimal control problem. *Minimize*

$$\int_0^T \ell\big(q(t), u(t)\big)\mathrm{d}t,$$

subject to $\dot{q} = \sum_{i=1}^m X_i(q)u_i$, $q \in Q$, $u = (u_1, \ldots, u_m) \in U \subset \mathcal{R}^m$, *and with fixed endpoints* $q(0) = q_0 \in N$ *and* $q(T) = q_T \in N$.

The embedded optimal control problem is well posed when (0.3) restricted to N is accessible from q_0. When this is the case we can impose on q_T that it needs to belong to the set of states reachable from q_0 in time T. The associated optimal control problem on N is now given by:

Associated optimal control problem. *Minimize*

$$\int_0^T \ell\big(q(t),u(t)\big)\mathrm{d}t,$$

subject to $\dot q = \sum_{i=1}^m X_i|_N(q)u_i$, $q \in N$, $u = (u_1,\dots,u_m) \in U \subset \mathcal{R}^m$, *and with fixed endpoints* $q(0) = q_0 \in N$ *and* $q(T) = q_T \in N$.

When $N \subset Q$ is an immersed submanifold of Q the inclusion $i_N : N \hookrightarrow Q$ is an immersion. The pullback bundle $i_N^*(T^*Q)$ is defined as the vector bundle over N whose fiber over $n \in N$ is given by $T^*_{i_N(n)}Q$. Since $T_n^*i_N : T^*_{i_N(n)}Q \to T_n^*N$, the dual of the tangent map of i_N is globally defined when restricted to $i_N^*(T^*Q)$. Furthermore since $i_N : N \hookrightarrow Q$ is an immersion we have that $T_ni_N : T_nN \to T_{i(n)}Q$ is injective for all $n \in N$ and therefore $T_n^*i_N : T^*_{i_N(n)}Q \to T_n^*N$ is surjective; that is $T^*i_N|_{i_N^*(T^*Q)}$ is surjective on fibers.

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). ■

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 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

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

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]).

Merging the Left and Right Problems. We will now show both the symmetric representation of the rigid body equations in both left and right invariant form arise from a rather general optimal control problem that includes the one above as a special case.

We begin by recalling a general optimal control problem on matrices:

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 2.3.*

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 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*

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

Define the optimal control problem

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

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*

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

6 Stiefel Manifolds

We introduce the variational and optimal control problems on a Stiefel manifold based on minimizing the time integral of the kinetic energy.

The metric on the manifold is given by the kinetic energy expression. We also give the extremal flows obtained in the limiting cases of the sphere/ellipsoid ($n = 1$), and the N dimensional rigid body ($n = N$). The extremal flows in these cases are well-known and integrable.

Variational Problems on a Stiefel Manifold

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 than 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 (see Marsden and Ratiu, 1999) on $V(1, N) = \mathbb{S}^{N-1}$. The equation of motion can be written as

$$\Lambda \ddot{q} = bq, \tag{6.8}$$

where b is a real scalar in this case. We get the Lagrangian (variational) equations for the geodesic flow on the sphere ($\mathbb{S}^{N-1}$) as

$$\ddot{q} = -\frac{\dot{q}^{\mathrm{T}}\dot{q}}{q^{\mathrm{T}}\Lambda^{-1}q}\Lambda^{-1}q. \tag{6.9}$$

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.

We can similarly obtain the equations for the general Stiefel case.
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 $g - g^{-1} \in \mathfrak{g}$.

We use the trace pairing in $\mathfrak{gl}(n)$:

$$\langle A, B\rangle = \operatorname{Trace}(A^{\mathsf{T}} B). \tag{7.2}$$

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

$$\min \int_0^T \frac{1}{4}\langle U, \Sigma(U)\rangle \mathrm{d}t \tag{7.3}$$

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 (7.3) is then defined as

$$\begin{aligned} H(P, Q, U) &= \langle P, QU\rangle - \frac{1}{4}\langle U, \Sigma(U)\rangle \\ &= \langle Q^{\mathsf{T}} P, U\rangle - \frac{1}{4}\langle U, \Sigma(U)\rangle. \end{aligned} \tag{7.4}$$

Proposition 7.1. *The necessary conditions for optimality of a solution to the optimal control problem (7.3) with costate $P \in \mathbb{R}^{n\times n}$ yield the following Hamilton's equations*

$$\dot{Q} = QU, \quad \dot{P} = -PU^{\mathsf{T}}. \tag{7.5}$$

Lemma 7.2. *The extremal controls for the optimal control problem (7.3) when $P \in G$ are given by*

$$U_{ext} = \Sigma^{-1}\left(Q^{\mathsf{T}}P - (Q^{\mathsf{T}}P)^{-1}\right). \tag{7.6}$$

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

$$\Omega_{\mathrm{can}}((X_1, Y_1), (X_2, Y_2)) = \langle Y_2, X_1\rangle - \langle Y_1, X_2\rangle. \tag{7.7}$$

Proposition 7.3. *The extremal flow (7.5) generated by the optimal control problem (7.3) which evolves on the canonical symplectic manifold* $(\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}, \Omega_{\mathrm{can}})$ *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

$$H(P, Q, U_{ext}) = \frac{1}{4}\langle M, \Sigma^{-1}(M)\rangle, \tag{7.8}$$

and the extremal control can be expressed as

$$U_{ext} = \Sigma^{-1}(M) \in \mathfrak{g}. \tag{7.9}$$

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

$$\widehat{\sigma} : \mathfrak{g} \to \mathfrak{g}; \quad \widehat{\sigma}(A) = -A^{\mathsf{T}}. \tag{7.10}$$

Can show:

Theorem 7.4. *The "generalized Euler" equations for the optimal control problem (7.3) are given by*

$$\dot{Q} = QU, \quad \dot{M} = [M, \widehat{\sigma}(U)], \quad U = \Sigma^{-1}(M). \tag{7.11}$$

To pass between the two formulations we consider the map

$$\Phi : G \times G \to G \times \mathfrak{g}, \quad (Q, P) \ \mapsto \ (Q, M) \tag{7.12}$$

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

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

$$P = \sigma(Q) \exp\left(\sinh^{-1} \frac{M}{2}\right), \tag{7.13}$$

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*

$$S \triangleq \left\{ (Q, P) \in G \times G \;\mid\; m = P\sigma(Q^{-1}) - \sigma(Q)P^{-1},\; \|m\| < 2 \right\}, \tag{7.14}$$

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

$$\min_{U_k} \sum_{k=1}^{N} \langle \Delta, U_k \rangle, \qquad \Delta = \frac{1}{2}(\Lambda + \Lambda^{\mathsf{T}}), \tag{7.15}$$

subject to

$$Q_{k+1} = Q_k U_k, \qquad Q_0 = \overline{Q}_0, \;\; Q_N = \overline{Q}_N. \tag{7.16}$$

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 (7.15) is given by a sequence of matrices (Q_k, P_k) in $G \times G$ satisfying the optimal evolution equations*

$$Q_{k+1} = Q_k U_k, \qquad P_{k+1} = P_k \sigma(U_k), \tag{7.17}$$

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

$$U_k \Delta - \Delta U_k^{-1} = P_k^{\mathsf{T}} Q_k - \left(P_k^{\mathsf{T}} Q_k\right)^{-1}. \tag{7.18}$$

8 Continuous Embedded Optimal Control Problems

Let Q denote a (finite dimensional) manifold and let $\mathfrak{X}(Q)$ denote the space of smooth vector fields on Q. Consider $D \subset \mathfrak{X}(Q)$ and $\mathcal{D} \subset TQ$ defined at every $q \in Q$ by

$$\mathcal{D}(q) = \text{span}\{X(q),\ X \in D\}.$$

This is an example of a generalized distribution; if the rank of $\mathcal{D}$ is constant on Q it is a distribution in the classical sense. Consider the control problem

$$\dot{q} = \sum_{i=1}^{m} X_i(q) u_i, \tag{8.1}$$

where $q \in Q$,$X_i \in \mathfrak{X}(Q)$, $u = (u_1, \ldots, u_m) \in U$ with $U \subset \mathcal{R}^m$ an open neighborhood of $0 \in \mathcal{R}^m$. Let $N \subset Q$ denote an immersed submanifold of Q which is invariant under the flow of (8.1). If N is a maximal integral manifold of an integrable generalized distribution $\mathcal{D}$ then N is an immersed submanifold of Q which is invariant by construction, see, e.g., **?**. Let $\ell : Q \times U \to \mathcal{R}$ be a cost function. For each choice of invariant submanifold $N \subset Q$ we introduce the concept of an embedded optimal control problem as follows:

Embedded optimal control problem. *Minimize*

$$\int_0^T \ell\big(q(t), u(t)\big)\mathrm{d}t,$$

subject to $\dot{q} = \sum_{i=1}^m X_i(q)u_i$, $q \in Q$, $u = (u_1, \ldots, u_m) \in U \subset \mathcal{R}^m$, *and with fixed endpoints* $q(0) = q_0 \in N$ *and* $q(T) = q_T \in N$.

The embedded optimal control problem is well posed when (8.1) restricted to N is accessible from q_0. When this is the case we can impose on q_T that it needs to belong to the set of states reachable from q_0 in time T. The associated optimal control problem on N is now given by:

Associated optimal control problem. *Minimize*

$$\int_0^T \ell\big(q(t), u(t)\big)\mathrm{d}t,$$

subject to $\dot{q} = \sum_{i=1}^m X_i|_N(q)u_i$, $q \in N$, $u = (u_1, \ldots, u_m) \in U \subset \mathcal{R}^m$, *and with fixed endpoints* $q(0) = q_0 \in N$ *and* $q(T) = q_T \in N$.

When $N \subset Q$ is an immersed submanifold of Q the inclusion $i_N : N \hookrightarrow Q$ is an immersion. The pullback bundle $i_N^*(T^*Q)$ is defined as the vector bundle over N whose fiber over $n \in N$ is given by $T^*_{i_N(n)}Q$. Since $T_n^* i_N : T^*_{i_N(n)}Q \to T_n^*N$, the dual of the tangent map of i_N is globally defined when restricted to $i_N^*(T^*Q)$. Furthermore since $i_N : N \hookrightarrow Q$ is an immersion we have that $T_n i_N : T_n N \to T_{i(n)}Q$ is injective for all $n \in N$ and therefore $T_n^* i_N : T^*_{i_N(n)}Q \to T_n^* N$ is surjective; that is $T^* i_N|_{i_N^*(T^*Q)}$ is surjective on fibers.

We define the map $\mathcal{P} : \mathfrak{X}(Q) \to \mathcal{F}(T^*Q)$ by

$$\mathcal{P}(X)(\alpha_q) = \langle \alpha_q, X(q) \rangle ,$$

where $\langle \cdot, \cdot \rangle$ is the natural pairing between covectors and vectors. The function $\mathcal{P}(X)$ is called the momentum function of X. Similarly we define the map $\bar{\mathcal{P}} : \mathfrak{X}(N) \to \mathcal{F}(T^*N)$ by

$$\bar{\mathcal{P}}(Y)(\beta_q) = \langle \beta_q, Y(q) \rangle ,$$

where again $\bar{\mathcal{P}}(Y)$ is called the momentum function of Y.

We define the following vector field on Q

$$X_u(q) = \sum_{i=1}^{m} X_i(q) u_i. \tag{8.2}$$

The Hamiltonian function $H : T^*Q \times U \to \mathcal{R}$ of interest is defined as

$$H(\alpha_q, u) := \langle \alpha_q, X_u(q) \rangle - \lambda_0 \ell(q, u) = \mathcal{P}(X_u)(\alpha_q) - \lambda_0 \ell(q, u),$$

where λ_0 is a constant with value either $\lambda_0 = 1$ or $\lambda_0 = 0$. Similarly the Hamiltonian function $\bar{H} : T^*N \times U \to \mathcal{R}$ is defined as

$$\bar{H}(\beta_q, u) := \langle \beta_q, X_u|_N(q) \rangle - \lambda_0 \ell(q, u) = \bar{\mathcal{P}}(X_u|_N)(\beta_q) - \lambda_0 \ell(q, u).$$

The solutions to the embedded optimal control problem and the associated optimal control problem as given by Pontryagin's principle, which utilizes these Hamiltonians, are called normal extremals for $\lambda_0 = 1$ and abnormal extremals for $\lambda_0 = 0$.

The variational derivative $\frac{\delta \ell}{\delta q}(q,u) \in T_q^* Q$ is defined as

$$\left\langle \frac{\delta \ell}{\delta q}(q,u), \delta q \right\rangle = \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \ell(\gamma_\epsilon, u), \tag{8.3}$$

where γ_ϵ depends smoothly on ϵ and $\gamma_0 = q$ and $\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \gamma_\epsilon = \delta q \in T_q Q$. Similarly we define the variational derivative $\frac{\bar{\delta} \ell}{\bar{\delta} q}(q,u) \in T_q^* N$ as

$$\left\langle \frac{\bar{\delta} \ell}{\bar{\delta} q}(q,u), \bar{\delta} q \right\rangle = \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \ell(\bar{\gamma}_\epsilon, u), \tag{8.4}$$

where $\bar{\gamma}_\epsilon$ depends smoothly on ϵ and $\bar{\gamma}_0 = q$ and $\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \bar{\gamma}_\epsilon = \bar{\delta} q \in T_q N$. For $q \in N$ we have

$$\frac{\bar{\delta} \ell}{\bar{\delta} q}(q,u) = T_q^* i_N \left(\frac{\delta \ell}{\delta q}(q,u) \right). \tag{8.5}$$

For $\alpha, \sigma \in T_q^*Q$ the vertical lift of σ relative to α is

$$\mathrm{ver}_\alpha(\sigma) := \left.\frac{d}{ds}\right|_{s=0} (\alpha + s\sigma) \in T_\alpha(T^*Q). \tag{8.6}$$

Similarly for $\beta, \eta \in T_q^*N$ the vertical lift of η relative to β is

$$\bar{\mathrm{ver}}_\beta(\eta) := \left.\frac{d}{ds}\right|_{s=0} (\beta + s\eta) \in T_\beta(T^*N). \tag{8.7}$$

Then we have the following result:

Theorem 8.1. *Let λ_0 be fixed as $\lambda_0 = 1$ or $\lambda_0 = 0$. An extremal (normal if $\lambda_0 = 1$, abnormal if $\lambda_0 = 0$) for the embedded optimal control problem as given by applying Pontryagin's maximum principle is a solution of*

$$\frac{\partial H}{\partial u} = 0, \quad \dot{\alpha} = X_{\mathcal{P}(X_u)}(\alpha) + \lambda_0 \mathrm{ver}_\alpha \left(\frac{\delta \ell}{\delta q} \right), \tag{8.8}$$

for $\alpha \in T^*Q$. The pullback bundle $i_N^*(T^*Q)$ is invariant under the flow of (8.8).

An extremal (normal if $\lambda_0 = 1$, abnormal if $\lambda_0 = 0$) for the associated optimal control problem is a solution of

$$\frac{\partial \bar{H}}{\partial u} = 0, \quad \dot{\beta} = X_{\bar{\mathcal{P}}(X_u|_N)}(\beta) + \lambda_0 \bar{\mathrm{ver}}_\beta \left(\frac{\bar{\delta} \ell}{\bar{\delta} q} \right), \tag{8.9}$$

for $\beta \in T^*N$.

If $\alpha(t) \in i_N^*(T^*Q)$ is a solution to (8.8) then $\beta(t) := T^* i_N\big(\alpha(t)\big) \in T^*N$ is a solution to (8.9).

Proof uses Pontryagin.

Note: The canonical symplectic structure on T^*Q and T^*N giving the two different extremal generating Hamiltonian vector fields can be very different, and thus it is not obvious that a solution to the equations for the embedded optimal control problem can be "projected" via T^*i_N to a solution of the equations for the associated problem.

Also this theorem shows that normal extremals "project" to normal extremals and abnormal extremals "project" to abnormal extremals. Since N is an immersed submanifold of Q we have that for any solution $\beta(t)$ to the associated optimal control problem there exists a solution $\alpha(t)$ to the embedded optimal control problem which "projects" via T^*i_N to $\beta(t)$.

Therefore we can find solutions to the associated optimal control problem in terms of the optimal trajectory $q(t)$ and control $u(t)$ by instead finding solutions (in terms of the optimal trajectory $q(t)$ and control $u(t)$) to the embedded optimal control problem. For example if Q is a linear space and N is a nonlinear space it will most likely be much easier to solve the embedded optimal control problem than the associated optimal control problem.

If $\mathcal{D}$, where $\mathcal{D}(q) = \text{span}\{X_1(q), \ldots, X_m(q)\}$, is completely integrable then the extremal generating equations for the embedded optimal control problem gives the solution to any associated optimal control problem for N being a maximal integral manifold of $\mathcal{D}$. This means that the embedded optimal control problem prescribes a foliation of solutions to all associated optimal control problems on leaves of $\mathcal{D}$.