The symmetric representation of the rigid body equations and their discretization

Anthony M Bloch[1]**, Peter E Crouch**[2]**, Jerrold E Marsden**[3] **and Tudor S Ratiu**[4]

[1] Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA
[2] Center for Systems Science and Engineering, Arizona State University, Tempe, AZ 85287, USA
[3] Control and Dynamical Systems 107-81, California Institute of Technology, Pasadena, CA 91125, USA
[4] Département de Mathématiques, Ecole Polytechnique Fédérale de Lausanne, Lausanne CH-1015, Switzerland

E-mail: abloch@math.lsa.umich.edu, peter.crouch@asu.edu, marsden@cds.caltech.edu and ratiu@dma.epfl.ch

Abstract
This paper analyses continuous and discrete versions of the generalized rigid body equations and the role of these equations in numerical analysis, optimal control and integrable Hamiltonian systems. In particular, we present a symmetric representation of the rigid body equations on the Cartesian product $SO(n) \times SO(n)$ and study its associated symplectic structure. We describe the relationship of these ideas with the Moser–Veselov theory of discrete integrable systems and with the theory of variational symplectic integrators. Preliminary work on the ideas discussed in this paper may be found in Bloch *et al* (Bloch A M, Crouch P, Marsden J E and Ratiu T S 1998 *Proc. IEEE Conf. on Decision and Control* **37** 2249–54).

1. Introduction

This paper presents an alternative formulation of the n-dimensional rigid body equations which we call the *symmetric representation of the rigid body equations* (equations (SRBn) in section 3) and an associated set of discrete equations called the *symmetric representation of the discrete rigid body equations* (equations (SDRBn) in section 4). Both the continuous and discrete equations evolve on a Cartesian product $G \times G$ of a Lie group G rather than on its cotangent bundle T^*G. One interesting way to derive the continuous equations is by means of the

(Pontryagin) maximum principle of optimal control theory. Likewise, the discrete equations can be derived from discrete optimal control theory.

Parts of this work were motivated by the work of Moser and Veselov (1991) on integrable discrete rigid body equations (the Moser–Veselov equations are (DRBn) in section 4) and through this, with problems in geometric numerical integration algorithms. In fact, there has been much interest in recent years in structured algorithms (such as variational and symplectic methods) for integrating Hamiltonian systems and, in particular, rigid body mechanics (see, e.g. Lewis and Simo (1995), McLaghlan and Scovel 1995, Marsden and Wendlandt (1997), Marsden *et al* (1998, 1999), Bobenko and Suris (1999), Kane *et al* (1999, 2000), Marsden and West (2001) and references therein). A particularly interesting feature of the rigid body equations is that the discrete form is still integrable in a precise sense, as shown by Moser and Veselov (1991) (see also Deift *et al* (1992) and Suris (2001)).

We show that the discrete rigid body equations (SDRBn) can not only be obtained as a discretization of our continuous equations (SRBn), but may be obtained directly by a discrete version of the maximum principle. We also establish the variational formulation of the discrete equations and the sense in which the continuous case is obtained as the time step tends to zero.

We also discuss some links with the Manakov equations with parameter in both their continuous and discrete setting. Recall that the Manakov equations are key to understanding the integrability of the n-dimensional rigid body (see Manakov (1976)).

Relationship to other optimal control problems. Our formulation is related to the development of optimal control problems on adjoint orbits of Lie groups defined using what is usually called the normal metric and a weighted generalization of it (see Brockett (1994), Bloch and Crouch (1996) and Bloch *et al* (1997)). In this work one considers the optimal control problem for a cost function that is the energy function associated with the normal metric. On adjoint orbits this gives rise to a set of *coupled double bracket equations*. We remark that this *pair* of coupled double bracket equations is quite different from the double bracket equations discussed in Brockett (1989), Bloch (1990) and Bloch *et al* (1990a,b, 1992). This pair of equations is discussed below (see equation (5.10)). These equations are always Hamiltonian and specialize to a number of integrable Hamiltonian systems of interest, such as the geodesic flow equations on adjoint orbits and Grassmannians.

Optimization vs optimal control problems. In contrast to the coupled double bracket equations, the double bracket equations arose historically from optimization (rather than optimal control) problems. The double bracket equations arose in fact from a steepest descent approach to optimization and are gradient in general (with respect to the normal metric). They are Hamiltonian only in certain situations, e.g. they yield the tridiagonal symmetric Toda lattice equation (see Bloch (1990) and Bloch *et al* (1990a)).

The coupled double bracket equations, on the other hand, arise naturally in optimal control problems via the maximum principle, as discussed in Bloch and Crouch (1996), and provide, in some cases, explicit solutions to these optimal control problems.

As a consequence of the maximum principle, it follows that the coupled double bracket equations are Hamiltonian. A related paper of interest on explicitly soluble optimal control problems is that of Faybusovich (1988), where the methods of integrable systems in optimal control problems are discussed (see also Jurdjevic (1997)).

Structure of the symmetric representation of the rigid body equations. Our particular interest in this paper is the development of a particular symmetric representation of the generalized

rigid body equations on the n-dimensional proper orthogonal group $SO(n)$ (whose Lie algebra is denoted $\mathfrak{so}(n)$). We show that these equations can be put into the following symmetric form:

$$\dot{Q} = Q\Omega, \qquad \dot{P} = P\Omega. \tag{1.1}$$

The notation in these equations is as follows: the matrices Q and P are the dynamical variables, where $Q \in SO(n)$ denotes the configuration of the body. For these equations to make sense as first-order equations on $SO(n) \times SO(n)$, one needs to specify how $\Omega = Q^{-1}\dot{Q} \in \mathfrak{so}(n)$, the body angular velocity, is a function of Q and P. This will be explained in the main text.

Setting $A = PQ^{\mathrm{T}}$, skew symmetry of Ω shows that $\dot{A} = 0$ and hence $A = P(0)Q(0)^{\mathrm{T}}$ and so $P(t) = P(0)Q(0)^{\mathrm{T}}Q(t)$. In other words, this 'quadrature' shows that P may be thought of as a function of Q. Thus, Ω, instead of being a function of (Q, P), becomes a function of Q alone. This indicates that the integrable structure of the symmetric representation of the rigid body equations depends simply on the integrability of the reduced order system:

$$\dot{Q} = Q\Omega(Q), \tag{1.2}$$

where the right-hand side turns out to be cubic in Q. This may be viewed as a reduction process under the group $SO(n)$. In stark contrast, for the classical rigid body equations the same symmetry reduction results in the standard Lax system $\dot{M} = [M, \Omega]$ and the reconstruction process requires the integration of the kinematic system $\dot{Q} = Q\Omega$.

One of the key developments in this paper is the establishment of a notion of equivalence between this usual system of rigid body equations

$$\dot{Q} = Q\Omega, \qquad \dot{M} = [M, \Omega], \tag{1.3}$$

and our symmetric representation of the rigid body equations (1.1). Indeed, Bloch and Crouch (1996) demonstrate that on suitable domains,

$$P = Q(\mathrm{e}^{\sinh^{-1} M/2}) \tag{1.4}$$

relates trajectories of (1.1) and (1.3). This equivalence, effected by a mapping Ψ, that we introduce in the paper, therefore demonstrates that the two-step process involved in integrating the classical equations (1.3) (solving for M first and then Q) may be replaced simply by integration of one system (1.2) followed by use of the identity $P = AQ$.

We further identify in this paper invariant sets $S \subset SO(n) \times SO(n)$ and $S_M \subset T^*SO(n)$ and demonstrate that S inherits a symplectic structure so that Ψ is a symplectomorphism from S to S_M mapping trajectories of (1.1)–(1.3).

While Q has the interpretation as the configuration of the body, P is not the rigid body angular momentum. As discussed in Bloch and Crouch (1996) and later in this paper, P is naturally a costate variable arising from studying the n-dimensional rigid body as an optimal control problem. One may want to consider how this result for the rigid body system may be extended to other integrable systems such as the heavy top (see, e.g. Lewis *et al* (1992) and Bobenko and Suris (1999)). While this is not clear from the work in this paper (due to the reduced symmetry in systems such as the heavy top), the key observation here is the means by which the variable P is identified, as a costate vector in an associated optimal control problem. Of course, every Euler–Lagrange system, coming from Hamilton's principle, may be thought of as an optimal control problem. We hypothesize that by identifying optimal control problems whose extremals yield the equations of motion of other integrable systems, further insight may be gained into their structure as is clearly the case for the rigid body. Clearly this is not an obvious process as the complicated relationship between the rigid body and the symmetric representation of the rigid body demonstrates.

Discrete representation of the symmetric rigid body equations. The discrete representation of the symmetric rigid body equations is of the form

$$Q_{k+1} = Q_k U_k, \qquad P_{k+1} = P_k U_k, \tag{1.5}$$

where again Q_k and P_k are in $SO(n)$ and U_k is a function of Q_k and P_k. Just as (SRBn) is locally equivalent in a precise sense to the standard rigid body equations, these equations, which we dub the symmetric representation of the discrete rigid body equations (SDRBn), are equivalent in a precise sense to the Moser–Veselov discrete rigid body equations and moreover define a new algorithm for integrating the rigid body equations. We also show how these equations arise from the discrete Pontryagin maximum principle.

Outline. The structure of this paper is as follows: in section 2 we describe the classical generalized rigid body equations on $SO(n)$. In section 3 we introduce the symmetric representation of the rigid body equations and discuss their equivalence with the standard rigid body equations. In section 4 we discuss both the discrete Moser–Veselov equations and the discrete representation of the symmetric rigid body equations. In section 5 we discuss how to derive the (SRBn) equations from optimal control theory and in section 6 we show how to derive (SDRBn) from the theory of discrete optimal control. In section 7 we discuss various relationships between (RBn) and (SRBn). Finally, in section 8, we present some conclusions and discuss some planned extensions of this work.

2. The *n*-dimensional rigid body

In this section we review the classical rigid body equations in three and, more generally, in n dimensions. We shall also compare the left and right invariant equations.

For convenience we shall use the following pairing (multiple of the Killing form) on $\mathfrak{so}(n)$, the Lie algebra of $n \times n$ real skew matrices regarded as the Lie algebra of the n-dimensional proper rotation group $SO(n)$:

$$\langle \xi, \eta \rangle = -\tfrac{1}{2} \operatorname{trace}(\xi \eta). \tag{2.1}$$

The factor of $\tfrac{1}{2}$ in (2.1) is to make this inner product agree with the usual inner product on $\mathbb{R}^3$ when it is identified with $\mathfrak{so}(3)$ in the following standard way: associate the 3×3 skew matrix $\hat{u}$ to the vector u by $\hat{u} \cdot v = u \times v$, where $u \times v$ is the usual cross product in $\mathbb{R}^3$. We use this inner product to identify the dual of the Lie algebra, namely $\mathfrak{so}(n)^*$, with the Lie algebra $\mathfrak{so}(n)$.

We recall from Manakov (1976) and Ratiu (1980) that the left invariant generalized rigid body equations on $SO(n)$ may be written as

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

where $Q \in SO(n)$ 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

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

is the body angular momentum. Here $J : \mathfrak{so}(n) \to \mathfrak{so}(n)$ is the symmetric (with respect to the inner product (2.1)), positive definite, and hence invertible, operator defined by

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

where Λ is a diagonal matrix satisfying $\Lambda_i + \Lambda_j > 0$ for all $i \neq j$. For $n = 3$ the elements of Λ_i are related to the standard diagonal moment of inertia tensor I by $I_1 = \Lambda_2 + \Lambda_3$, $I_2 = \Lambda_3 + \Lambda_1$ and $I_3 = \Lambda_1 + \Lambda_2$.

The equations $\dot{M} = [M, \Omega]$ are readily checked to be the Euler–Poincaré equations on $\mathfrak{so}(n)$ for the Lagrangian $l(\Omega) = \frac{1}{2}\langle \Omega, J(\Omega)\rangle$. This corresponds to the Lagrangian on $TSO(n)$ given by

$$L(g, \dot{g}) = \tfrac{1}{2}\langle g^{-1}\dot{g}, J(g^{-1}\dot{g})\rangle. \qquad (2.2)$$

It follows from the general Euler–Poincaré theory (see, e.g. Marsden and Ratiu (1999)) that the equations (RBn) are the geodesic equations on $TSO(n)$, left trivialized as $SO(n) \times \mathfrak{so}(n)$, relative to the left invariant metric whose expression at the identity is

$$\langle\langle \Omega_1, \Omega_2\rangle\rangle = \langle \Omega_1, J(\Omega_2)\rangle. \qquad (2.3)$$

According to Mishchenko and Fomenko (1978), there is a similar formalism for any semisimple Lie group and that in that context, one has integrability on the generic coadjoint orbits.

Right invariant system. The system (RBn) has a right invariant counterpart. This right invariant system is given as follows. Consider the right invariant Riemannian metric on $SO(n)$ whose value at the identity is given by (2.3). The geodesic equations of this metric on $TSO(n)$, right trivialized as $SO(n) \times \mathfrak{so}(n)$, are given by

$$\dot{Q}_r = \Omega_r Q_r, \qquad \dot{M}_r = [\Omega_r, M_r], \qquad \text{(RightRBn)}$$

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

The proof is a straightforward verification.

The relation between the left and right systems given in this proposition is not to be confused with the right trivialized representation of the left invariant rigid body equations, i.e. the left invariant system written in spatial representation. For a discussion of this distinction, see, for example, Holm *et al* (1986). One can also view the right invariant system as the *inverse* representation of the standard left invariant rigid body.

Remark. It is a remarkable fact that the dynamic rigid body equations on $SO(n)$ and indeed on any semisimple Lie group are integrable (Mishchenko and Fomenko (1976)). A key observation in this regard, due to Manakov, was that one could write the generalized rigid body equations as Lax equations with parameter

$$\frac{\mathrm{d}}{\mathrm{d}t}(M + \lambda\Lambda^2) = [M + \lambda\Lambda^2, \Omega + \lambda\Lambda], \qquad (2.4)$$

where $M = J(\Omega) = \Lambda\Omega + \Omega\Lambda$, as in section 2. 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 a generic adjoint orbit of $SO(n)$ (identified with the corresponding coadjoint orbit). (We remark that the the $SO(n)$ rigid body equations were in fact written down by F Frahm in 1874 who also proved integrability for the case $n = 4$. In addition, F Schottky in 1891 showed how to obtain explicit theta-function solutions in this case. For references to this work see Bogayavlenski (1994) and Federov and Kozlov (1995).) Moser and Veselov (1991) show that there is a corresponding formulation of the discrete rigid body equations with parameter. We shall return to this issue in the conclusion section.

3. Rigid body equations: symmetric representation

In this section we introduce a Hamiltonian system that will be related to the system (RBn) and, later, to optimal control problems. We will call this system (SRBn), standing for *the symmetric representation of the rigid body in n-dimensions.*

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

$$\dot{Q} = Q\Omega, \qquad \dot{P} = P\Omega, \tag{SRBn}$$

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

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

It is easy to check that this system of equations on the space $SO(n) \times SO(n)$ is invariant under the left diagonal action of $SO(n)$.

Proposition 3.1. *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 $M = Q^{\mathrm{T}}P - P^{\mathrm{T}}Q$ and using the equations (SRBn) gives the second of the equations (RBn).   □

It is because of this proposition that the equations (SRBn) are called the symmetric representation of the rigid body equations on $SO(n) \times SO(n)$ in left invariant form.

Recall that the *spatial angular momentum* for the standard left invariant rigid body equations (RBn) is *defined* to be the value of momentum map for the cotangent lifted left action of $SO(n)$ on $T^*SO(n)$[5].

Proposition 3.2. *For a solution of the left invariant rigid body equations (RBn) obtained by means of proposition 3.1, the spatial angular momentum is given by $m = PQ^{\mathrm{T}} - QP^{\mathrm{T}}$ and hence m is conserved along the rigid body flow.*

Proof. If we start with a solution $(Q(t), P(t))$ of the symmetric representation of the rigid body system, and map this solution to $(Q(t), M(t))$ where $M(t) = Q^{\mathrm{T}}P - P^{\mathrm{T}}Q$, then as we have seen, M satisfies the rigid body system, and so M is the body angular momentum, i.e. it is the value of the momentum map for the *right* action.

By the general Euler–Poincaré and Lie–Poisson theory, m, which is the value of the momentum map for the left action, is obtained from M using the coadjoint action of $SO(n)$ on $\mathfrak{so}(n)^* \cong \mathfrak{so}(n)$, namely $m = QMQ^{\mathrm{T}} = Q(Q^{\mathrm{T}}P - P^{\mathrm{T}}Q)Q^{\mathrm{T}} = PQ^{\mathrm{T}} - QP^{\mathrm{T}}$. From Noether's theorem, $\dot{m} = 0$; one can also verify this directly by differentiating m along (SRBn).   □

Note that in fact PQ^{T} and QP^{T} are also conserved separately along the flow.

The system (RightSRBn).   By definition, the *symmetric representation of the rigid body equations in right invariant form* on $SO(n) \times SO(n)$ 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 $M_r = P_r Q_r^{\mathrm{T}} - Q_r P_r^{\mathrm{T}}$.

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

---

[5] See, for example, Marsden and Ratiu (1999) for these basic notions.

Proposition 3.3. *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).*

In the right invariant case it follows that $m_r := Q_r^{\mathrm{T}} M_r Q_r = Q_r^{\mathrm{T}} P_r - P_r^{\mathrm{T}} Q_r$ is conserved along the flow of either (RightSRBn) or (RightRBn).

Relating the left and the right systems.

Proposition 3.4. *If $(Q(t), P(t))$ satisfies (SRBn) then the pair $(Q_r(t), P_r(t))$, where $Q_r(t) = Q(t)^{\mathrm{T}}$ and $P_r(t) = P(t)^{\mathrm{T}}$, satisfies (RightSRBn) with $\Omega_r = -\Omega = \Omega^{\mathrm{T}}$.*

This is a straightforward verification.

Local equivalence of the rigid body and the representation of the symmetric rigid body equations. Above we saw that solutions of (SRBn) can be mapped to solutions of the rigid body system. Now we consider the converse question. Thus, suppose we have a solution (Q, M) of the standard left invariant rigid body equations. We seek to solve for P in the expression

$$M = Q^{\mathrm{T}} P - P^{\mathrm{T}} Q. \tag{3.1}$$

For the following discussion, it will be convenient to make use of the operator norm on matrices. Recall that this norm is given by $\|A\|_{\mathrm{op}} = \sup\{\|Ax\| \mid \|x\| = 1\}$, where the norms on the right-hand side are the usual Euclidean space norms.

Since elements of $SO(n)$ have operator norms bounded by 1 and since the operator norm satisfies $\|AB\|_{\mathrm{op}} \leqslant \|A\|_{\mathrm{op}} \|B\|_{\mathrm{op}}$, we see that *if M satisfies $M = Q^{\mathrm{T}} P - P^{\mathrm{T}} Q$, then $\|M\|_{\mathrm{op}} \leqslant 2$. Therefore, $\|M\|_{\mathrm{op}} \leqslant 2$ is a necessary condition for solvability of (3.1) for P.*

Definition 3.5. *Let C denote the set of (Q, P) that map to M with operator norm equal to 2 and let S denote the set of (Q, P) that map to M with operator norm strictly less than 2. Also denote by S_M the set of points $(Q, M) \in T^*SO(n)$ with $\|M\|_{\mathrm{op}} < 2$. For the left invariant system we trivialize $T^*SO(n) \cong SO(n) \times \mathfrak{so}(n)^*$ by means of left translation to the identity and we identify $\mathfrak{so}(n)^*$ with $\mathfrak{so}(n)$ using the Killing metric (2.1), as earlier.*

Note that C contains pairs (Q, P) with the property that $Q^{\mathrm{T}} P$ is both skew and orthogonal.

Recall that $\sinh : \mathfrak{so}(n) \to \mathfrak{so}(n)$ is defined by $\sinh \xi = (\mathrm{e}^{\xi} - \mathrm{e}^{-\xi})/2$. One sees that indeed sinh takes values in $\mathfrak{so}(n)$ by using, for example, its series expansion:

$$\sinh \xi = \xi + \tfrac{1}{3!} \xi^3 + \tfrac{1}{5!} \xi^5 + \cdots.$$

Recall from calculus that the inverse function $\sinh^{-1}(u)$ has a convergent power series expansion for $|u| < 1$ that is given by integrating the power series expansion of the function $1/\sqrt{1 + u^2}$ term by term. This power series expansion shows that the map $\sinh : \mathfrak{so}(n) \to \mathfrak{so}(n)$ has an inverse on the set $U = \{u \in \mathfrak{so}(n) \mid \|u\|_{\mathrm{op}} < 1\}$. We shall denote this inverse, naturally, by $\sinh^{-1}$, so $\sinh^{-1} : U \to \mathfrak{so}(n)$.

Example of $SO(3)$. As an example, let us consider $\mathfrak{so}(3)$ which we parameterize as follows: we write an element of $\mathfrak{so}(3)$ as $\mu\hat{c}$ where $\hat{c}$ is an element of $\mathfrak{so}(3)$ of unit operator norm (so c, the corresponding three-vector has vector norm one) and μ is a positive scalar. One checks that the operator norm of $\hat{c}$ is equal to the Euclidean norm of c. Hence, the set U consists of the set of elements $\mu\hat{c}$ where c is a unit vector and μ is a real number with $0 \leqslant \mu < 1$. From Rodrigues' formula one finds that

$$\mathrm{e}^{\mu\hat{c}} = I + \sin(\mu)\hat{c} + (I - cc^{\mathrm{T}})(\cos \mu - 1). \tag{3.2}$$

Thus, one sees that $\sinh(\mu\hat{c}) = \sin(\mu)\hat{c}$. Notice that from this formula, sinh is not globally one to one. However, it has an inverse defined on the set U explicitly given by

$$\sinh^{-1}(\mu\hat{c}) = \sin^{-1}(\mu)\hat{c}.$$

Proposition 3.6. *For* $\|M\|_{\mathrm{op}} < 2$, *the equation (3.1) has the solution*

$$P = Q(\mathrm{e}^{\sinh^{-1} M/2}). \tag{3.3}$$

Proof. Notice that $M = \mathrm{e}^{\sinh^{-1} M/2} - \mathrm{e}^{-\sinh^{-1} M/2}$. ☐

Similarly, in the right invariant case, we obtain the formula

$$P_r = (\mathrm{e}^{\sinh^{-1} M_r/2})Q_r. \tag{3.4}$$

Example of $SO(3)$**.** We now show that for $SO(3)$ the set C is not empty, even though there are no points Q, P such that $Q^{\mathrm{T}}P$ is both skew and orthogonal (because in $SO(3)$ there are no skew orthogonal matrices, as all three by three skew matrices are singular). Let $Q^{\mathrm{T}}P = \mathrm{e}^{\mu\hat{c}}$ where $\mu = \pi/2$. Then by equation (3.2), $Q^{\mathrm{T}}P = I + \hat{c}$ and hence is not skew. Now for x such that $c^{\mathrm{T}}x = 0$ we have

$$\|(Q^{\mathrm{T}}P - P^{\mathrm{T}}Q)x\| = 2\|\hat{c}x\| = 2\|x\|,$$

and thus $\|(Q^{\mathrm{T}}P - P^{\mathrm{T}}Q)\|_{\mathrm{op}} = 2$.

In fact, reversing the argument above shows that for $SO(3)$ *the set C consists entirely of elements of form* $Q^{\mathrm{T}}P = I + \hat{c}$ *for some c.*

Proposition 3.7. *The sets C and S are invariant under the double rigid body equations.*

Proof. Notice that the operator norm is invariant under conjugation, i.e. for $Q \in SO(n)$ and $M \in \mathfrak{so}(n)$, we have $\|QMQ^{-1}\|_{\mathrm{op}} = \|M\|_{\mathrm{op}}$. This is readily checked from the definition of the operator norm. Recall that under the identification of the dual $\mathfrak{so}(n)^*$ with the space $\mathfrak{so}(n)$, the coadjoint action agrees with conjugation. Thus, the map $f : \mathfrak{so}(3) \to \mathbb{R}$; $M \mapsto \|M\|_{\mathrm{op}}$ is a Casimir function and so is invariant under the dynamics. In particular, its level sets are invariant and so the sets S and C are invariant. ☐

One can see that the operator norm is invariant under the dynamics by a direct argument as well. This is done by writing the operator norm as $\|M\|_{\mathrm{op}} = \sqrt{\lambda}$, where λ is the maximum eigenvalue of $M^{\mathrm{T}}M$ (by the Rayleigh–Ritz quotient). Then one differentiates the equation $M^{\mathrm{T}}Mv = \lambda v$ along the flow of the rigid body equations, subject to the constraint $\|v\|^2 = 1$ to see that $\dot{\lambda} = 0$.

Example of $SO(3)$**.** For the rotation group, the trace norm (up to a factor of 2) and the operator norm both agree with the standard Euclidean norm under the identification $v \in \mathbb{R}^3 \mapsto \hat{v} \in \mathfrak{so}(3)$. The standard norm is indeed a Casimir function for the rotation group and is invariant under the rigid body equations by conservation of angular momentum.

The Hamiltonian form of (SRBn). Recall that the classical rigid body equations are Hamiltonian on $T^*SO(n)$ with respect to the canonical symplectic structure on the cotangent bundle of $SO(n)$. The following result gives the corresponding theorem for (SRBn).

Proposition 3.8. *Consider the Hamiltonian system on the symplectic vector space* $\mathfrak{gl}(n) \times \mathfrak{gl}(n)$ *with the symplectic structure*

$$\Omega_{\mathfrak{gl}(n)}(\xi_1, \eta_1, \xi_2, \eta_2) = \tfrac{1}{2} \operatorname{trace} \left(\eta_2^{\mathrm{T}} \xi_1 - \eta_1^{\mathrm{T}} \xi_2 \right), \tag{3.5}$$

where (ξ_i, η_i), $i = 1, 2$ *are elements of* $\mathfrak{gl}(n) \times \mathfrak{gl}(n)$ *and Hamiltonian*

$$H(\xi, \eta) = -\tfrac{1}{8} \operatorname{trace}[(J^{-1}(\xi^{\mathrm{T}} \eta - \eta^{\mathrm{T}} \xi))(\xi^{\mathrm{T}} \eta - \eta^{\mathrm{T}} \xi)]. \tag{3.6}$$

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

Proof. We first compute the Hamiltonian vector field for the given Hamiltonian. Denote it by $X_H(\xi, \eta) = X(\xi, \eta), Y(\xi, \eta))$. Now one computes that

$$\mathbf{d}H(\xi, \eta) \cdot (\delta\xi, \delta\eta) = -\tfrac{1}{4} \operatorname{trace}[J^{-1}(\xi^{\mathrm{T}} \eta - \eta^{\mathrm{T}} \xi)((\delta\xi)^{\mathrm{T}} \eta - \eta^{\mathrm{T}} \delta\xi)]$$
$$-\tfrac{1}{4} \operatorname{trace}[J^{-1}(\xi^{\mathrm{T}} \eta - \eta^{\mathrm{T}} \xi)(\xi^{\mathrm{T}} \delta\eta - (\delta\eta)^{\mathrm{T}} \xi)]. \tag{3.7}$$

The condition that X_H be the Hamiltonian vector field, namely,

$$\Omega_{\mathfrak{gl}(n)}((X(\xi, \eta), Y(\xi, \eta)), (\delta\xi, \delta\eta)) = \mathbf{d}H(\xi, \eta) \cdot (\delta\xi, \delta\eta)$$

gives

$$X(\xi, \eta) = \xi J^{-1}(\xi^{\mathrm{T}} \eta - \eta^{\mathrm{T}} \xi), \qquad Y(\xi, \eta) = \eta J^{-1}(\xi^{\mathrm{T}} \eta - \eta^{\mathrm{T}} \xi). \tag{3.8}$$

Keeping in mind that $J^{-1}(\xi^{\mathrm{T}} \eta - \eta^{\mathrm{T}} \xi)$ is an element of $\mathfrak{so}(n)$, and that the tangent space to $SO(n) \times SO(n)$ at the point (Q, P) may be identified with $Q\mathfrak{so}(n) \times P\mathfrak{so}(n)$, we see that the Hamiltonian vector field X_H is tangent to $SO(n) \times SO(n)$ at each of its points (Q, P). Moreover, the equations

$$\dot{\xi} = \xi J^{-1}(\xi^{\mathrm{T}} \eta - \eta^{\mathrm{T}} \xi), \qquad \dot{\eta} = \eta J^{-1}(\xi^{\mathrm{T}} \eta - \eta^{\mathrm{T}} \xi) \tag{3.9}$$

become, on this submanifold, the symmetric representation of the rigid body system. $\qquad \square$

Note that the above Hamiltonian is equivalent to $H = \tfrac{1}{4}\langle J^{-1}M, M \rangle$, as in Ratiu (1980).

The symplectic structure on $S \subset SO(n) \times SO(n)$. There are two important remarks to be made about the symplectic nature of the phase space of the symmetric representation of the symmetric rigid body system.

In what follows, we will show that *the space* $SO(n) \times SO(n)$ *is not a symplectic submanifold of* $\mathfrak{gl}(n) \times \mathfrak{gl}(n)$, *but, on the other hand, the open subset* $S \subset SO(n) \times SO(n)$ *in definition 3.5 is symplectic.*

The argument for the first statement is as follows. If $SO(n) \times SO(n)$ were symplectic, then $\Omega_{\mathfrak{gl}(n)}$, when restricted to $SO(n) \times SO(n)$, would be nondegenerate. This would mean that for each $(Q, P) \in SO(n) \times SO(n)$, and each $\xi_1, \eta_1 \in \mathfrak{so}(n)$, the statement

$$\Omega_{\mathfrak{gl}(n)}((Q\xi_1, P\eta_1), (Q\xi_2, P\eta_2)) = 0,$$

for all $\xi_2, \eta_2 \in \mathfrak{so}(n)$, would imply that $\xi_1 = 0$ and $\eta_1 = 0$. Using the definition of the symplectic form, this condition becomes

$$\operatorname{trace}((P\eta_2)^{\mathrm{T}} Q\xi_1 - (P\eta_1)^{\mathrm{T}} Q\xi_2) = 0,$$

for all $\xi_2, \eta_2 \in \mathfrak{so}(n)$. This, in turn, is equivalent to the vanishing of each term separately, implying that $\xi_1 = 0$ and $\eta_1 = 0$. However, the condition that $\operatorname{trace}((P\eta_2)^{\mathrm{T}} Q\xi_1) = 0$ for all $\eta_2 \in \mathfrak{so}(n)$ is equivalent to $P^{\mathrm{T}} Q\xi_1$ being symmetric. To show that this does not generally hold, take $n = 3$ and P to be the identity. Then the condition requires that $Q\xi_1 Q = -\xi_1$ imply $\xi_1 = 0$. However, this is not true: let Q be a rotation about the z-axis through 180°, so

that Q is both symmetric and orthogonal. In this case, choosing ξ_1 to be the element of $\mathfrak{so}(3)$ corresponding to the vector $(1, 0, 0)$ we see that indeed $Q\xi_1 Q = -\xi_1$ holds; note that $Q\xi_1 Q$ just rotates the vector $(1, 0, 0)$ to its negative.

Even though the set $SO(n) \times SO(n)$ is not a symplectic submanifold, the open set S is, as the next proposition shows.

Proposition 3.9. *The set* $S \subset SO(n) \times SO(n) \subset \mathfrak{gl}(n) \times \mathfrak{gl}(n)$ *is a symplectic submanifold of* $\mathfrak{gl}(n) \times \mathfrak{gl}(n)$.

Proof. Let $(P, Q) \in S$. We need to show that
$$\Omega_{\mathfrak{gl}(n)}((Q\xi_1, P\eta_1), (Q\xi_2, P\eta_2)) = \tfrac{1}{2} \operatorname{trace}((P\eta_2)^{\mathrm{T}} Q\xi_1 - (P\eta_1)^{\mathrm{T}} Q\xi_2) = 0,$$
for all $\xi_2, \eta_2 \in \mathfrak{so}(n)$ implies $\xi_1 = 0$ and $\eta_1 = 0$. This in turn is equivalent to the vanishing of each term separately. That is, we need to show that
$$\operatorname{trace}((P\eta_2)^{\mathrm{T}} Q\xi_1) = \operatorname{trace}\left(\eta_2^{\mathrm{T}} P^{\mathrm{T}} Q\xi_1\right) = 0,$$
for all $\eta_2 \in \mathfrak{so}(n)$ implies $\xi_1 = 0$. Since η_2 is skew symmetric and arbitrary, this is equivalent to the condition that $P^{\mathrm{T}} Q\xi_1$ being symmetric implies $\xi_1 = 0$. The matrix $P^{\mathrm{T}} Q\xi_1$ is symmetric when $Q^{\mathrm{T}} P\xi_1 Q^{\mathrm{T}} P = -\xi_1$, which by equation (3.3) and the fact that $(Q, P) \in S$ is the same as $\mathrm{e}^{\sinh^{-1} M/2} \xi_1 \mathrm{e}^{\sinh^{-1} M/2} = -\xi_1$. Thus, the proof of the proposition will be complete if we prove the following lemma. $\square$

Lemma 3.10. *For* $A, B \in \mathfrak{so}(n)$, *if* $\|A\|_{\mathrm{op}} < 1$ *and*
$$\mathrm{e}^{\sinh^{-1} A} B \mathrm{e}^{\sinh^{-1} A} = -B, \tag{3.10}$$
then $B = 0$.

Proof. Since A is a normal matrix with pure imaginary eigenvalues, it can be written in the form $A = Q\Lambda Q^{\mathrm{T}}$, where Q is orthogonal and $\Lambda = i\operatorname{diag}\lambda$. Set $\tilde{B} = Q^{\mathrm{T}} B Q$ and then (3.10) implies
$$D\tilde{B}D = -\tilde{B}, \qquad \text{where } D = \mathrm{e}^{\sinh^{-1}\Lambda} = \operatorname{diag} d.$$
This is equivalent to $(1 + d_k d_l)\tilde{B}_{k,l} = 0$, $k, l = 1, \dots, n$. Note that $d_k = \mathrm{e}^{\sinh^{-1} i\lambda_k} = \mathrm{e}^{i\sin^{-1}\lambda_k}$. Recall however that $\|A\|_{\mathrm{op}} < 1$ which, for normal matrices, is equivalent to $\rho(A) < 1$. In other words, $|\lambda_k| < 1$ for all k, and this implies that $\sinh^{-1}(i\lambda_k) = \sin^{-1}\lambda_k \in (-\pi/2, \pi/2)$. Suppose now that $1 + d_k d_l = 0$ for some k and l. Then $\mathrm{e}^{i(\sin^{-1}\lambda_k + \sin^{-1}\lambda_l)} = \mathrm{e}^{\pm i\pi}$; in other words $\sin^{-1}\lambda_k + \sin^{-1}\lambda_l = \pm\pi$. This is a contradiction, hence $1 + d_k d_l \neq 0$ for all $k, l = 1, 2, \dots, n$. This implies $\tilde{B} = O$ and, since Q is nonsingular, $B = O$. This is the proof[6] of lemma 3.10. $\square$

The pull-back symplectic structure. We now show that the pull back of the canonical symplectic structure on $T^*SO(n)$ to $SO(n) \times SO(n)$ gives the symplectic structure on the space S. More precisely, we have the following proposition.

Proposition 3.11. *Let* $\Psi : (Q, P) \mapsto (Q, M) = (Q, Q^{\mathrm{T}} P - P^{\mathrm{T}} Q)$ *denote the map from* $SO(n) \times SO(n)$ *to* $T^*SO(n) \cong SO(n) \times \mathfrak{so}(n)$ *whose restriction to* S *is 1–1 and onto* S_M. *Then the restriction of this map to* S *is a diffeomorphism between* S *and* S_M.

Further, let ω *denote the canonical symplectic form on* $T^*SO(n)$ *and let* $\Omega_{\mathfrak{gl}(n)}|_S$ *denote the restriction of the form (3.5) to* S. *Then*
$$\Psi^*(\omega|_{S_M}) = 2\Omega_{\mathfrak{gl}(n)}|_S. \tag{3.11}$$

---

[6] We thank Arieh Iserles for informing us of this short proof.

Proof. Firstly we prove that the restriction of the map Ψ to S is a diffeomorphism between S and S_M. The 1–1 nature of the map is proved using similar arguments to lemma 3.10 and the map is onto by construction.

Now we prove differentiability of the inverse map by the proving that the Jacobian of the inverse map is nonsingular on S_M. This map is given by (3.3) and hence the derivative of P with respect to M is

$$\frac{1}{2} Q(\mathrm{e}^{\sinh^{-1} M/2}) \left(1 + \frac{M^2}{4}\right)^{-1/2}, \tag{3.12}$$

where the last factor is interpreted via a power series. Recall that $\sinh M/2$ is nonsingular on S_M by our earlier arguments (see the discussion following definition (3.5)), so $\mathrm{e}^{\sinh^{-1} M/2}$ is defined. The first factor in (3.12), $\frac{1}{2} Q(\mathrm{e}^{\sinh^{-1} M/2})$, is the product of two nonsingular factors, so is obviously nonsingular.

$I + M^2/4$ is nonsingular and positive on S_M by similar arguments used in lemma 3.10.

Now we turn to the proof of equation (3.11). Let $(Q\xi_1, P\eta_1)$ and $(Q\xi_2, P\eta_2)$, $\xi_i, \eta_i \in \mathfrak{so}(n)$ be tangent vectors to $SO(n) \times SO(n)$. Then by (3.5), we have

$$\Omega|_S((Q\xi_1, P\eta_1), (Q\xi_2, P\eta_2)) = \tfrac{1}{2} \operatorname{trace}((P\eta_2)^{\mathrm{T}} Q\xi_1 - (P\eta_1)^{\mathrm{T}} Q\xi_2)$$
$$= \tfrac{1}{2} \operatorname{trace}(P^{\mathrm{T}} Q)(-\xi_1\eta_2 + \xi_2\eta_1).$$

The canonical symplectic form on $T^*SO(n)$ left trivialized to $SO(n) \times \mathfrak{so}(n)$ is given by (see, e.g. Abraham and Marsden (1978), proposition 4.4.2ii, p 316)

$$\omega_{(Q,M)}((V_1, Z_1)(V_2, Z_2)) = \langle Z_2, Q^{-1}V_1 \rangle - \langle Z_1, Q^{-1}V_2 \rangle + \langle M, [Q^{-1}V_1, Q^{-1}V_2] \rangle$$
$$= -\tfrac{1}{2} \operatorname{trace}(Z_2 Q^{-1}V_1) + \tfrac{1}{2} \operatorname{trace}(Z_1 Q^{-1}V_2)$$
$$- \tfrac{1}{2} \operatorname{trace}(M[Q^{-1}V_1, Q^{-1}V_2]), \tag{3.13}$$

where $V_i \in T_Q SO(n)$ (identified with $T_Q^* SO(n)$) and $Z_i \in T_M\mathfrak{so}(n) \equiv \mathfrak{so}(n)$ and where we used the pairing (2.1).

Observing that the derivative of Ψ is given by

$$T\Psi : (Q\xi, P\eta) \mapsto (Q\xi, -\xi Q^{\mathrm{T}} P + Q^{\mathrm{T}} P\eta + \eta P^{\mathrm{T}} Q - P^{\mathrm{T}} Q\xi) \tag{3.14}$$

and substituting in (3.13) with $V_i = Q\xi_i$ and $Z_i = -\xi_i Q^{\mathrm{T}} P + Q^{\mathrm{T}} P\eta_i + \eta_i P^{\mathrm{T}} Q - P^{\mathrm{T}} Q\xi_i$, we obtain

$$\Psi^*\omega_{(Q,P)}((V_1, Z_1), (V_2, Z_2)) = -\tfrac{1}{2} \operatorname{trace}\{(-\xi_2 Q^{\mathrm{T}} P + Q^{\mathrm{T}} P\eta_2 + \eta_2 P^{\mathrm{T}} Q - P^{\mathrm{T}} Q\xi_2)\xi_1\}$$
$$+ \tfrac{1}{2} \operatorname{trace}\{(-\xi_1 Q^{\mathrm{T}} P + Q^{\mathrm{T}} P\eta_1 + \eta_1 P^{\mathrm{T}} Q - P^{\mathrm{T}} Q\xi_1)\xi_2\}$$
$$- \tfrac{1}{2} \operatorname{trace}\{(Q^{\mathrm{T}} P - P^{\mathrm{T}} Q)(\xi_1\xi_2 - \xi_2\xi_1)\}$$
$$= \operatorname{trace}\{(P^{\mathrm{T}} Q)(\xi_2\eta_1 - \xi_1\eta_2)\},$$

which gives the result. $\qquad\square$

There are similar results in the right invariant case. Summarizing the above arguments we have the following theorem.

Theorem 3.12. *The symmetric representation of the rigid body equations (SDRBn) on the invariant set S is equivalent to the rigid body equations (RBn) on the invariant set S_M where S and S_M are defined in definition 3.5. Equivalence means that every trajectory of (SRBn) lying in S is mapped by Ψ onto a trajectory of (RBn) lying in S_M and similarly in the reverse direction under the map Ψ^{-1}. Further, Ψ is a symplectomorphism between the symplectic manifolds S and S_M.*

Scaling. The rigid body equations are homogeneous. Precisely, if $M(t)$, $Q(t)$ is a solution of the rigid body system, then so is $\epsilon M(\epsilon t)$, $Q(\epsilon t)$. To check this statement, one needs to show that $\bar{M}(t) := \epsilon M(\epsilon t)$, $\bar{Q}(t) := Q(\epsilon t)$ satisfies $\dot{\bar{Q}} = \bar{Q}\bar{\Omega}$ and $\dot{\bar{M}} = [\bar{M}, \bar{\Omega}]$, where $\bar{\Omega} := \bar{Q}^{-1}\dot{\bar{Q}} \in \mathfrak{so}(n)$ and where $\bar{M} := J(\bar{\Omega}) = \Lambda\bar{\Omega} + \bar{\Omega}\Lambda \in \mathfrak{so}(n)$. Differentiating, using the chain rule, $\dot{\bar{M}}(t) = \epsilon^2 \dot{M}(\epsilon t)$ and $\dot{\bar{Q}}(t) = \epsilon \dot{Q}(\epsilon t)$. Note that

$$\bar{\Omega}(t) = \bar{Q}(t)^{-1}\dot{\bar{Q}}(t) = Q^{-1}(\epsilon t)\epsilon \dot{Q}(\epsilon t) = \epsilon \Omega(\epsilon t).$$

Substitute these in the above equation for $\bar{M}$ to see it remains satisfied—each side picks up a factor of ϵ^2, which cancels. Similarly we verify the equation for $\bar{Q}$ and the relation between $\bar{M}$ and $\bar{\Omega}$.

From the above we have the following proposition.

Proposition 3.13. *Given a solution of (RBn) outside the set S_M, there is an associated scaled solution that lies in the set S_M and hence can be realized as the image under the map Ψ of a solution of (SRBn).*

The proof follows from the fact that the operator norm is dynamically invariant.

Dirac bracket. We shall now give the Poisson bracket on the open set $S \subset SO(n) \times SO(n)$, which is obtained by means of the Dirac bracket formula.

Proposition 3.14. *Let F, K be smooth real valued functions on $\mathfrak{gl}(n) \times \mathfrak{gl}(n)$ and $F|S$, $K|S$ be their restrictions to the open set $S \subset SO(n) \times SO(n)$. Then their Poisson bracket at a point $(Q, P) \in S$ is given by*

$$
\begin{aligned}
\{F|S, K|S\} = &\langle \nabla_2 K, \nabla_1 F \rangle - \langle \nabla_1 K, \nabla_2 F \rangle \\
&- \tfrac{1}{2}\langle Q(\nabla_2 K)^{\mathrm{T}} + (\nabla_2 K)Q^{\mathrm{T}}, (I + R \otimes R^{\mathrm{T}})^{-1}R(P(\nabla_1 F)^{\mathrm{T}} + (\nabla_1 F)P^{\mathrm{T}})\rangle \\
&+ \tfrac{1}{2}\langle P(\nabla_1 K)^{\mathrm{T}} + (\nabla_1 K)P^{\mathrm{T}}, (I + R \otimes R^{\mathrm{T}})^{-1}(Q(\nabla_2 F)^{\mathrm{T}} + (\nabla_2 F)Q^{\mathrm{T}})R\rangle,
\end{aligned}
$$

(3.15)

where $R = QP^{\mathrm{T}}$ and ∇_1 and ∇_2 are gradients on $\mathfrak{gl}(n)$ with respect to the standard metric with respect to the first and second factor, respectively.

This is seen by using the general Dirac bracket formula (see, e.g. Marsden and Ratiu (1999), proposition 8.5.1). In fact, this formula may be obtained by using the constraint submersion $\psi : \mathfrak{gl}(n) \times \mathfrak{gl}(n) \to \mathrm{Sym} \times \mathrm{Sym}$, where Sym is the space of symmetric matrices, given by $(A, B) \mapsto (AA^{\mathrm{T}}, BB^{\mathrm{T}})$ so that S is an open subset of $\mathrm{O}(n) \times \mathrm{O}(n) = \psi^{-1}(\mathrm{Id} \times \mathrm{Id})$. One then uses formula (8.5.12) of Marsden and Ratiu (1999) to do the computation.

Remark.

(a) The matrix $I + R \otimes R^{\mathrm{T}}$ has a well-defined inverse precisely on the set S of interest by the arguments in lemma 3.10.
(b) In the Dirac bracket formula (3.15), the matrices $R(P(\nabla_1 F)^{\mathrm{T}} + (\nabla_1 F)P^{\mathrm{T}})$ and $(Q(\nabla_2 F)^{\mathrm{T}} + (\nabla_2 F)Q^{\mathrm{T}})R$ should be viewed as n^2-vectors (which are multiplied on the left by $n^2 \times n^2$ matrices).
(c) As with the general Dirac formula, for F chosen to be our Hamiltonian (3.6), the last two terms of the Dirac bracket vanish since our flow is tangent to the manifold $SO(n) \times SO(n)$.

4. Moser–Veselov and the symmetric representation of the discrete rigid body

In this section we discuss some aspects of discrete aspects of our problem following the treatment of Veselov (1988) and Moser and Veselov (1991). This general method is

closely related to the development of variational integrators for the integration of mechanical systems, as in Marsden *et al* (1999) and Kane *et al* (2000). Another approach to integrating differential equations on manifolds is discussed in Crouch and Grossman (1993). See also Iserles *et al* (1999), Budd and Iserles (1999) and Bobenko and Suris (1999).

Review of the Moser–Veselov discrete rigid body.   We briefly review the Moser and Veselov (1991) discrete rigid body equations, a system that will be called (DRBn). Discretize the configuration matrix and let $Q_k \in SO(n)$ denote the rigid body configuration at time k, let $\Omega_k \in SO(n)$ 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:

$$\Omega_k = Q_k^{\mathrm{T}} Q_{k-1} \tag{4.1}$$

$$M_k = \Omega_k^{\mathrm{T}} \Lambda - \Lambda \Omega_k \tag{4.2}$$

$$M_{k+1} = \Omega_k M_k \Omega_k^{\mathrm{T}}. \tag{4.3}$$

$$\text{(DRBn)}$$

These equations may be viewed as defining two different algorithms.

MV-algorithm 1. Define the step ahead map

$$(Q_k, Q_{k+1}) \mapsto (Q_{k+1}, Q_{k+2}) \tag{4.4}$$

as follows: compute Ω_{k+1} from (4.1), compute M_{k+1} from (4.2), compute M_{k+2} from (4.3), compute Ω_{k+2} from (4.2) and then compute Q_{k+2} from (4.1).

Remark. Given M_k, conditions under which equation (4.2) is solvable for Ω_k are discussed in Moser and Veselov (1991) and Cardoso and Leite (2001). We will return to this point later.

MV-algorithm 2. Define the map

$$(Q_k, M_k) \mapsto (Q_{k+1}, M_{k+1}) \tag{4.5}$$

as follows: compute Ω_k from (4.2), compute M_{k+1} from (4.3), compute Ω_{k+1} from (4.2) and compute Q_{k+1} from (4.1).

Discrete variational principle.   The Moser–Veselov equations (4.1)–(4.3) can be obtained by a discrete variational principle, as was done in Moser and Veselov (1991). This variational principle has the general form of that in discrete mechanics described in, for example, Marsden and Wendlandt (1997), Bobenko and Suris (1999) and Marsden and West (2001). See also the following sections on optimal control. Namely, stationary points of the functional

$$S = \sum_k \operatorname{trace}(Q_k \Lambda Q_{k+1}^{\mathrm{T}}) \tag{4.6}$$

on sequences of orthogonal $n \times n$ matrices give the Moser–Veselov equations.

This variational approach can be justified as in Marsden *et al* (1999). We shall justify it here from the optimal control point of view in section 5. We consider the *left* invariant generalized rigid body equations on $SO(n)$.

The symmetric representation of the discrete rigid body.   We now define the symmetric representation of the discrete rigid body equations as follows:

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

where U_k is defined by

$$U_k \Lambda - \Lambda U_k^{\mathrm{T}} = Q_k^{\mathrm{T}} P_k - P_k^{\mathrm{T}} Q_k. \tag{4.7}$$

We will write this as

$$J_D U_k = Q_k^{\mathrm{T}} P_k - P_k^{\mathrm{T}} Q_k, \tag{4.8}$$

where $J_D : SO(n) \to \mathfrak{so}(n)$ (the discrete version of J) is defined by $J_D U = U\Lambda - \Lambda U^{\mathrm{T}}$. 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 neighbourhood of the identity in $SO(n)$ to a neighbourhood 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 (4.7), compute Q_{k+1} and P_{k+1} using (SDRBn). Note that the update map for Q and P is done in parallel.

Proposition 4.1. *The symmetric representation of the discrete rigid body equations (SDRBn) on the invariant set S is equivalent to the Moser–Veselov equations (4.1)–(4.3) (DRBn) on the invariant set S_M where S and S_M are defined in definition 3.5. Equivalence means that every trajectory of (SDRBn) lying in S is mapped by Ψ onto a trajectory of (DRBn) lying in S_M and similarly in the reverse direction under the map Ψ^{-1}. Further, Ψ is a symplectomorphism between the symplectic manifolds S and S_M.*

Proof. Suppose that we have a solution (Q_k, P_k) to (SDRBn). We will now produce a solution (Q_{k+1}, M_{k+1}) of the Moser–Veselov equations. We claim that

$$M_{k+1} = Q_k^{\mathrm{T}} P_k - P_k^{\mathrm{T}} Q_k \tag{4.9}$$

will give us the required M_{k+1} that does the job. To see this, let

$$\Omega_{k+1} = Q_{k+1}^{\mathrm{T}} Q_k = U_k^{\mathrm{T}}. \tag{4.10}$$

Now substitute (4.10) into (4.7) and use (4.9) to give (4.2) with k replaced by $k + 1$. Next, substitute the equations (SDRBn) into (4.9) with k replaced by $k + 1$ to yield (4.3) with k replaced by $k + 1$. Clearly (4.1) with k replaced by $k + 1$ is the same as (4.10). Thus, we have shown that (SDRBn) imply the Moser–Veselov equations.

The following remark will be useful for what follows. Recall from MV-algorithm 2 that $M_{k+1} = \Omega_k M_k \Omega_k^{\mathrm{T}}$, so M_{k+1} is obtained from M_k by conjugation, so has the same operator norm. Thus, MV-algorithm 2 leaves the set S_M invariant, as in the continuous rigid body equations (RBn). By the first part of this proof, it follows that the system (SDRBn) leaves the set S invariant.

To prove the converse, assume we have a solution $(Q_{k+1}, M_{k+1}) \in S_M$ of the Moser–Veselov equations. Note that because $(Q_{k+1}, M_{k+1}) \in S_M$, we can solve equation (4.9) for P_k, as in the continuous case, to give

$$P_k = Q_k \mathrm{e}^{\sinh^{-1} M_{k+1}/2}. \tag{4.11}$$

This then gives us a sequence (Q_k, P_k), which we claim satisfies the system (SDRBn). To see this, we note from (4.1) that $Q_{k+1} = Q_k \Omega_{k+1}^{\mathrm{T}}$. We need to show that $\Omega_{k+1}^{\mathrm{T}}$ satisfies the defining equation (4.7) for U_k. That is, we must show

$$\Omega_{k+1}^{\mathrm{T}} \Lambda - \Lambda \Omega_{k+1} = Q_k^{\mathrm{T}} P_k - P_k^{\mathrm{T}} Q_k. \tag{4.12}$$

That is, in view of (4.2),

$$M_{k+1} = Q_k^{\mathrm{T}} P_k - P_k^{\mathrm{T}} Q_k. \tag{4.13}$$

But this is valid since P_k was chosen to satisfy this equation. Therefore, the first equation in (SDRBn) holds and we have shown that $\Omega_{k+1}^{\mathrm{T}} = U_k$.

To prove the second equation of (SDRBn), we proceed as follows. We have

$$P_{k+1} = Q_{k+1}e^{\sinh^{-1} M_{k+2}/2} = Q_k U_k e^{\sinh^{-1} M_{k+2}/2} \tag{4.14}$$

using the first of (SDRBn). Using (4.3), this becomes

$$P_{k+1} = Q_k U_k \exp\left(\sinh^{-1} U_k^{\mathrm{T}} M_{k+1} \frac{U_k}{2}\right) = Q_k \exp\left(\sinh^{-1} \frac{M_{k+1}}{2}\right) U_k = P_k U_k. \tag{4.15}$$

$$\square$$

Note that if we define $m_{k+1} = P_k Q_k^{\mathrm{T}} - Q_k P_k^{\mathrm{T}}$, then $m_{k+1} = Q_k M_{k+1} Q_k^{\mathrm{T}}$. Thus m_k may be interpreted as a discrete analog of the spatial momentum and from (4.3) this is conserved under the algorithm.

Convergence of the discrete system to the continuous system. We now show how to obtain the representation of the symmetric rigid body equations (in left invariant form) (SRBn) from their discrete counterpart (SDRBn). The key to doing this is the introduction of a time step h. It is interesting that the second-order Moser–Veselov equations (4.1)–(4.3) do not explicitly involve a time step—the time step is determined by the fact that one needs to specify initial data at two time points: Q_0 and Q_1 say. Formally, the two points determine a velocity field using a time step h, as we explained above in the discretization of ξ.

We define U_k^h by

$$U_k^h = J_D^{-1}\left(h\left(Q_k^{\mathrm{T}} P_k - P_k^{\mathrm{T}} Q_k\right)\right). \tag{4.16}$$

We also define

$$\Omega = \lim_{h \to 0}\left(\frac{U_k^h - \mathrm{Id}}{h}\right), \tag{4.17}$$

where Id denotes the identity. Then we have the following theorem.

Theorem 4.2. *Taking the derivative with respect to h in (SDRBn) yields (SRBn).*

Proof. Using (SDRBn), we have

$$\frac{Q_{k+1} - Q_k}{h} = Q_k\left(\frac{U_k^h - I}{h}\right), \qquad \frac{P_{k+1} - P_k}{h} = P_k\left(\frac{U_k^h - I}{h}\right). \tag{4.18}$$

Taking the limit on both sides with respect to h yields (SRBn) subject to checking that the formula for U_k^h tends to that for Ω. This is a consequence of the following computation (using (4.16)):

$$\lim_{h \to 0} \frac{\left(U_k^h - \mathrm{Id}\right)\Lambda - \Lambda\left(U_k^h - \mathrm{Id}\right)^{\mathrm{T}}}{h} = \lim_{h \to 0} \frac{h\left(Q_k^{\mathrm{T}} P_k - P_k^{\mathrm{T}} Q_k\right)}{h}. \tag{4.19}$$

Taking the limit we obtain

$$\Omega\Lambda - \Lambda\Omega^{\mathrm{T}} = Q^{\mathrm{T}} P - P^{\mathrm{T}} Q, \qquad \text{i.e. } \Omega\Lambda + \Lambda\Omega = Q^{\mathrm{T}} P - P^{\mathrm{T}} Q \tag{4.20}$$

as desired. In taking this limit we write $Q_k = Q(kh)$, where $kh = t$ and similarly for P_k. $\square$

5. Optimal control

In this section we briefly review, from Bloch and Crouch (1996), two results which link the above formulation of the symmetric representation of the rigid body equations with the theory

of optimal control. The first result shows how to derive these equations directly from an optimal control problem and the second shows how to derive them as a special case of a more general optimal control problem. The latter problem is characterized by extremals which are of the coupled double bracket form studied in Bloch *et al* (1997).

Definition 5.1. *Let $T > 0$, Q_0, $Q_T \in SO(n)$ be given and fixed. Let the rigid body optimal control problem be given by*

$$\min_{U \in \mathfrak{so}(n)} \frac{1}{4} \int_0^T \langle U, J(U) \rangle \, \mathrm{d}t, \tag{5.1}$$

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

$$\dot{Q} = QU, \qquad Q(0) = Q_0, \qquad Q(T) = Q_T. \tag{5.2}$$

Proposition 5.2. *The rigid body optimal control problem (5.1) 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

$$U = J^{-1}(Q^\mathrm{T} P - P^\mathrm{T} Q). \tag{5.3}$$

Remark. The proof (see Bloch and Crouch (1996)) simply involves writing the Hamiltonian of the maximum principle (see e.g. Bryson and Ho (1975)) as

$$H = \langle P, QU \rangle + \tfrac{1}{4} \langle U, J(U) \rangle, \tag{5.4}$$

where the costate vector P is a multiplier enforcing the dynamics, and then maximizing with respect to U in the standard fashion (see, e.g. Brockett (1973)). While in general there are no constraints on the costate vector $P \in \mathfrak{gl}(n)$, one can consider the restriction of the extremal flows to invariant submanifolds. This limits possible extremal trajectories that can be recovered. For example (SRBn) restricts to a system on $SO(n) \times SO(n)$. One can make other assumptions on the costate vector. For example, suppose we assume a costate vector B such that $Q^\mathrm{T} B$ is skew. Then it is easy to check that the extremal evolution equations become

$$\dot{Q} = QJ^{-1}(Q^\mathrm{T} B), \qquad \dot{B} = BJ^{-1}(Q^\mathrm{T} B), \tag{5.5}$$

and that these equations restrict to an invariant submanifold defined by the condition that $Q^\mathrm{T} B$ is skew symmetric. These are the McLachlan–Scovel equations (McLaghlan and Scovel 1995). Comparing these equations with (SRBn) we see that $B = P - QP^\mathrm{T}Q$. There is a similar result for the right invariant case.

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. In addition, as we shall see, this sheds light on the question of integrability of the *n*-dimensional rigid body.

We begin by recalling a general optimal control problem on matrices (see Bloch and Crouch (1996)).

Definition 5.3. *Let $\mathfrak{u}(n)$ denote the Lie algebra of the unitary group $U(n)$.*

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 $\langle A, B \rangle = \frac{1}{2} \mathrm{trace}(A^\dagger B)$, where $A^\dagger$ is the adjoint, i.e. the transpose conjugate.

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

$$\min_{U,V} \frac{1}{4} \int \{\langle U, J_U U\rangle + \langle V, J_V V\rangle\}\, dt \tag{5.6}$$

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

$$\dot{Q} = UQ - QV, \qquad Q(0) = Q_0, \qquad Q(T) = Q_T. \tag{5.7}$$

This problem was motivated by an optimal control problem on adjoint orbits of compact Lie groups as discussed in Brockett (1994).

Theorem 5.4. *The optimal control problem (5.3) has optimal controls given by*

$$U = J_U^{-1}(PQ^\dagger - QP^\dagger), \qquad V = J_V^{-1}(P^\dagger Q - Q^\dagger P), \tag{5.8}$$

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

$$\begin{aligned}
\dot{Q} &= J_U^{-1}(PQ^\dagger - QP^\dagger)Q - QJ_V^{-1}(P^\dagger Q - Q^\dagger P), \\
\dot{P} &= J_U^{-1}(PQ^\dagger - QP^\dagger)P - PJ_V^{-1}(P^\dagger Q - Q^\dagger P).
\end{aligned} \tag{5.9}$$

Note also that J_U and J_V are in general different operators acting on different spaces. In certain case (see the rigid body below) the spaces and the operators may be taken to be the same.

Corollary 5.5. *Equations (5.9) are given by the coupled double bracket equations*

$$\dot{\hat{Q}} = [\hat{Q}, \hat{J}^{-1}[\hat{P}, \hat{Q}]], \qquad \dot{\hat{P}} = [\hat{P}, \hat{J}^{-1}[\hat{P}, \hat{Q}]], \tag{5.10}$$

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

$$\hat{Q} = \begin{bmatrix} 0 & Q \\ -Q^\dagger & 0 \end{bmatrix} \in \mathfrak{u}(p+q) \tag{5.11}$$

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

A formal limiting argument setting $J_V = J$ and $J_U \to \infty$ gives the symmetric representation of the rigid body equation in left invariant form. Similarly to obtain the equations in their right invariant form set $J_U = J$ and let $J_V \to \infty$. One sees in fact that equations (5.9) are literally the sum of the symmetric representations of the rigid body equations in their left and right invariant forms.

6. Discrete optimal control problems

One can obtain the symmetric representation of the rigid body equations as a special case of a general class of discrete optimal control equations.

Definition 6.1. *Let N be a positive integer and X_0, $X_N \in \mathbb{R}^n$ be given. Let $f(x_k, u_k)$, $g(x_k, u_k)$ be smooth mappings from $\mathbb{R}^n \times \mathbb{R}^m$ into $\mathbb{R}^n$ and $\mathbb{R}^+$, respectively. Let E denote a control constraint set with $E \subset \mathbb{R}^m$; specifically, assume that E is defined as the zero set of a smooth submersion $k : \mathbb{R}^m \to \mathbb{R}^l$, i.e. $u \in E$ if and only if $k(u) = 0$. Let $\langle \cdot, \cdot \rangle$ denote the pairing between vectors in $\mathbb{R}^n$.*

Define the optimal control problem:

$$\min_{u_k \in E} \sum_{k=0}^{N} g(x_k, u_k) \tag{6.1}$$

subject to $x_{k+1} = f(x_k, u_k)$, with $x_0 = X_0$ and $x_N = X_N$, for $u_k \in E$.

Proposition 6.2. *A solution to problem (6.1) satisfies the following extremal solution equations*

$$p_k = \frac{\partial H}{\partial x_k}\left(p_{k+1}, x_k, u_k^*\right), \qquad x_{k+1} = \frac{\partial H}{\partial p_{k+1}}\left(p_{k+1}, x_k, u_k^*\right), \tag{6.2}$$

where

$$H(p_{k+1}, x_k, u_k) = \langle p_{k+1}, f(x_k, u_k)\rangle - g(x_k, u_k). \tag{6.3}$$

In these equations, u_k^* is determined as follows: let

$$\hat{H}(p_{k+1}, x_k, u_k, \sigma) = H(p_{k+1}, x_k, u_k) + \langle \sigma, k(u_k)\rangle, \tag{6.4}$$

for $\sigma \in \mathbb{R}^l$, σ a Lagrange multiplier; then u_k^* and σ are solutions of the equations:

$$\frac{\partial \hat{H}}{\partial u_k}\left(p_{k+1}, x_k, u_k^*, \sigma\right) = 0, \qquad k\left(u_k^*\right) = 0. \tag{6.5}$$

The proof is an application of the discrete maximum principle (see appendix A).

We assume that both u_k^* and σ are determined uniquely by equations (6.5). Also note that $u_k^* = u_k^*(p_{k+1}, x_k)$, $\sigma = \sigma(p_{k+1}, x_k)$. With this notation, we have the following consequence.

Corollary 6.3. *Assume that the extremal equations (6.2) hold. Then,*

$$\begin{aligned}
\mathrm{d}p_k &= \frac{\partial^2 \hat{H}}{\partial x_k^2}\left(p_{k+1}, x_k, u_k^*\right)\mathrm{d}x_k + \frac{\partial^2 \hat{H}}{\partial p_{k+1}\partial x_k}\left(p_{k+1}, x_k, u_k^*\right)\mathrm{d}p_{k+1}, \\
\mathrm{d}x_{k+1} &= \frac{\partial^2 \hat{H}}{\partial p_{k+1}\partial x_k}\left(p_{k+1}, x_k, u_k^*\right)\mathrm{d}x_k + \frac{\partial^2 \hat{H}}{\partial p_{k+1}^2}\left(p_{k+1}, x_k, u_k^*\right)\mathrm{d}p_{k+1}
\end{aligned} \tag{6.6}$$

We remark that the implicit advance map $\Phi : (x_k, p_k) \mapsto (x_{k+1}, p_{k+1})$ generated by the extremal evolution (6.2) is symplectic, i.e.

$$\Phi^*\mathrm{d}x_{k+1} \wedge \mathrm{d}p_{k+1} = \mathrm{d}x_k \wedge \mathrm{d}p_k. \tag{6.7}$$

This is easily demonstrated by using corollary 6.3. One can also derive symplecticity directly from Hamilton's phase space principle (see Marsden and West (2001)).

We can then obtain the discrete rigid body (SDRBn) equations as follows.

Definition 6.4. *Let Λ be a positive definite diagonal matrix. Let $\bar{Q}_0, \bar{Q}_N \in SO(n)$ be given and fixed. Let*

$$\hat{V} = \sum_{k=1}^{N} \mathrm{trace}(\Lambda U_k). \tag{6.8}$$

Define the optimal control problem,

$$\min_{U_k} \hat{V} = \min_{U_k} \sum_{k=1}^{N} \mathrm{trace}(\Lambda U_k), \tag{6.9}$$

subject to dynamics and initial and final data

$$Q_{k+1} = Q_k U_k, \qquad Q_0 = \bar{Q}_0, \qquad Q_N = \bar{Q}_N, \tag{6.10}$$

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

Theorem 6.5. *A solution of the optimal control problem (6.4) satisfies the optimal evolution equations (SDRBn)*

$$Q_{k+1} = Q_k U_k, \qquad P_{k+1} = P_k U_k, \tag{6.11}$$

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

$$U_k\Lambda - \Lambda U_k^\mathrm{T} = Q_k^\mathrm{T} P_k - P_k^\mathrm{T} Q_k. \tag{6.12}$$

Equation (6.12) can be solved for U_k under certain circumstances, as discussed in Moser and Veselov (1991) and Cardoso and Leite (2001); we discuss this issue further below.

Proof. Applying proposition 6.2, we get

$$H(P_{k+1}, Q_k, U_k) = \text{trace}\left(P_{k+1}^{\mathrm{T}} Q_k U_k\right) - \text{trace}\left(\Lambda U_k\right) = \text{trace}\left(\left(P_{k+1}^{\mathrm{T}} Q_k - \Lambda\right) U_k\right). \quad (6.13)$$

Note that $\hat{V} = \sum_k \text{trace}(\Lambda U_k) = \sum_k \text{trace}) = \sum_k \text{trace}(Q_k \Lambda Q_{k+1}^{\mathrm{T}})$, the Moser–Veselov functional, but that the functional is *linear* in the controls.

We need to find the critical points of $H(P_{k+1}, Q_k, U_k)$ where $U_k^{\mathrm{T}} U_k = I$ since $U_k \in SO(n)$. Thus, we need to minimize a functional of the form $\text{trace}(AU)$, A fixed, subject to $U^{\mathrm{T}} U = I$. Set

$$\tilde{V} = \text{trace}(AU) + \tfrac{1}{2}\,\text{trace}(\Sigma(U^{\mathrm{T}} U - I)),$$

where $\Sigma = \Sigma^{\mathrm{T}}$ is a matrix of Lagrange multipliers. Then $\delta\tilde{V} = \text{trace}(A\delta U + \Sigma U^{\mathrm{T}} \delta U) = 0$ implies $A + \Sigma U^{\mathrm{T}} = 0$ where $U^{\mathrm{T}} U = 0$. Hence $\Sigma = -AU$. But since $\Sigma = \Sigma^{\mathrm{T}}$, the extrema of our optimization problem are obtained when $AU = U^{\mathrm{T}} A^{\mathrm{T}}$. Applying this observation to our case, we have $\nabla_U H = 0$ when

$$\left(P_{k+1}^{\mathrm{T}} Q_k - \Lambda\right) U_k = U_k^{\mathrm{T}} \left(Q_k^{\mathrm{T}} P_{k+1} - \Lambda\right),$$

i.e.

$$U_k^{\mathrm{T}} \Lambda - \Lambda U_k = U_k^{\mathrm{T}} Q_k^{\mathrm{T}} P_{k+1} - P_{k+1}^{\mathrm{T}} Q_k U_k$$

or, equivalently

$$U_k \Lambda - \Lambda U_k^{\mathrm{T}} = -Q_k^{\mathrm{T}} P_{k+1} U_k^{\mathrm{T}} + U_k P_{k+1}^{\mathrm{T}} Q_k. \quad (6.14)$$

Also,

$$P_k = \nabla_{Q_k} H = \left(U_k P_{k+1}^{\mathrm{T}}\right)^{\mathrm{T}} = P_{k+1} U_k^{\mathrm{T}}. \quad (6.15)$$

Hence we obtain equations (6.11). Combining (6.11) with (6.14), we get

$$U_k \Lambda - \Lambda U_k^{\mathrm{T}} = P_k^{\mathrm{T}} Q_k - Q_k^{\mathrm{T}} P_k. \quad (6.16)$$

Changing the sign of P_k in the extremal equations yields the result. $\qquad\square$

Corollary 6.6. *The Hamiltonian for the flow (6.11) is given by*

$$H(Q_k, P_k) = \text{trace}(e^{\sinh^{-1} M_k/2}) - \text{trace}(\Lambda (J_D)^{-1}(M_k)). \quad (6.17)$$

Proof. From (6.13), we have

$$\begin{aligned}
H(Q_k, P_{k+1}, U_k) &= \text{trace}\left(P_{k+1}^{\mathrm{T}} Q_k U_k\right) - \text{trace}(\Lambda U_k) \\
&= \text{trace}\left(U_k^{\mathrm{T}} P_k^{\mathrm{T}} Q_k U_k\right) - \text{trace}(\Lambda U_k) \\
&= \text{trace}\left(P_k^{\mathrm{T}} Q_k\right) - \text{trace}(\Lambda U_k),
\end{aligned}$$

where

$$U_k = (J_D)^{-1}(P_k^{\mathrm{T}} Q_k - Q_k^{\mathrm{T}} P_k). \quad (6.18)$$

The result follows by noting that $U_k = (J_D)^{-1}(M_k)$ and $P_k = Q_k e^{\sinh^{-1} M_k/2}$. $\qquad\square$
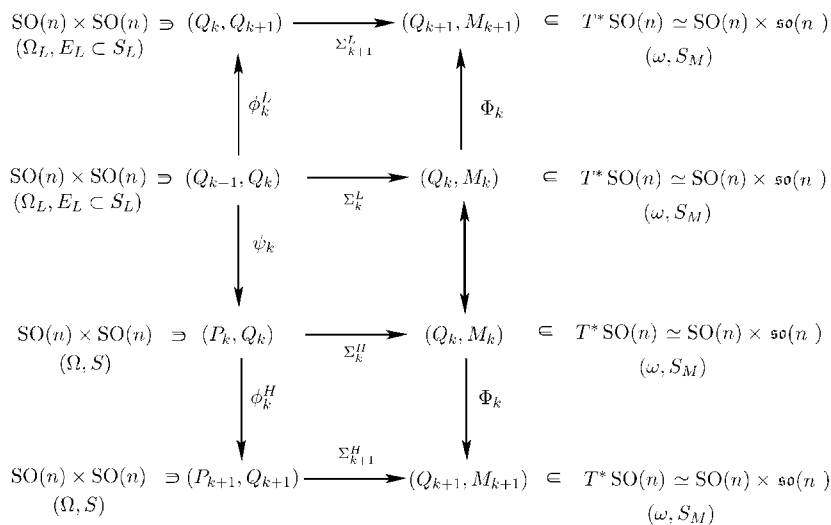
7. Relationships between the rigid body systems

In this section we discuss relationships between the various discrete rigid body systems as well as their symplectic nature. For ease of analysis, we shall make use of the Moser–Veselov system in different notation, which we refer to as the *modified Moser–Veselov system* (MDRBn); it is simply obtained through an index shift. This will be done throughout this section only. Precisely, we change the index on Q_k to Q_{k+1}. The index on M_k is unchanged. When we pass to the symmetric representation of the discrete rigid body, there is a corresponding shift in the index of P_k to P_{k+1}.

In this modified notation, the (MDRBn) equations read as follows:

$$\begin{aligned} J_D\Omega_k^{\mathrm{T}} &= M_k, \\ Q_{k+1} &= Q_k\Omega_k^{\mathrm{T}}, \\ M_{k+1} &= \Omega_k M_k \Omega_k^{\mathrm{T}}. \end{aligned} \qquad \text{(MDRBn)}$$

We have the following commutative diagram showing the relation between various discrete rigid body models:

$$
\begin{array}{ccccc}
\mathrm{SO}(n)\times\mathrm{SO}(n)\supset (Q_k,Q_{k+1}) & \xrightarrow{\ \Sigma^L_{k+1}\ } & (Q_{k+1},M_{k+1}) & \subseteq & T^*\mathrm{SO}(n)\simeq\mathrm{SO}(n)\times\mathfrak{so}(n) \\
(\Omega_L,E_L\subset S_L) & & & & (\omega,S_M) \\
\Big\uparrow{\phi^L_k} & & \Big\uparrow{\Phi_k} & & \\
\mathrm{SO}(n)\times\mathrm{SO}(n)\supset (Q_{k-1},Q_k) & \xrightarrow{\ \Sigma^L_k\ } & (Q_k,M_k) & \subseteq & T^*\mathrm{SO}(n)\simeq\mathrm{SO}(n)\times\mathfrak{so}(n) \\
(\Omega_L,E_L\subset S_L) & & & & (\omega,S_M) \\
\Big\downarrow{\psi_k} & & \Big\updownarrow & & \\
\mathrm{SO}(n)\times\mathrm{SO}(n)\supset (P_k,Q_k) & \xrightarrow{\ \Sigma^H_k\ } & (Q_k,M_k) & \subseteq & T^*\mathrm{SO}(n)\simeq\mathrm{SO}(n)\times\mathfrak{so}(n) \\
(\Omega,S) & & & & (\omega,S_M) \\
\Big\downarrow{\phi^H_k} & & \Big\downarrow{\Phi_k} & & \\
\mathrm{SO}(n)\times\mathrm{SO}(n)\supset (P_{k+1},Q_{k+1}) & \xrightarrow{\ \Sigma^H_{k+1}\ } & (Q_{k+1},M_{k+1}) & \subseteq & T^*\mathrm{SO}(n)\simeq\mathrm{SO}(n)\times\mathfrak{so}(n) \\
(\Omega,S) & & & & (\omega,S_M)
\end{array}
$$

We now detail each of the maps occurring in the diagram. First of all, the map

$$\Phi_k : (Q_k,M_k)\mapsto (Q_{k+1},M_{k+1})$$

is defined by the (MDRBn) algorithm. Second, the map Σ^H_k is defined by

$$\Sigma^H_k : (Q_k,P_k)\mapsto (Q_k, Q_k^{\mathrm{T}}P_k - P_k^{\mathrm{T}}Q_k = M_k).$$

Next, the map $\phi^H_k : (Q_k,P_k)\mapsto (Q_{k+1},P_{k+1})$ is defined by the (SDRBn) equations:

$$\begin{aligned} U_k &= (J_D)^{-1}\left(Q_k^{\mathrm{T}}P_k - P_k^{\mathrm{T}}Q_k\right), \\ Q_{k+1} &= Q_kU_k, \\ P_{k+1} &= P_kU_k. \end{aligned} \qquad (7.1)$$

The map $\phi^L_k : (Q_{k-1},Q_k)\mapsto (Q_k,Q_{k+1})$, the solution of the discrete Euler–Lagrange equations, is defined by the equations (MDRBn) as follows:

$$\Omega_{k-1} = Q_k^{\mathrm{T}}Q_{k-1},$$

$$M_{k-1} = J_D(\Omega_{k-1}^\mathrm{T}),$$
$$M_k = \Omega_{k-1} M_{k-1} \Omega_{k-1}^\mathrm{T}, \tag{7.2}$$
$$\Omega_k^\mathrm{T} = (J_D)^{-1}(M_k),$$
$$Q_{k+1} = Q_k \Omega_k^\mathrm{T}.$$

In other words, we have

$$Q_{k+1} = Q_k (J_D)^{-1}\left(Q_k^\mathrm{T} Q_{k-1} J_D(Q_{k-1}^\mathrm{T} Q_k) Q_{k-1}^\mathrm{T} Q_k\right). \tag{7.3}$$

Next, we turn to the map $\psi_k : (Q_{k-1}, Q_k) \mapsto (P_k, Q_k)$. We know from (SDRBn) that $P_k = Q_k \mathrm{e}^{\sinh^{-1} M_k/2}$ and from (7.2) that

$$M_k = Q_k^\mathrm{T} Q_{k-1} J_D(Q_{k-1}^\mathrm{T} Q_k) Q_{k-1}^\mathrm{T} Q_k. \tag{7.4}$$

Thus, P_k is determined in the mapping ψ_k by

$$P_k = Q_k \exp\left(\sinh^{-1} \tfrac{1}{2}\left(Q_k^\mathrm{T} Q_{k-1} J_D\left(Q_{k-1}^\mathrm{T} Q_k\right) Q_{k-1}^\mathrm{T} Q_k\right)\right). \tag{7.5}$$

The map $\Sigma_k^L : (Q_{k-1}, Q_k) \mapsto (Q_k, M_k)$ is defined by noting that M_k is determined in the mapping Σ_k^L by (7.4).

It is convenient to introduce a particular differential operator D acting on real-valued functions defined on a Lie group G. As usual, we will denote the Lie algebra by $\mathfrak{g}$, and assume that there is an Ad invariant inner product on $\mathfrak{g}$, denoted by $\langle \cdot, \cdot \rangle$.

If ϕ is a smooth function on G, we set

$$D\phi_g(X) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \phi(g\mathrm{e}^{tX}) \triangleq \langle (D\phi_g)^\#, X \rangle,$$

where $g \in G$ and $X \in \mathfrak{g}$. Thus, $D\phi : G \to \mathfrak{g}^*$ and we write its value at g as $D\phi_g \in \mathfrak{g}^*$ and $(D\phi_g)^\# \in \mathfrak{g}$ is the corresponding element of g determined by the inner product on $\mathfrak{g}$. We call $(D\phi_g)$ the 'right' derivative of ϕ at g.

We may also define a 'left' derivative, but for the purposes here where we deal with the left invariant form of the rigid body equations, this is not necessary.

Note that $(D(D\phi))_g \in \mathfrak{g}^* \times \mathfrak{g}^*$ with

$$(DD\phi)_g(X, Y) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \phi(g\mathrm{e}^{sY}\mathrm{e}^{tX}).$$

Clearly we have

$$DD\phi(X, Y) - DD\phi(Y, X) = D\phi([X, Y]). \tag{7.6}$$

In our situation we are interested in a product group $\bar{G} = G \times G$, with Lie algebra $\bar{\mathfrak{g}} = \mathfrak{g} \times \mathfrak{g}$. If $g = (g_1, g_2) \in \bar{G}$ and $X = X_1 + X_2 \in \bar{\mathfrak{g}}$, then

$$D_1\phi_g(X_1) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \phi(g(\mathrm{e}^{tX_1}, e)), \qquad D_2\phi_g(X_2) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \phi(g(e, \mathrm{e}^{tX_2})),$$

where e is used to also represent the identity element in both G_1 and G_2. Thus

$$D_1 D_2 \phi_g(X_1, X_2) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \phi(g(\mathrm{e}^{sX_1}, \mathrm{e}^{tX_2})).$$

Clearly $D_1 D_2 \phi = D_2 D_1 \phi$.

Given a function L on $\bar{G} = G \times G$, we define a two-form Ω_L on $\bar{G}$ as follows. Let

$$(g_1\xi_1, g_2\eta_1), (g_1\xi_2, g_2\eta_2) \in T_g\bar{G} \cong T_{g_1}G \times T_{g_2}G$$

and set

$$\Omega_L((g_1\xi_1, g_2\eta_1), (g_1\xi_2, g_2\eta_2)) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0}\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(L(g_1\mathrm{e}^{t\xi_1}, g_2\mathrm{e}^{s\eta_2}) - L(g_1\mathrm{e}^{t\xi_2}, g_2\mathrm{e}^{s\eta_1}))$$

$$= \tfrac{1}{2}D_2D_1L_g(\xi_1, \eta_2) - \tfrac{1}{2}D_2D_1L_g(\xi_2, \eta_1).$$

If $G = SO(n) \subset GL(n)$, then we may consider the special case of the function L defined by

$$L(g_1, g_2) = \mathrm{trace}\left(g_1\Lambda g_2^{\mathrm{T}}\right), \tag{7.7}$$

where Λ is a positive definite matrix. Now

$$D_2D_1L_g(\xi_1, \eta_2) = -\mathrm{trace}\left(g_1\xi_1\Lambda\eta_2 g_2^{\mathrm{T}}\right) = -\mathrm{trace}\left(g_2^{\mathrm{T}}g_1\xi_1\Lambda\eta_2\right),$$

$$D_2D_1L_g(\xi_2, \eta_1) = -\mathrm{trace}\left(g_2^{\mathrm{T}}g_1\xi_2\Lambda\eta_1\right).$$

Thus,

$$\Omega_L((g_1\xi_1, g_2\eta_1), (g_1\xi_2, g_2\eta_2)) = \tfrac{1}{2}\mathrm{trace}\left(g_2^{\mathrm{T}}g_1(\xi_2\Lambda\eta_1 - \xi_1\Lambda\eta_2)\right). \tag{7.8}$$

In particular, when $\Lambda = $ identity we have $\Omega_L = \Omega_{\mathfrak{gl}(n)}$ as defined earlier. Locally the two form associated with this Lagrangian $L(q_k, q_{k+1}) \stackrel{\sim}{=} L(Q_k, Q_{k+1})$ is given by $\mathrm{d}q_k \wedge (\partial^2 L/\partial q_k\partial q_{k+1})\mathrm{d}q_{k+1}$. We have already seen that the two-form $\Omega_{\mathfrak{gl}(n)}$ is symplectic on $S \subset SO(n) \times SO(n)$. We consider the issue of determining where Ω_L is symplectic later.

The Map Σ_k^L as the Moser–Veselov Legendre transform. The realization of Σ_k^L as the Moser–Veselov Legendre transform is given in the following theorem.

In the following we identify $\mathfrak{so}(n)^*$ with $\mathfrak{so}(n)$ via the trace form: $M \mapsto M^*, M^*(X) \equiv \mathrm{trace}(MX)$, where $M, X \in \mathfrak{so}(n)$ and $M^* \in \mathfrak{so}(n)^*$.

Let $S(g_1, g_2) = 2\,\mathrm{trace}(g_1\Lambda g_2^{\mathrm{T}})$ and set $S_k = S(Q_{k-1}, Q_k)$. We write

$$D_1S_k = (D_1S)_{(Q_{k-1}, Q_k)}, \qquad D_2S_k = (D_2S)_{(Q_{k-1}, Q_k)}.$$

Theorem 7.1. *If*

$$S_k = 2\mathrm{trace}\left(Q_{k-1}\Lambda Q_k^{\mathrm{T}}\right), \tag{7.9}$$

then

$$D_1S_k = -M_{k-1}, \qquad D_2S_k = M_k. \tag{7.10}$$

Proof.

$$D_1S_k(X) = 2\,\mathrm{trace}\left(Q_{k-1}X\Lambda Q_k^{\mathrm{T}}\right) = 2\,\mathrm{trace}\left(Q_k^{\mathrm{T}}Q_{k-1}X\Lambda\right) = 2\,\mathrm{trace}\left(\Omega_{k-1}X\Lambda\right)$$

$$= 2\,\mathrm{trace}(\Lambda\Omega_{k-1}X) = -\mathrm{trace}\left(\left(\Omega_{k-1}^{\mathrm{T}}\Lambda - \Lambda\Omega_{k-1}\right)X\right) = -\mathrm{trace}(M_{k-1}X).$$

Thus, $D_1S_k = -M_{k-1}$ as required.

$$D_2S_k(X) = -2\,\mathrm{trace}\left(Q_{k-1}\Lambda X Q_k^{\mathrm{T}}\right) = -2\,\mathrm{trace}\left(Q_k^{\mathrm{T}}Q_{k-1}\Lambda X\right) = -2\,\mathrm{trace}(\Omega_{k-1}\Lambda X)$$

$$= -\mathrm{trace}\left(\left(\Omega_{k-1}\Lambda - \Lambda\Omega_{k-1}^{\mathrm{T}}\right)X\right).$$

But from (MDRBn) we have $M_k = \Lambda\Omega_{k-1}^{\mathrm{T}} - \Omega_{k-1}\Lambda$, so $D_2S_k(X) = \mathrm{trace}\,M_k X$, and so $D_2S_k = M_k$ as required. $\qquad\square$

Note. Equations (7.10) are obtained using the modified Moser–Veselov equations (MDRBn) equations. If (DRBn) were used we would have

$$D_1S_k = -M_k, \qquad D_2S_k = M_{k+1}.$$

We now return to using only (MDRBn). From theorem 7.1 we may express Σ_k^L in the form

$$\Sigma_k^L : (Q_{k-1}, Q_k) \mapsto (Q_k, D_2S_k), \tag{7.11}$$

where S_k is given in (7.9).

Pull backs of two forms. We observe that all the maps in our commutative diagram are symplectic (see appendix B). In particular one can demonstrate the following pull-back relationships:

$$\left(\Sigma_k^L\right)^* \omega = 2\Omega_L, \qquad (\psi_k)^* \Omega_{\mathfrak{gl}(n)} = \Omega_L,$$
$$\left(\phi_k^L\right)^* \Omega_L = \Omega_L, \qquad (\Phi_k)^* \omega = \omega. \qquad (7.12)$$

Note. The proof that $(\Sigma_k^H)^* \omega = 2\Omega_{\mathfrak{gl}(n)}$ was accomplished in proposition 3.11 in the continuous case, since for this case the mapping in the continuous and discrete cases are identical. The proof that $(\phi_k^H)^* \Omega_{\mathfrak{gl}(n)} = \Omega_{\mathfrak{gl}(n)}$ is a remark after corollary 6.3.

Domains for discrete algorithms. We now consider the domains of validity of the discrete algorithms we have discussed. We introduce the following sets:

$$U_\alpha = \{A \mid A \in \mathfrak{so}(n), \|A\|_{\mathrm{op}} < \alpha\},$$
$$D_L = \{\Omega \mid \Omega \in SO(n), J_D(\Omega^{\mathrm{T}}) \in U_2\},$$
$$S_L = \{(Q, \bar{Q}) \mid \bar{\Omega} := \bar{Q}^{\mathrm{T}} Q \in D_L\}, \qquad Q, \bar{Q} \in SO(n),$$
$$E_L = \{(Q, \bar{Q}) \mid (Q, \bar{Q}) \in S_L, W \mapsto \Lambda W \bar{\Omega}^{\mathrm{T}} + \bar{\Omega} W \Lambda$$
$$\text{is invertible as a mapping of } \mathfrak{so}(n) \text{ to } \mathfrak{so}(n)\}.$$

Domain on which Ω_L is symplectic. We know that $\Omega_{\mathfrak{gl}(n)}$ is symplectic on $S \subset SO(n) \times SO(n)$, and

$$\psi_k^* \Omega_{\mathfrak{gl}(n)} = \tfrac{1}{2} \left(\Sigma_k^L\right)^* \omega = \Omega_L.$$

It follows that Ω_L cannot be symplectic outside the set

$$\psi_k^{-1}(S) = \{(Q_{k-1}, Q_k) \mid \psi_k(Q_{k-1}, Q_k) \in S\}.$$

Note $\psi_k^{-1}(S) \subset SO(n) \times SO(n)$ is independent of k. We characterize this set. Since $(Q_k, P_k) \in S$ if and only if $M_k \in U_2$ and $M_k = \Omega_{k-1} M_{k-1} \Omega_{k-1}^{\mathrm{T}}$ so

$$(Q_k, P_k) \in S \qquad \text{if and only if } M_{k-1} = J_D\left(\Omega_{k-1}^{\mathrm{T}}\right) \in U_2.$$

But $\Omega_{k-1}^{\mathrm{T}} = Q_{k-1}^{\mathrm{T}} Q_k$, so

$$(Q_{k-1}, Q_k) \in \psi_k^{-1}(S) \qquad \text{if and only if } (Q_{k-1}, Q_k) \in S_L.$$

Within the set S_L, we need to know where Ω_L is nondegenerate, or in other words, where the mapping $\psi_k : S_L \to S; (Q_{k-1}, Q_k) \mapsto (Q_k, P_k)$ has full rank. Clearly we need only determine where the mapping $Q_{k-1} \mapsto P_k$ is of full rank. From the relation $P_k = Q_k e^{\sinh^{-1} M_k/2}$, we know that

$$P_k \delta P_k = \frac{1}{2} P_k \left(1 + \left(\frac{M_k}{2}\right)^2\right)^{-1/2} \delta M_k.$$

Also, from the relation $M_k = \Lambda Q_{k-1}^{\mathrm{T}} Q_k - Q_k^{\mathrm{T}} Q_{k-1} \Lambda$, we obtain

$$\delta M_k = -\Lambda \delta Q_{k-1} Q_{k-1}^{\mathrm{T}} Q_k - Q_k^{\mathrm{T}} Q_{k-1} \delta Q_{k-1} \Lambda.$$

If $W = \delta Q_{k-1} \in so(n)$, we have

$$-2\left(1 + \left(\frac{M_k}{2}\right)^2\right)^{1/2} \delta P_k = \Lambda W \Omega_{k-1}^{\mathrm{T}} + \Omega_{k-1} W \Lambda.$$

Thus ψ_k has full rank when $W \mapsto \Lambda W \Omega_{k-1}^{\mathrm{T}} + \Omega_{k-1} W \Lambda$ is invertible as a map on $so(n)$. Note that we require $\|M_k\|_{\mathrm{op}} < 2$ for invertibility of the term $(1 + (M_k/2)^2)$, but this follows from the condition $(Q_{k-1}, Q_k) \in S_L$. We have proven the following theorem.

Theorem 7.2. *Ω_L is symplectic on $E_L \subset S_L \subset SO(n) \times SO(n)$.*

This result can also be seen using the relation $\frac{1}{2}(\Sigma_k^L)^*\omega = \Omega_L$. Since $\Sigma_k^L : (Q_{k-1}, Q_k) \mapsto (Q_k, M_k)$ and $M_k = \Lambda Q_{k-1}^T Q_k - Q_k^T Q_{k-1}\Lambda$, we need to check that the mapping $Q_{k-1} \mapsto M_k$ is of full rank, as in the analysis above. However, since $\Sigma_k^L = \Sigma_k^H \circ \psi_k$, Σ_k^H must also have full rank. Thus, the mapping $(Q_k, P_k) \mapsto (Q_k, M_k = Q_k^T P_k - P_k^T Q_k)$ must be of full rank. It is sufficient that the mapping $P_k \mapsto Q_k^T P_k - P_k^T Q_k$ is of full rank. Thus it is necessary that the mapping

$$\delta P_k \mapsto Q_k^T P_k \delta P_k + \delta P_k P_k^T Q_k$$

is of full rank, with $\delta P_k \in SO(n)$. But this is of course the condition that $\Omega_{\mathfrak{gl}(n)}$ is nondegenerate, $(Q_k, P_k) \in S$, which is again equivalent to the condition that $(Q_{k-1}, Q_k) \in S_L$. We again obtain the result of theorem 7.2.

The solution of the equation $J_D(U) = M$.   Here we give explicit conditions on M and Λ under which the system

$$J_D(U) = U\Lambda - \Lambda U^T = M, \qquad M \in SO(n), \quad U \in SO(n) \tag{7.13}$$

has explicit solutions for U.

In the case $\Lambda = I$, the solution is given by $M/2 = \sinh A$, $U = e^A$ for $\|M\|_{op} < 2$ as was demonstrated in section 3.

In the general case, we extract some results from Cardoso and Leite (2001).

Lemma 7.3. *The expression*

$$U = \left(\frac{M}{2} + S\right)\Lambda^{-1}$$

is an orthogonal solution of (7.13) if and only if S is a symmetric solution of the Riccati equation:

$$S^2 + S\left(\frac{M}{2}\right) + \left(\frac{M}{2}\right)^T S - \left(\frac{M^2}{4} + \Lambda^2\right) = 0. \tag{7.14}$$

It is well known that the solution of the Riccati equation is

$$SDS + SA + A^T S - C = 0,$$

where $D \geqslant 0$, $C = C^T$ are governed by the spectral analysis of the Hamiltonian matrix

$$H = \begin{bmatrix} A & D \\ C & -A^T \end{bmatrix}.$$

See Lancaster and Rodman (1980, 1995) and Kucera (1972) for details on the algebraic Riccati equation.

In particular, by lemma 7.3, solutions of equation (7.13) are determined through the spectral analysis of the Hamiltonian matrix

$$\bar{H} = \begin{bmatrix} \dfrac{M}{2} & I \\ \dfrac{M^2}{4} + \Lambda^2 & \dfrac{M}{2} \end{bmatrix}.$$

Theorem 7.4 (Cardoso and Leite (2001)). *Equation (7.13) has a unique solution $U \in SO(n)$ if and only if the spectrum of $\bar{H}$ is pure imaginary, and the size of the Jordan blocks associated to each nonzero eigenvalue is even.*

The analysis of Moser and Veselov (1991) is also recovered, corresponding to the case where $\bar{H}$ has no pure imaginary eigenvalues. Indeed, in Cardoso and Leite (2001) it is shown that the condition that the size of the Jordan blocks associated to each nonzero imaginary eigenvalue of $\bar{H}$ is even is equivalent to existence of special orthogonal solutions of equation (7.13). Thus if $\bar{H}$ has no nonzero imaginary eigenvalues, there are multiple solutions of (7.13).

Theorem 7.5 (Cardoso and Leite 2001). *If $M^2/4 + \Lambda^2$ is positive definite, then equation (7.14) has a unique solution $S \geqslant 0$ such that the eigenvalues of $S + M/2$ have positive real parts and $U = (S + M/2)\,\Lambda^{-1}$ is a special orthogonal solution of equation (7.13).*

As an application of this result, we determine a condition under which all three algorithms, $\Phi_k, \phi_k^L, \phi_k^H$, are well defined, i.e. the operator J_D may be inverted at each step of the algorithm.

We first note that since (MDRBn) (and (DRBn) imply $M_{k+1} = \Omega_k M_k \Omega_k^T$ and $M_{k+1}^2 = \Omega_k M_k^2 \Omega_k^T$. Thus, if $M_k^2/4 + \alpha I > 0$ for some real number α, then $M_{k+1}^2/4 + \alpha I > 0$ also.

Lemma 7.6. *Assume $M_0^2/4 + \alpha I > 0$ for some $\alpha > 0$, where $\Lambda^2 > \alpha I$, then for each algorithm $\Phi_k, \phi_k^L, \phi_k^H$, at every step J_D is invertible and $M_k^2/4 + \Lambda^2$ is positive definite for each k.*

Proof. If $M_0^2/4 + \alpha I > 0$, it is clear from the previous analysis that $M_k^2/4 + \alpha I > 0$ for each k. But then

$$\frac{M_k^2}{4} + \Lambda^2 = \left(\frac{M_k^2}{4} + \alpha I \right) + (\Lambda^2 - \alpha I) > 0.$$

Thus, by theorem 7.5, equation (7.13) with $M = M_k$ may be solved for each k, i.e. J_D is invertible. □

Further relations between algorithms. We recall the relationship in equation (B.8):

$$\left(P_k^T + \Lambda Q_{k-1}^T \right) Q_k - Q_h^T (P_k + Q_{k-1}\Lambda) = 0.$$

This states that the matrix

$$\Sigma_k = Q_k^T (P_k + Q_{k-1}\Lambda) \tag{7.15}$$

is symmetric: $\Sigma_k = \Sigma_k^T$. We may rewrite equation (7.15) in the form

$$P_k = Q_k \Sigma_k - Q_{k-1}\Lambda. \tag{7.16}$$

Thus P_k is an interpolation of Q_{k-1} and Q_k, with symmetric weights Σ_k and Λ. But from (7.15) we also have

$$Q_k^T P_k = \Sigma_k - Q_k^T Q_{k-1}\Lambda = \Sigma_k - \Omega_{k-1}\Lambda.$$

Since $Q_k^T P_k \in SO(n)$, we have the identity $(\Sigma_k - \Omega_{k-1}\Lambda)^T (\Sigma_k - \Omega_{k-1}\Lambda) = I$, i.e.

$$\Sigma_k^2 - \Lambda\Omega_{k-1}^T \Sigma_k - \Sigma_k \Omega_{k-1}\Lambda + \Lambda^2 - I = 0. \tag{7.17}$$

Thus given the ϕ_k^L algorithm we may realize the map $\psi_k : (Q_{k-1}, Q_k) \mapsto (Q_k, P_k)$ by solving the quadratic equation (7.17) and using the interpolation formula (7.16). This should be compared with formula (7.5) which employs the operator $\sinh^{-1}$.

We have, however, proved the following result in corollary 6.6. The Hamiltonian for the discrete Hamiltonian flow ϕ_k^H is given by

$$H(Q_k, P_k) = \text{trace}(e^{\sinh^{-1} M_k/2}) - \text{trace}(\Lambda U_k).$$

Now we know that if $S_k = e^{\sinh^{-1} M_k/2}$ we have $S_k - S_k^{\mathrm{T}} = M_k$. Thus,

$$S_k - S_k^{\mathrm{T}} = M_k = U_k \Lambda - \Lambda U_k^{\mathrm{T}},$$

so $S_k - U_k \Lambda = S_k^{\mathrm{T}} - \Lambda U_k^{\mathrm{T}}$ and hence, $\Lambda_k \stackrel{\triangle}{=} S_k - U_k \Lambda = \Lambda_k^{\mathrm{T}}$. It follows that

$$H(Q_k, P_k) = \mathrm{trace}\,\Lambda_k. \tag{7.18}$$

Since S_k is orthogonal, we have

$$(\Lambda_k + U_k \Lambda)^{\mathrm{T}}(\Lambda_k + U_k \Lambda) = I,$$

i.e.

$$\Lambda_k^2 + \Lambda U_k \Lambda_k + \Lambda_k U_k \Lambda + \Lambda^2 - I = 0. \tag{7.19}$$

Moreover, we know that $S_k = Q_k^{\mathrm{T}} P_k$ and $U_k = Q_k^{\mathrm{T}} Q_{k+1}$, so

$$\Lambda_k = Q_k^{\mathrm{T}}(P_k - Q_{k+1}\Lambda) \tag{7.20}$$

and

$$P_k = Q_k \Lambda_k + Q_{k+1}\Lambda. \tag{7.21}$$

Thus again we may recover a relationship between the sequence (Q_{k-1}, Q_k) and the sequence (Q_k, P_k) on $SO(n) \times SO(n)$, via an interpolation (7.21) by symmetric matrices, where Λ_k is determined from the quadratic equation (7.19). Equation (7.19) should be compared with the similar quadratic equation (7.17). The importance of this second formulation, equations (7.19)–(7.21), lies in their relation to the Hamiltonian of the flow given by the sum of the eigenvalues of Λ_k in (7.18).

8. Conclusions

In this paper we have presented the theory of the symmetric representation of the rigid body equations on $SO(n) \times SO(n)$ in both the continuous and discrete setting and established a concrete link with the discrete Moser–Veselov equations. In addition, we described the relationship between the various symplectic structures on the underlying phase spaces.

We now make a few further remarks on integrability and future research plans in this area as well as possible future research on optimal control and numerical issues.

We recall from equation (1.2) in the introduction that the reduced equations of motion for (SRBn) are

$$\dot{Q} = Q\Omega(Q). \tag{8.1}$$

From (3.1) and the expression $P = AQ$ where $A = P(0)Q(0)^{\mathrm{T}}$, we can rewrite this equation as

$$\dot{Q} = QJ^{-1}(Q^{\mathrm{T}}(A - A^{\mathrm{T}})Q). \tag{8.2}$$

This is the natural $SO(n)$ reduction of the symmetric representation of the rigid body equations. Further, the Manakov integrals naturally pull back to this flow. Explicitly they take the form

$$\mathrm{trace}((A - A^{\mathrm{T}}) + \lambda Q\Lambda^2 Q^{\mathrm{T}})^k. \tag{8.3}$$

Note that it is easy to see that equation (8.2) implies the Lax equation for the rigid body by using the identity $\Omega = J^{-1}(Q^{\mathrm{T}}(A - A^{\mathrm{T}})Q)$.

In a future publication we intend to analyse the complete integrability of this reduced (SRBn) system as well as any natural Hamiltonian structure on the group $SO(n)$. This

Hamiltonian structure may be related to the almost Poisson structures for the unreduced system (SRRBn) discussed in Bloch *et al* (2000).

We make few further remarks on this issue here. One may reformulate the Manakov Lax pair formulation in the symmetric representation setting as follows. For an alternative approach to the Lax pair formulation see Federov (2002).

We consider the equations with parameter:

$$\dot{Q}_\lambda = Q_\lambda(\Omega + \lambda\Lambda), \qquad \dot{P}_\lambda = P_\lambda(\Omega + \lambda\Lambda), \tag{8.4}$$

where

$$\Omega = J^{-1}(M_\lambda - \lambda\Lambda^2) \qquad \text{and} \qquad M_\lambda = Q_\lambda^{-1}P_\lambda - P_\lambda^{-1}Q_\lambda. \tag{8.5}$$

Note, however, that at this point we view Q_λ and P_λ simply as elements of $GL(n)$. There is no implied functional dependence on Q and P appearing in the symmetric representation of the rigid body equations.

Then we can easily show that M_λ satisfies the rigid body equations with parameter

$$\dot{M}_\lambda = [M_\lambda, \Omega + \lambda\Lambda], \tag{8.6}$$

where $\Omega = J^{-1}(M_\lambda - \lambda\Lambda^2)$.

Note that the spatial momentum with parameter $m_\lambda = P_\lambda Q_\lambda^{-1} - Q_\lambda P_\lambda^{-1}$ is conserved under the flow. Note also that in the definition of M_λ here we use inverses rather than transpose. This is crucial when we generalize from the pure finite-dimensional group setting.

By comparing (2.4) and (8.6) we see that $M_\lambda = M + \lambda\Lambda^2$. However, there is no similar relationship between Q_λ, P_λ and Q, P.

We note also that the double bracket formulation discussed in section 5 turns out to be useful for analysing the integrability of the full rigid body equations. We can obtain not only the Manakov integrals but the spatial conserved momenta which are needed for noncommutative integrability (we intend to return to this issue in a future paper). Consider the equation

$$[\hat{P}, \hat{Q}] = \begin{bmatrix} QP^{\mathrm{T}} - PQ^{\mathrm{T}} & 0 \\ 0 & Q^{\mathrm{T}}P - P^{\mathrm{T}}Q \end{bmatrix}. \tag{8.7}$$

In the left invariant formulation, the lower right block is the body momentum while the upper left block is (minus) the spatial momentum.

To obtain the Manakov integrals, set $\hat{J} = \mathrm{diag}(0, J)$ as before. The Manakov integrals (and Casimirs) are then given by

$$\mathrm{trace}([\hat{P}, \hat{Q}] + \lambda\hat{J})^k, \tag{8.8}$$

while the spatial momenta are given by $\mathrm{trace}(\pi_U([\hat{P}, \hat{Q}])A)$ where A runs through a basis of spatial momenta and π_U is projection onto the upper-left block. Because of this upper/lower structure, one sees that the spatial and body integrals are in involution.

We can make some similar observation in the discrete setting. Moser and Veselov showed that (4.2) is equivalent to the equation of isospectral deformations

$$M_{k+1}^\lambda = \Omega_k^\lambda M_k^\lambda \left(\Omega_k^\lambda\right)^{-1}, \tag{8.9}$$

where $M_k^\lambda = M_k + \lambda\Lambda^2$ and $\Omega_k^\lambda = \Omega_k - \lambda\Lambda$.

Similarly we can define the symmetric representation of the discrete rigid body equations with parameters

$$Q_{k+1}^{\lambda} = Q_k^{\lambda}(U_k + \lambda \Lambda), \qquad P_{k+1}^{\lambda} = P_k^{\lambda}(U_k + \lambda \Lambda), \qquad (8.10)$$

where $U_k = (J_D)^{-1}(M_k^{\lambda} - \lambda \Lambda^2)$, $M_k^{\lambda} := (Q_k^{\lambda})^{-1}P_k^{\lambda} - (P_k^{\lambda})^{-1}Q_k^{\lambda}$ and $J_D(\Omega) = \Omega^{\mathrm{T}}\Lambda - \Lambda\Omega$. Then we have

$$M_{k+1}^{\lambda} = (U_k + \lambda \Lambda)^{-1} M_k^{\lambda}(U_k + \lambda \Lambda). \qquad (8.11)$$

Again, there is no functional dependence between Q_k^{λ} and P_k^{λ} implied so this is quite formal in some sense. In a forthcoming paper we hope to discuss integrability in the discrete case and its connection with the work of Moser and Veselov (1991) and Deift *et al* (1992).

In addition, we intend to consider the details of the symmetric representation of the rigid body equations as a numerical algorithm. In a future paper we also intend to consider generalizations of the theory here to arbitrary continuous and discrete Lagrangians on manifolds embedded in a vector space V. We hope to use the theory of optimal control to formulate such systems on $V \times V$ or $V \times V^*$.

Acknowledgments

Appendix A. Proof of proposition 6.2

Proof. We use the discrete maximum principle. We wish to minimize $\sum_{k=0}^{N} g(x_k, u_k)$ subject to the given discrete dynamics and control constraint set E. To implement the constraints defining E, we consider maximizing the augmented functional

$$V(p_{k+1}, x_k, u_k) = \sum_{k=0}^{N} \langle p_{k+1}, f(x_k, u_k) - x_{k+1} \rangle + \langle \sigma, k(u_k) \rangle - g(x_k, u_k)$$

$$= \sum_{k=0}^{N} -\langle p_{k+1}, x_{k+1} \rangle + \hat{H}(p_{k+1}; x_k, u_k, \sigma)$$

$$= \langle p_0, x_0 \rangle + \left(\sum_{k=0}^{N} -\langle p_k, x_k \rangle + \hat{H}(p_{k+1}, x_k, u_k, \sigma) \right),$$

where we set $p_{N+1} = 0$ for notational convenience since the $(N+1)$th equation for x_k is not needed. Thus,

$$\delta V = \sum_{k=0}^{N} -\langle p_k, \delta x_k \rangle - \langle \delta p_k, x_k \rangle + \frac{\partial \hat{H}}{\partial p_{k+1}} \delta p_{k+1} + \frac{\partial \hat{H}}{\partial x_k} \delta x_k + \frac{\partial \hat{H}}{\partial u_k} \delta u_k, \qquad (A.1)$$

where $k(u_k) = 0$. This gives extremal equations (6.2) and (6.5) since

$$\frac{\partial \hat{H}}{\partial p_{k+1}}(p_{k+1}, x_k, u_k, \sigma) = \frac{\partial H}{\partial p_{k+1}}(p_{k+1}, x_k, u_k),$$

$$\frac{\partial \hat{H}}{\partial x_k}(p_{k+1}, x_k, u_k, \sigma) = \frac{\partial H}{\partial x_k}(p_{k+1}, x_k, u_k). \qquad (A.2)$$

$\square$

Appendix B. Pull-back identities

Here we prove the identities in equation (7.12).

Lemma B.1. $\left(\Sigma_k^L\right)^* \omega = 2\Omega_L.$

Proof. Above, we defined the map

$$\Sigma_k^L : SO(n) \times SO(n) \to T^*SO(n) \simeq SO(n) \times \mathfrak{so}(n)^*.$$

Thus, the tangent is a map on the following spaces:

$$T\Sigma_k^L : TSO(n) \times TSO(n) \to TSO(n) \times T\mathfrak{so}(n)^*.$$

Since $\Sigma_k^L : (Q_{k-1}, Q_k) \mapsto (Q_k, M_k = D_2 S_k)$, we find that

$$T\Sigma_k^L : T_{Q_{k-1}}SO(n) \times T_{Q_k}SO(n) \to T_{Q_k}SO(n) \times \mathfrak{so}(n)^*.$$

For $i = 1, 2$, let $V_i = (Q_{k-1}\xi_i, Q_k\eta_i) \in T_{Q_{k-1}}SO(n) \times T_{Q_k}SO(n)$, then we see that

$$T\Sigma_k^L \cdot V_i = (Q_k \eta_i, D_1 D_2 S_k \xi_i + D_2 D_2 S_k \eta_i). \tag{B.1}$$

Note that in this computation it is essential that we represent $T_{Q_{k-1}}SO(n) \times T_{Q_k}SO(n)$ by the 'left' translates of vectors $(\xi_i, \eta_i) \in \mathfrak{so}(n) \times \mathfrak{so}(n)$, to be compatible with the use of the 'right' derivative.

Write $T\Sigma_k^L \cdot V_i = (A_i, Z_i) \in T_{Q_k}SO(n) \times \mathfrak{so}(n)^*$ and identify $\mathfrak{so}(n)^*$ with $\mathfrak{so}(n)$ using the trace form, as above. Using the canonical structure of ω we have

$$\omega((A_1, Z_1), (A_2, Z_2)) = \langle Z_2, Q_k^{-1} A_1 \rangle - \langle Z_1, Q_k^{-1} A_2 \rangle + \langle M_k, [Q_k^{-1} A_1, Q_k^{-1} A_2] \rangle.$$

Substituting the expressions (B.1), we obtain

$$\omega(T\Sigma^L \cdot V_1, T\Sigma^L \cdot V_2) = -\tfrac{1}{2}(D_1 D_2 S_k(\xi_2, \eta_1) + D_2 D_2 S_k(\eta_2, \eta_1))$$
$$+ \tfrac{1}{2}(D_1 D_2 S_k(\xi_1, \eta_2) + D_2 D_2 S_k(\eta_1, \eta_2)) - \tfrac{1}{2}\operatorname{trace}(M_k[\eta_1, \eta_2]).$$

But by identity (7.6) and the fact that $D_2 S_k = M_k$ from (7.10), we obtain

$$\omega(T\Sigma^L \cdot V_1, T\Sigma^L \cdot V_2) = -\tfrac{1}{2}(D_1 D_2 S_k(\xi_2, \eta_1) - D_1 D_2 S_k(\xi_1, \eta_2)).$$

But this is simply $2\Omega_L$. $\qquad\square$

Lemma B.2. $(\phi_k^L)^* \Omega_L = \Omega_L.$

Proof. Recall that $\phi_k^L : (Q_{k-1}, Q_k) \mapsto (Q_k, Q_{k+1})$. We may express the necessary conditions of the Moser–Veselov variational problem as $D_2 L_{(Q_{k-1}, Q_k)} + D_1 L_{(Q_k, Q_{k+1})} = 0$. Note that this identity is an identity in the $so(n)^*$ and is obtained through an operation in the same space $T_{Q_k}SO(n)$. We may rewrite this identity in the general form of a product group $\bar{G} = G \times G$:

$$D_2 S(g_1, g_2) + D_1 S(g_2, g_3) = 0. \tag{B.2}$$

Let $\xi_i, \eta_i \in \mathfrak{g}$. Differentiating (B.2) we obtain

$$D_1 D_2 S(g_1, g_2)\xi_1 + D_2 D_2 S(g_1, g_2)\xi_2 + D_1 D_1 S(g_2, g_3)\xi_2 + D_2 D_1 S(g_2, g_3)\xi_3 = 0. \tag{B.3}$$

We obtain a similar expression with η_i substituting ξ_i, $i = 1, 2, 3$.

We have

$$4\Omega_{L(g_1, g_2)}(((g_1\xi_1), (g_2\xi_2)), ((g_1\eta_1), (g_2\eta_2)))(=A(g_1, g_2))$$
$$= D_2 D_1 S_{(g_1, g_2)}(\xi_1, \eta_2) - D_2 D_1 S_{(g_1, g_2)}(\xi_2, \eta_1)$$

and

$$4\Omega_{L_{(g_2,g_3)}}(((g_2\xi_2),(g_3\xi_3)),((g_2\eta_2),(g_3\eta_3)))(=B(g_2,g_3))$$
$$= D_2 D_1 S_{(g_2,g_3)}(\xi_2,\eta_3) - D_2 D_1 S_{(g_2,g_3)}(\xi_3,\eta_2).$$

But from (B.3) we obtain

$$D_1 D_2 S(g_1,g_2)(\xi_1,\eta_2) + D_2 D_2 S(g_1,g_2)(\xi_2,\eta_2) + D_1 D_1 S(g_2,g_3)(\xi_2,\eta_2)$$
$$+ D_2 D_1 S(g_2,g_3)(\xi_3,\eta_2) = 0$$

and

$$D_1 D_2 S(g_1,g_2)(\eta_1,\xi_2) + D_2 D_2 S(g_1,g_2)(\eta_2,\xi_2) + D_1 D_1 S(g_2,g_3)(\eta_2,\xi_2)$$
$$+ D_2 D_1 S(g_2,g_3)(\eta_3,\xi_2) = 0.$$

Subtracting the last two identities we obtain

$$A(g_1,g_2) = B(g_2,g_3) - D_2 S(g_1,g_2)([\xi_2,\eta_2]) - D_1 S(g_2,g_3)([\xi_2,\eta_2]).$$

But from (B.2) we have $A(g_1,g_2) = B(g_2,g_3)$. If $\phi : \bar{G} \to \bar{G}$ is given by $\phi(g_1,g_2) = (g_2,g_3)$ where (g_1,g_2,g_3) satisfy (B.2), we have

$$\tfrac{1}{4}B(g_2,g_3) = (\phi^*\Omega_L)(((g_1\xi_1),(g_2\xi_2)),((g_1\eta_1),(g_2\eta_2))),$$
$$\tfrac{1}{4}A(g_1,g_2) = \Omega_L(((g_1\xi_1),(g_2\xi_2)),((g_1\eta_1),(g_2,\eta_2))). \qquad \square$$

Lemma B.3. $\Phi_k^*\omega = \omega.$

Proof. Recall that $\Phi_k : (Q_k, M_k) \longrightarrow (Q_{k+1}, M_{k+1})$. We have, from (7.10)

$$M_{k+1} = D_2 S_{k+1} = D_2 S(Q_k, Q_{k+1}),$$
$$-M_k = D_1 S_{k+1} = D_1 S(Q_k, Q_{k+1}). \tag{B.4}$$

Differentiating these expressions, we get, for $i = 1, 2$,

$$\delta_i M_{k+1} = D_1 D_2 S(Q_{k_1} Q_{k+1})\delta_i Q_k + D_2 D_2 S(Q_k, Q_{k+1})\delta_i Q_{k+1},$$
$$-\delta_i M_k = D_1 D_1 S(Q_k, Q_{k+1})\delta_i Q_k + D_2 D_1 S(Q_k, Q_{k+1})\delta_i Q_{k+1}, \tag{B.5}$$

where

$$v\delta_i Q_k, \delta_i Q_{k+1} \in \mathfrak{so}(n) \qquad \text{and} \qquad \delta_i M_k, \delta_i M_{k+1} \in \mathfrak{so}(n)^*.$$

In terms of these quantities, we may write

$$\hat{\omega}_{(Q_k,M_k)} = -2\omega((Q_k\delta_1 Q_k, \delta_1 M_k),(Q_k\delta_2 Q_k, \delta_2 M_k))$$
$$= \delta_2 M_k(\delta_1 Q_k) - \delta_1 M_k(\delta_2 Q_k) - M_k([\delta_1 Q_k, \delta_2 Q_k]) \tag{B.6}$$

and

$$\hat{\omega}_{(Q_{k+1},M_{k+1})} = -2\omega((Q_{k+1}\delta_1 Q_{k+1}, \delta_1 M_{k+1}),(Q_{k+1}\delta_2 Q_{k+1}, \delta_2 M_{k+1}))$$
$$= \delta_2 M_{k+1}(\delta_1 Q_{k+1}) - \delta_1 M_{k+1}(\delta_2 Q_{k+1}) - M_{k+1}([\delta_1 Q_{k+1}, \delta_2 Q_{k+1}]). \tag{B.7}$$

We substitute $M_k = -D_1 S(Q_k, Q_{k+1})$ in (B.6) and substitute $M_{k+1} = D_2 S(Q_k, Q_{k+1})$ in (B.7). Letting $D_i(D_j S(Q_k, Q_{k+1})) = D_i(D_j S_k)$, and substituting (B.5) into expressions (B.6) and (B.7) yields

$$\hat{\omega}_{(Q_k,M_k)} = -D_1 D_1 S_k(\delta_1 Q_k, \delta_2 Q_k) - D_2 D_1 S_k(\delta_1 Q_k, \delta_2 Q_{k+1})$$
$$+ D_1 D_1 S_k(\delta_2 Q_k, \delta_1 Q_k) + D_2 D_1 S_k(\delta_2 Q_k, \delta_1 Q_{k+1}) + D_1 S_k([\delta_1 Q_k, \delta_2 Q_k]),$$
$$\hat{\omega}_{(Q_{k+1},M_{k+1})} = D_1 D_2 S_k(\delta_1 Q_{k+1}, \delta_2 Q_k) + D_2 D_2 S_k(\delta_1 Q_{k+1}, \delta_2 Q_{k+1}) - D_1 D_2 S_k(\delta_2 Q_{k+1}, \delta_1 Q_k)$$
$$- D_2 D_2 S_k(\delta_2 Q_{k+1}, \delta_1 Q_{k+1}) - D_2 S_k([\delta_1 Q_{k+1}, \delta_2 Q_{k+1}]).$$

Again, using identity (7.6),

$$\hat{\omega}(Q_{k+1}, M_{k+1}) = -2(\Phi_k^*\omega)((Q_k\delta_1 Q_k, \delta_1 M_k), (Q_k\delta_2 Q_k, \delta_2 M_k))$$
$$= -2\omega((Q_k\delta_1 Q_k, \delta_1 M_k), (Q_k\delta_2 Q_k, \delta_2 M_k))$$
$$= \hat{\omega}(Q_k, M_k). \qquad \square$$

Lemma B.4. $\psi_k^*\Omega_{\mathfrak{gl}(n)} = \Omega_L$.

Proof. From (SDRBn) and (MDRBn), we have

$$M_k = Q_k^{\mathrm{T}} P_k - P_k^{\mathrm{T}} Q_k = \Omega_{k-1} M_{k-1} \Omega_{k-1}^{\mathrm{T}} = \Omega_{k-1}(\Omega_{k-1}^{\mathrm{T}}\Lambda - \Lambda\Omega_{k-1})\Omega_{k-1}^{\mathrm{T}}$$
$$= \Lambda\Omega_{k-1}^{\mathrm{T}} - \Omega_{k-1}\Lambda = \Lambda Q_{k-1}^{\mathrm{T}} Q_k - Q_k^{\mathrm{T}} Q_{k-1}\Lambda.$$

Hence,

$$\left(P_k^{\mathrm{T}} + \Lambda Q_{k-1}^{\mathrm{T}}\right) Q_k - Q_k^{\mathrm{T}}(P_k + Q_{k-1}\Lambda) = 0. \tag{B.8}$$

Take variations of (B.8) with

$$\delta Q_k = Q_k\eta, \qquad \delta P_k = P_k\xi, \qquad \delta Q_{k-1} = Q_{k-1}\mu$$

where $\eta, \xi, \mu \in \mathfrak{so}(n)$. Thus,

$$-\xi P_k^{\mathrm{T}} Q_k + P_k^{\mathrm{T}} Q_k\eta - \Lambda\mu Q_{k-1}^{\mathrm{T}} Q_k + \Lambda Q_{k-1}^{\mathrm{T}} Q_k\eta + \eta Q_k^{\mathrm{T}} P_k - Q_k^{\mathrm{T}} P_k\xi$$
$$+\eta Q_k^{\mathrm{T}} Q_{k-1}\Lambda - Q_k^{\mathrm{T}} Q_{k-1}\mu\Lambda = 0. \tag{B.9}$$

Write $X_i = (Q_k\eta_i, P_k\xi_i)$, $Z_i = (Q_{k-1}\mu_i, Q_k\eta_i)$. We have

$$\Omega_{\mathfrak{gl}(n)}(Q_k, P_k)(X_1, X_2) = \tfrac{1}{2}\,\mathrm{trace}\left(P_k^{\mathrm{T}} Q_k(\eta_2\xi_1 - \eta_1\xi_2)\right), \tag{B.10}$$

$$\Omega_L(Q_{k-1}, Q_k)(Z_1, Z_2) = \tfrac{1}{2}\,\mathrm{trace}\left(Q_k^{\mathrm{T}} Q_{k-1}(\mu_2\Lambda\eta_1 - \mu_1\Lambda\eta_2)\right). \tag{B.11}$$

Since $\psi_k : (Q_{k-1}, Q_k) \to (Q_k, P_k)$, with the derivative defined implicitly by (B.9), $\psi_k^*\Omega_{\mathfrak{gl}(n)} = \Omega_L$ is equivalent to showing that expressions (B.10) and (B.11) are equal. By evaluating (B.9) at $\eta = \eta_1, \xi = \xi_1, \mu = \mu_1$ and taking the trace inner product with η_2 yields

$$\mathrm{trace}\left(P_k^{\mathrm{T}} Q_k\eta_2\xi_1\right) - \mathrm{trace}\left(P_k^{\mathrm{T}} Q_k\eta_1\eta_2\right)$$
$$= -\mathrm{trace}\left(Q_k^{\mathrm{T}} Q_{k-1}\mu_1\Lambda\eta_2\right) + \mathrm{trace}\left(\Lambda Q_{k-1}^{\mathrm{T}} Q_k\eta_1\eta_2\right),$$

i.e.

$$\mathrm{trace}\left(P_k^{\mathrm{T}} Q_k\eta_2\xi_1\right) + \mathrm{trace}\left(Q_k^{\mathrm{T}} Q_{k-1}\mu_1\Lambda\eta_2\right) = \mathrm{trace}\left(\left(P_k^{\mathrm{T}} + \Lambda Q_{k-1}^{\mathrm{T}}\right) Q_k\eta_1\eta_2\right).$$

By subtracting this expression from itself, but with the indices interchanged yields

$$\Omega_{\mathfrak{gl}(n)}(Q_k, P_k)(X_1, X_2) - \Omega_L(Q_{k-1}, Q_k)(Z_1, Z_2)$$
$$= \tfrac{1}{2}\,\mathrm{trace}\left(\left(P_k^{\mathrm{T}} + \Lambda Q_{k-1}^{\mathrm{T}}\right) Q_k(\eta_1\eta_2 - \eta_2\eta_1)\right).$$

However, from (B.8) $(P_k^{\mathrm{T}} + \Lambda Q_{k-1}^{\mathrm{T}})Q_k$ is symmetric, and $\eta_1\eta_2 - \eta_2\eta_1$ is skew symmetric so

$$\Omega_{\mathfrak{gl}(n)}(Q_k, P_k)(X_1, X_2) = \Omega_L(Q_{k-1}, Q_k)(Z_1, Z_2). \qquad \square$$

References

Abraham R and Marsden J E 1978 *Foundations of Mechanics* 2nd edn (Reading, MA: Addison-Wesley) (reprinted by Persius 1997)

Anderson H 1983 Rattle: a velocity version of the shake algorithm for molecular dynamics calculations *J. Comput. Phy.* **52** 24–34

Bloch A M 1990 Steepest descent, linear programming and Hamiltonian flows *Contemp. Math. AMS* **114** 77–88

Bloch A M, Brockett R and Crouch P 1997 Double bracket equations and geodesic flows on symmetric spaces *Comm. Math. Phys.* **187** 357–73

Bloch A M, Brockett R W and Ratiu T S 1990a A new formulation of the generalized Toda Lattice equations and their fixed point analysis via the momentum map *Bull. Am. Math. Soc.* **23** 477–85

Bloch A M, Brockett R W and Ratiu T S 1992 Completely integrable gradient flows *Comm. Math. Phys.* **147** 57–74

Bloch A M and Crouch P 1996 Optimal control and geodesic flows *Systems Control Lett.* **28** 65–72

Bloch A M, Crouch P, Marsden J E and Ratiu T S 1998 Discrete rigid body dynamics and optimal control *Proc. IEEE Conf. on Decision and Control* **37** 2249–54

Bloch A M, Crouch P E, Marsden J E and Ratiu T S 2000 An almost Poisson structure for the generalized rigid body equations *Proc. IFAC Workshop on Lagrangian and Hamiltonian Method in Nonlinear Control* pp 92–7

Bloch A M, Flaschka H and Ratiu T S 1990b A convexity theorem for isospectral manifolds of Jacobi matrices in a compact Lie algebra *Duke Math. J.* **61** 41–65

Bobenko A I and Suris Y B 1999 Discrete time Lagrangian mechanics on Lie groups with an application to the Lagrange top *Comm. Math. Phys.* **204** 147–88

Bogayavlenski O 1994 *Breaking Solitons* (Berlin: Springer)

Brockett R 1989 Least squares matching problems *Linear Algebra Appl.* **122/123/124** 761–77

Brockett R 1994 The double bracket equation as the solution of a variational problem *Hamiltonian and Gradient Flows, Algorithms and Control* (Providence, RI: American Mathematical Society) pp 69–76

Bryson A E and Ho Y-C 1975 *Applied Optimal Control: Optimization, Estimation, and Control* (Washington, DC: Hemisphere Publishing)

Budd C and Iserles A 1999 Geometric integration: numerical solution of differential equations on manifolds *R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.* **357** 945–56

Cardoso J and Leite F 2001 The Moser–Veselov equation *Pre-Publicacoes do Departmento de Matematic, Universidade de Coimbra*

Crouch P and Grossman R 1993 Numerical integration of ordinary differential equations on manifolds *J. Nonlinear Sci.* **3** 1–33

Deift P, Li L C and Tomei C 1992 Loop groups, discrete versions of some classical integrable systems, and rank 2 extensions *Mem. Am. Math. Soc.* **100** 1–101

Faybusovich L E 1988 Explicitly solvable nonlinear optimal control problems *Int. J. Control.* **48** 2507–26

Federov Yu N 2002 Integrable flows and Backlund transformations on extended Stiefel varieties with applications to the Euler top on the Lie group $SO(3)$ *Preprint*

Federov Yu N and Kozlov V 1995 Various aspects of n-dimensional rigid body dynamics *Am. Math. Soc. Transl.* **168** 141–71

Holm D D, Marsden J E and Ratiu T 1986 The Hamiltonian structure of continuum mechanics in material, inverse material, spatial and convective representations, Les Presses de L'Université de Montrèal *Séminaire de Mathématiques Supérieurs* **100** 11–122

Iserles A, Mclachlan R and Zanna A 1999 Approximately preserving symmetries in the numerical integration of ordinary differential equations *Euro. J. Appl. Math.* **10** 419–45

Jurdjevic V 1997 *Geometric Control Theory (Cambridge Studies in Advanced Mathematics vol 52)* (Cambridge: Cambridge University Press)

Kane C, Marsden J E and Ortiz M 1999 Symplectic energy-momentum integrators *J. Math. Phys.* **40** 3353–71

Kane C, Marsden J E, Ortiz M and West M 2000 Variational integrators and the Newmark algorithm for conservative and dissipative mechanical systems *Int. J. Num. Math. Eng.* **49** 1295–325

Kucera V 1972 A contribution to matrix quadratic equations *IEEE Trans. Automatic Control* **17** 344–7

Lancaster P and Rodman L 1980 Existence and uniqueness theorems for the algebraic Riccati equation *Int. J. Control.* **32** 285–309

Lancaster P and Rodman L 1995 *Algebraic Riccati Equations* (New York: Oxford University Press)

Leimkuhler B and Skeel R 1994 Symplectic numerical integrators in constrained Hamiltonian systems *J. Comput. Phys.* **112** 117–25

Lewis D, Ratiu T S, Simo J C and Marsden J E 1992 The heavy top: a geometric treatment *Nonlinearity* **5** 1–48

Lewis D and Simo J C 1995 Conserving algorithms for the dynamics of Hamiltonian systems on Lie groups *J. Nonlinear Sci.* **4** 253–99

Manakov S V 1976 Note on the integration of Euler's equations of the dynamics of an n-dimensional rigid body *Funct. Anal. Appl.* **10** 328–9

Marsden J E, Patrick G W and Shkoller S 1998 Mulltisymplectic geometry, variational integrators and nonlinear PDEs *Comm. Math. Phys.* **199** 351–95

Marsden J E, Pekarsky S and Shkoller S 1999 Discrete Euler–Poincaré and Lie–Poisson equations *Nonlinearity* **12** 1647–62

Marsden J E and Ratiu T S 1999 *Introduction to Mechanics and Symmetry* 2nd edn (Berlin: Springer) *(Texts in Applied Mathematics vol 17, 1994)*

Marsden J E and Wendlandt J M 1997 Mechanical systems with symmetry, variational principles and integration algorithms *Current and Future Directions in Applied Mathematics* ed M Alber, B Hu and J Rosenthal (Bask: Birkhäuser) pp 219–61

Marsden J E and West M 2001 Discrete variational mechanics and variational integrators *Acta Numerica* **10** 357–514

McLaghlan R I and Scovel C 1995 Equivariant constrained symplectic integration *J. Nonlinear Sci.* **5** 233–56

Mishchenko A S and Fomenko A T 1976 On the integration of the Euler equations on semisimple Lie algebras *Sov. Math. Dokl.* **17** 1591–3

Mishchenko A S and Fomenko A T 1978 Generalized Liouville method of integration of Hamiltonian systems *Funct. Anal. Appl.* **12** 113–21

Moser J and Veselov A P 1991 Discrete versions of some classical integrable systems and factorization of matrix polynomials *Comm. Math. Phys.* **139** 217–43

Ratiu T 1980 The motion of the free n-dimensional rigid body *Indiana U. Math. J.* **29** 609–27

Reich S 1994 Momentum preserving symplectic integrators *Physica* D **76** 375–83

Suris Yu B 2001 Integrable discretizations of some cases of the rigid body dynamics *J. Nonlinear Math. Phys.* **8** 534–60

Veselov A P 1988 Integrable discrete-time systems and difference operators *Funct. Anal. Appl.* **22** 83–94

Wendlandt J M and Marsden J E 1997 Mechanical integrators derived from a discrete variational principle *Physica* D **106** 223–46