On an isospectral Lie–Poisson system and its Lie algebra

Anthony M. Bloch* Arieh Iserles†

January 29, 2005

Abstract

In this paper we analyse the matrix differential system $X' = [N, X^2]$, where N is skew-symmetric and $X(0)$ is symmetric. We prove that it is isospectral and that it is endowed with a Poisson structure, and discuss its invariants and Casimirs.

Formulation of the Poisson problem in a Lie–Poisson setting, as a flow on a dual of a Lie algebra, requires a computation of its faithful representation. Although the existence of a faithful representation is assured by the Ado theorem and a symbolic algorithm, due to de Graaf, exists for general computation of faithful representations of Lie algebras, the practical problem of forming a 'tight' representation, convenient for subsequent analytic and numerical work, belongs to numerical algebra. We solve it for the Poisson structure corresponding to the equation $X' = [N, X^2]$.

1 Introduction

The subject matter of this paper is the matrix system of ordinary differential equations

$$X' = [N, X^2], \quad t \geq 0, \qquad X(0) = X_0 \in \mathrm{Sym}(n), \tag{1.1}$$

where $N \in \mathfrak{so}(n)$ is given. Here $\mathrm{Sym}(n)$ and $\mathfrak{so}(n)$ denote the symmetric space of real $n \times n$ symmetric matrices and the Lie algebra of real $n \times n$ skew-symmetric matrices, respectively. It is easy to observe (and will be formally proved in the sequel) that $X(t) \in \mathrm{Sym}(n)$ for all $t \geq 0$.

Our interest in the system (1.1) is motivated by four reasons. Firstly, it is trivial to verify that (1.1) can be rewritten in the form

$$X' = [N, X]X + X[N, X], \quad t \geq 0, \qquad X(0) = X_0 \in \mathrm{Sym}(n), \tag{1.2}$$

*Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA, email abloch@umich.edu.

†Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, England, email A.Iserles@damtp.cam.ac.uk.

where $[A,B]=AB-BA$ is the standard matrix commutator. Since $[N,X]\in\mathrm{Sym}(n)$ for $N\in\mathfrak{so}(n)$, $X\in\mathrm{Sym}(n)$, the system (1.2) is a special case of of a *congruent flow*

$$X'=A(X)X+XA^\top(X),\quad t\geq 0,\qquad X(0)=X_0\in\mathrm{Sym}(n), \tag{1.3}$$

where $A:\mathrm{Sym}(n)\to\mathrm{M}(n)$ is sufficiently smooth: here $\mathrm{M}(n)$ denotes the set of real $n\times n$ matrices.

The terminology is explained as follows. Let V be the solution of the matrix differential system

$$V'=A(VX_0V^\top)V,\quad t\geq 0,\qquad V(0)=I.$$

Then, at least in an interval $[0,t^*)$ where V is nonsingular, we have $A(VX_0V^T)=V'V^{-1}$. Direct differentiation demonstrates that, letting

$$X(t)=V(t)X_0V^\top(t),\qquad 0\leq t<t^*, \tag{1.4}$$

we have

$$X'=V'X_0V^\top+VX_0{V'}^\top=A(X)X+XA^\top(X).$$

Since the initial condition $X(0)=X_0$ is satisfied as well, we thus deduce that (1.4) is the solution of (1.3). In other words, the solution of (1.1) is an outcome of the *general linear group* $\mathrm{GL}(n)$ acting on $\mathrm{Sym}(n)$ by congruence.

Once $\mathrm{GL}(n)$ acts on $\mathrm{Sym}(n)$ by congruence, the flow is automatically endowed with a number of invariants. Thus, the *signature* of $X(t)$ is constant, and so is the *angular field of values*

$$F'(X)=\{\boldsymbol{y}^*X\boldsymbol{y}\,:\,\boldsymbol{y}\in\mathbb{C}^n\setminus\{\boldsymbol{0}\}\}$$

(Horn & Johnson 1991).

Action by congruence *per se* is not very interesting. However, intriguingly, the differential system (1.1) is subject to another group action: it is acted on by the *special orthogonal group* $\mathrm{SO}(n)$ by similarity. Specifically, letting $B(X)=NX+XN$, we can write it in the form

$$X'=[B(X),X],\quad t\geq 0,\qquad X(0)=X_0\in\mathrm{Sym}(n). \tag{1.5}$$

Any system of the form (1.5), where $B(X):\mathrm{Sym}(n)\to\mathfrak{so}(n)$ (as it does in our case) is *isospectral:* the eigenvalues of $X(t)$ are invariant. The underlying reason, which we have already indicated, is an $\mathrm{SO}(n)$ action by similarity,

$$X(t)=Q(t)X_0Q^{-1}(t),\qquad t\geq 0, \tag{1.6}$$

where $Q\in\mathrm{SO}(n)$ is the solution of

$$Q'=B(QX_0Q^{-1})Q,\quad t\geq 0,\qquad Q(0)=I.$$

Note, incidentally, that, by virtue of orthogonality, $Q^{-1}=Q^\top$, therefore (1.6) is a special case of (1.4) and, by the same token, an isospectral flow is a special case of congruent flow. Having said so, isospectral flows have received a much greater attention than their congruent counterparts, since they feature in a number of important applications,

- The Toda lattice equations of near-neighbour interaction between unit-mass particles can be translated to this form. This is an important tool in their analysis (Flaschka 1974, Toda 1987).
- The familiar QR method for the computation of matrix eigenvalues can be interpreted as sampling a specific isospectral flow at unit intervals (Symes 1981/82). This insight into the connection between processes in numerical linear algebra and differential flows is fundamental and has spawned much further research (Deift, Nanda & Tomei 1983, Watkins 1984, Chu 1994).
- The identification of linear-algebraic processes with differential flows has led to new algorithms for several important problems. An excellent example is the work of Chu (1992) on the inverse eigenvalue problem for symmetric Toeplitz matrices. Another relevant reference is (Bloch, Brockett & Ratiu 1992).
- Letting $B(X) = [N, X]$, where $N \in \mathrm{Sym}(n)$, leads to the *double-bracket flows,* which have attracted a great deal of attention in the last decade (Brockett 1991, Chu & Driessel 1991, Bloch et al. 1992, Bloch, Brockett & Crouch 1997, Bloch & Iserles 2003).

To recap, the differential system (1.1) is evolving under two distinct group actions. This makes it fairly unusual (although by no means unique) and worthy of further analysis.

Another aspect of equations (1.1) that renders them interesting is that they are, in a sense, dual to the *generalized rigid-body equations*

$$M' = [\Omega, M], \quad t \geq 0, \qquad M(0) \in \mathfrak{so}(n),$$

where $M = \Omega J + J\Omega$, $J \in \mathrm{Sym}(n)$ (therefore $\Omega \in \mathfrak{so}(n)$) (Manakov 1976). This point will be further elaborated in Section 3.

Our final motivation for the study of (1.1) originates in the apparent behaviour of their solutions. Thus, for example, we have computed numerically the solution of the system for

$$N = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}.$$

In Fig. 1 we display the phase portraits of the solution. Clearly, the solution displays a great deal of regularity and apparently it evolves on invariant tori. Similar behaviour is obtained for other matrices N, also in larger number of dimensions.

It is fair to say that this kind of regular behaviour seldom arises at random. Typically it is an indication of a deeper structure, often either Hamiltonian or related to Hamiltonian. The purpose of this paper is to probe and understand this structure.

In Section 2 we prove that the matrix system (1.1) can be at the first instance written as a set of almost-Poisson equations,

$$\boldsymbol{x}' = S(\boldsymbol{x})\boldsymbol{\nabla} H(\boldsymbol{x}), \quad t \geq 0, \qquad \boldsymbol{x}(0) = \boldsymbol{x}_0 \in \mathbb{R}^{\frac{1}{2}n(n+1)}. \tag{1.7}$$

Here $\boldsymbol{x}$ is the upper-triangular portion of the symmetric matrix X, 'stretched' as a vector, the smooth function $S \in \mathfrak{so}(\frac{1}{2}n(n+1))$ is linear and homogeneous and the

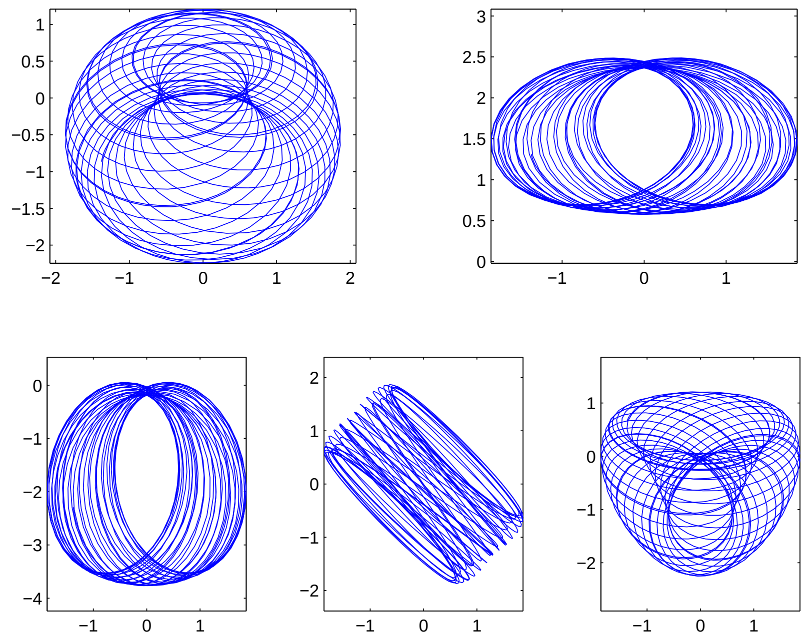

Figure 1: The phase portraits $(x_{1,2}, x_{k,l})$ for $(k,l) = (1,1),(1,3),(2,2),(2,3),(3,3)$, with a random initial condition.

Hamiltonian H is smooth and nonnegative. Moreover, we verify that the *structure matrix* S obeys the *Jacobi identity:* for every $p,q,r = 1,2,\ldots,\frac{1}{2}n(n+1)$ it is true that

$$\sum_{k=1}^{\frac{1}{2}n(n+1)} \left[\frac{\partial S_{p,q}}{\partial x_k} S_{k,r} + \frac{\partial S_{q,r}}{\partial x_k} S_{k,p} + \frac{\partial S_{r,p}}{\partial x_k} S_{k,q} \right] = 0. \tag{1.8}$$

Therefore

$$\{f,g\} = (\boldsymbol{\nabla} f)^\top S (\boldsymbol{\nabla} g)$$

defines a Poisson bracket and (1.7) is a *Poisson system.* This has important implications, which we discuss briefly and which will be discussed more comprehensively elsewhere.

Section 3 is devoted to a preliminary investigation of a critical aspect of Poisson flows, focusing on their invariants and Casimirs. We prove that (1.1) has $\approx \frac{1}{4}n^2$ invariants and at least two Casimirs: $\det X$ and $\mathbf{1}^\top N X \mathbf{1}$.

The main computational content of the paper is Section 4. A Poisson system can be written in a Lie–Poisson form, as a flow over an orbit in a dual $\mathfrak{g}^*$ of a Lie algebra $\mathfrak{g}$ which is determined by the elements of the structure matrix S. Specifically, recalling that S is linear and homogeneous in $\boldsymbol{x}$, there exist *structure constants* $c_{p,q}^r$,

$p, q, r = 1, 2, \ldots, m$, where $m = \frac{1}{2}n(n+1)$, such that

$$S_{p,q}(\boldsymbol{x}) = \sum_{r=1}^{m} c_{p,q}^{r} x_r, \qquad p, q = 1, 2, \ldots, m$$

Skew-symmetry of S is equivalent to

$$c_{p,q}^{r} + c_{q,p}^{r} = 0, \qquad p, q, r = 1, \ldots, m, \tag{1.9}$$

while the Jacobi identity (1.8) is equivalent to

$$\sum_{i=1}^{m} (c_{p,q}^{i} c_{i,r}^{s} + c_{q,r}^{i} c_{i,p}^{s} + c_{r,p}^{i} c_{i,q}^{s}) = 0, \qquad p, q, r, s = 1, 2, \ldots, n. \tag{1.10}$$

Now, (1.9) and (1.10) are precisely the conditions for the $\{c_{p,q}^{r}\}$ to form structure constants of a finite-dimensional Lie algebra (Olver 1995). Let us denote this algebra by $\mathfrak{g}$ and assume that $\{E_1, E_2, \ldots, E_m\}$ is its basis. Moreover, let us suppose that $\{F_1, F_2, \ldots, F_m\}$ form the basis of the dual $\mathfrak{g}^*$ (the linear space of linear functionals acting on $\mathfrak{g}$) and that $\langle F_k, E_l \rangle = \delta_{k,l}$, $k, l = 1, 2, \ldots, m$, where

$$\langle Y, Z \rangle = \mathrm{tr}\,(Y^{\top} Z)$$

and $\delta_{k,l}$ is the familiar Kronecker delta. The following construction is general to all Poisson systems (Marsden & Ratiu 1994). We associate with $\boldsymbol{x}$, the solution of (1.7), an element of $\mathfrak{g}^*$, namely

$$Y(t) = \sum_{k=1}^{m} x_k(t) F_k.$$

With minor abuse of notation, we let $H(Y) = H(\boldsymbol{x})$. Then (1.7) is equivalent to

$$Y' = -\mathrm{ad}^{\,*}_{H'(Y)} Y, \quad t \ge 0, \qquad Y(0) = \sum_{k=1}^{m} x_k(0) F_k, \tag{1.11}$$

where $\mathrm{ad}^{\,*}$ is the adjoint operator in $\mathfrak{g}^*$ and $H'(Y) = (\partial H(Y)/\partial Y_{p,q})_{p,q=1,\ldots,m}$.

The formulation of a Poisson system as a flow in $\mathfrak{g}^*$ is advantageous for a number of analytic and numerical reasons (Engø & Faltinsen 2001, Lewis & Simo 1994, Marsden & Ratiu 1994). However, practical implementation of (1.11) requires a representation of $F_1, F_2, \ldots, F_m$ by matrices. In principle, this can be done since, by the *Ado theorem,* every finite-dimensional Lie algebra possesses a faithful representation (Varadarajan 1984). Yet, the Ado theorem falls short of providing such a representation in an explicit form. In principle, it is possible to use symbolic algebra to this end: in a beautiful paper, de Graff (1997) demonstrated how to render Ado's original proof into a constructive algorithm. Yet, this falls short of our requirements. Firstly, the algorithm delivers the E_ks, but not a biorthogonal basis of $\mathfrak{g}^*$. Secondly, and more importantly, the size of matrices in de Graff's algorithm increases exponentially with m, while practical work with (1.11) requires either a minimal representation or, at the very least, one which is fairly small. In Section 4 we introduce a numerical, linear-algebraic algorithm that performs this task for our specific Lie–Poisson flow and produces matrices of size $(2n-2)\times(2n-2)$.

Inasmuch as this is a paper on a specific Lie–Poisson system and the representation of 'its' Lie algebra, we address ourselves to a particular instance of a considerably more general problem. Given structure constants $\{c_{p,q}^r\}$ that obey (1.9) and (1.10), hence being within the conditions of the Ado theorem, determine a *small* faithful representation of the underlying free Lie algebra. This is, essentially, a problem of numerical algebra which, to our knowledge, has been never addressed by numerical algebraists. A major goal of the present paper is to demonstrate that, at least in one case, this problem is tractable. The general case remains as a substantive challenge.

2 From a differential flow to a Poisson system

An ordinary differential system is said to be *almost Poisson* if it is of the form

$$\boldsymbol{x}' = S(\boldsymbol{x})\boldsymbol{\nabla} H(\boldsymbol{x}), \quad t\geq 0, \qquad \boldsymbol{x}(0)=\boldsymbol{x}_0\in\mathbb{R}^m, \tag{2.1}$$

where $H:\mathbb{R}^m\to\mathbb{R}$ (the *Hamiltonian*) and the linear, homogeneous matrix function $S:\mathbb{R}^m\to\mathfrak{so}(m)$ (the *structure matrix*) are sufficiently smooth. The constants $\{c_{p,q}^r\}_{p,q,r=1,\ldots,m}$ such that

$$S_{p,q}(\boldsymbol{x}) = \sum_{r=1}^m c_{p,q}^r x_r, \qquad p,q=1,\ldots,m,$$

are called the *structure constants* of (2.1). If, in addition, S obeys the Jacobi identity (1.8) (with $\frac{1}{2}n(n+1)$ replaced by m), (2.1) is said to be a *Poisson* (or, alternatively, *Kostant–Kirillov–Lie–Soriau*) *system* (Marsden & Ratiu 1994). In that case, as already stated in Section 1, the structure constants obey the identities

$$c_{p,q}^r + c_{q,p}^r = 0, \qquad p,q,r=1,\ldots,m, \tag{2.2}$$

$$\sum_{i=1}^m (c_{p,q}^i c_{i,r}^s + c_{q,r}^i c_{i,p}^s + c_{r,p}^i c_{i,q}^s) = 0, \qquad p,q,r,s=1,\ldots,m. \tag{2.3}$$

Our contention is that (1.1) is a Poisson system. To this end, we first demonstrate that it is almost Poisson. Writing (1.1) in a coordinate-by-coordinate fashion, it becomes

$$x_{p,q}' = \sum_{r=1}^n\sum_{s=1}^n (n_{p,r}x_{r,s}x_{q,s} - n_{r,s}x_{p,r}x_{q,s} - n_{r,s}x_{p,s}x_{q,r} + n_{q,r}x_{p,s}x_{r,s})$$

for $p,q=1,\ldots,n$. By virtue of $N\in\mathfrak{so}(n)$, however, once we swap the indices r and s we have

$$\sum_{r=1}^n\sum_{s=1}^n n_{r,s}x_{p,s}x_{q,r} = \sum_{s=1}^n\sum_{r=1}^n n_{s,r}x_{p,r}x_{q,s} = -\sum_{r=1}^n\sum_{s=1}^n n_{r,s}x_{p,r}x_{q,s}.$$

Thus, taking into account the symmetry of X, we obtain $m = \frac{1}{2}n(n+1)$ equations

$$\begin{aligned} x'_{p,q} = &\sum_{r=1}^{n} (n_{p,r}x_{q,r} + n_{q,r}n_{p,r})x_{r,r} \\ &+ 2\sum_{r=1}^{n-1}\sum_{s=r+1}^{n} (n_{p,r}x_{q,s} + n_{q,r}x_{p,s})x_{r,s}, \qquad (p,q) \in \mathbb{J}_n, \end{aligned} \tag{2.4}$$

where

$$\mathbb{J}_n = \{(p,q) \,:\, 1 \le p \le q \le n\}.$$

Writing (1.1) in an almost-Poisson form, we need to 'stretch' $\{x_{p,q}\}_{(p,q)\in\mathbb{J}_n}$ into a vector $\boldsymbol{x} \in \mathbb{R}^m$. It is helpful to retain both indices of the elements of X, while ordering them lexicographically: for example, for $n = 3$ we have $m = 6$ and

$$\boldsymbol{x} = [\; x_{1,1} \quad x_{1,2} \quad x_{1,3} \quad x_{2,2} \quad x_{2,3} \quad x_{3,3} \;]^\top.$$

Note that we may occasionally use $x_{p,q}$ with $p > q$, since this makes our presentation more transparent. In that case, of course, $x_{p,q} = x_{q,p}$.

We set

$$H(\boldsymbol{x}) = \tfrac{1}{2}\|X\|_{\mathsf{F}}^2 = \tfrac{1}{2}\sum_{r=1}^{n} x_{r,r}^2 + \sum_{r=1}^{n-1}\sum_{s=r+1}^{n} x_{r,s}^2, \tag{2.5}$$

hence (2.4) can be rewritten in the form

$$x'_{p,q} = \sum_{r=1}^{n}\sum_{s=r}^{n} (n_{p,r}x_{q,s} + n_{q,r}x_{p,s})\frac{\partial H(\boldsymbol{x})}{\partial x_{r,s}}, \qquad (p,q) \in \mathbb{J}_n.$$

This form looks tantalisingly similar to (2.1), except that letting $S_{(p,q),(r,s)}(\boldsymbol{x}) = n_{p,r}x_{q,s} + n_{q,r}x_{p,s}$ does not lead to a skew-symmetric matrix. However, identifying $x_{q,p} = x_{p,q}$ yields

$$x'_{p,q} = x'_{q,p} = \sum_{r=1}^{n}\sum_{s=r}^{n} (n_{q,r}x_{p,s} + n_{p,r}x_{q,s})\frac{\partial H(\boldsymbol{x})}{\partial x_{r,s}}, \qquad (p,q) \in \mathbb{J}_n.$$

We now average the two expressions: the outcome is

$$x'_{p,q} = \tfrac{1}{2}\sum_{(r,s)\in\mathbb{J}_n} (n_{p,r}x_{q,s} + n_{p,s}x_{q,r} + n_{q,r}x_{p,s} + n_{q,s}x_{p,r})\frac{\partial H(\boldsymbol{x})}{\partial x_{r,s}}, \qquad (p,q) \in \mathbb{J}_n.$$

This is consistent with the almost-Poisson form (2.1), once we define the Hamiltonian H by (2.5) and let

$$S_{(p,q),(r,s)}(\boldsymbol{x}) = \tfrac{1}{2}(n_{p,r}x_{q,s} + n_{p,s}x_{q,r} + n_{q,r}x_{p,s} + n_{q,s}x_{p,r}), \qquad (p,q),(r,s) \in \mathbb{J}_n. \tag{2.6}$$

Theorem 1 *The system (2.1) with the Hamiltonian (2.5) and the structure matrix (2.6) is Poisson.*

Proof The matrix S is linear and homogeneous in $\boldsymbol{x}$. Moreover, $N \in \mathfrak{so}(n)$ and $X \in \mathrm{Sym}(n)$ imply at once that $S \in \mathfrak{so}(m)$. The system is thus almost Poisson.

The structure constants associated with S are

$$c^{(k,l)}_{(p,q),(r,s)} = \tfrac{1}{2}[\delta^{(k,l)}_{(q,s)} n_{p,r} + \delta^{(k,l)}_{(q,r)} n_{p,s} + \delta^{(k,l)}_{(p,s)} n_{q,r} + \delta^{(k,l)}_{(p,r)} n_{q,s}], \qquad (p,q),(r,s),(k,l) \in \mathbb{J}_n, \tag{2.7}$$

where $\delta^{(k,l)}_{(i,j)}$ is the 'symmetrized' Kronecker's delta,

$$\delta^{(k,l)}_{(i,j)} = \begin{cases} 1, & (i,j) = (k,l) \text{ or } (i,j) = (l,k), \\ 0, & \text{otherwise.} \end{cases}$$

It is possible to prove directly that the above structure constants satisfy the Jacobi condition (2.3). This requires a great deal of fairly tedious algebra. Instead, we follow a (gratefully acknowledged) suggestion of Peter Olver, which identifies the structure constants (2.7) with an unusual Lie algebra. Let $Y, Z \in \mathrm{Sym}(n)$, $N \in \mathfrak{so}(n)$, and let

$$[Y, Z]_N = YNZ - ZNY. \tag{2.8}$$

This is a proper Lie bracket, since $[Z, Y]_N = -[Y, Z]_N$ and

$$\begin{aligned} &[X,[Y,Z]_N]_N + [Y,[Z,X]_N]_N + [Z,[X,Y]_N]_N \\ &= \{XN(YNZ - ZNY) - (YNZ - ZNY)NX\} + \{YN(ZNX - XNZ) \\ &\quad - (ZNX - XNZ)NY\} + \{ZN(XNY - YNX) - (XNY - YNX)NZ\} \\ &= O \end{aligned}$$

– hence the Jacobi identity. Since $\mathrm{Sym}(n)$ is a linear space, it follows at once that, once accompanied by the bracket (2.8), it is a Lie algebra, $\mathfrak{h}_n(N)$, say.

We now define

$$H_{p,q} = \tfrac{1}{2}(\boldsymbol{e}_p \boldsymbol{e}_q^\top + \boldsymbol{e}_q \boldsymbol{e}_p^\top), \qquad (p,q) \in \mathbb{J}_n, \tag{2.9}$$

where $\boldsymbol{e}_k \in \mathbb{R}^n$ is the nth unit vector, and note that $\{H_{p,q} : (p,q) \in \mathbb{J}_n\}$ form a basis of $\mathrm{Sym}(n)$. A straightforward computation confirms that

$$[H_{p,q}, H_{r,s}]_N = \tfrac{1}{2}(n_{p,r}H_{q,s} + n_{p,s}H_{q,r} + n_{q,r}H_{p,s} + n_{q,s}H_{p,r}), \qquad (p,q),(r,s) \in \mathbb{J}_n,$$

where we identify $H_{p,q} = H_{q,p}$. Therefore, (2.7) are the structure constants associated with the underlying basis. Hence they satisfy (2.3) (with obvious amendments to cater for 'double' indices), we deduce the Jacobi identity and conclude that the system (2.1) is indeed Poisson. □

We are not aware of any applications of the bracket (2.8), although they might well exist.

In Section 4 we are concerned with a faithful representation of the Lie algebra associated with the structure constants (2.7). Here we just state the obvious: (2.9) is *not* a representation, faithful or otherwise: a representation is defined with the standard matrix commutator $[Y, Z] = YZ - ZY$. To the contrary! Once we derive a faithful representation of the Lie algebra associated with the structure constants (2.7), motivated by our goal to represent (1.1) as a Lie–Poisson flow, we simultaneously obtain *gratis* a faithful representation of $\mathfrak{h}_n(N)$.

3 Invariants

3.1 Lax pairs and their consequences

It is of interest to compute the invariants for the flow of the system (1.1). We bear in mind that, by virtue of the isospectral representation (1.2), we already know that the eigenvalues of X, or alternatively $\operatorname{tr} X^k$ for $k = 1, 2, \ldots, n-1$, are invariants.

One way to compute additional invariants is to rewrite the system as a Lax pair with parameter. One can do this is a fashion which is similar to that for the generalized rigid body equations (see (Manakov 1976)).

Theorem 2 *Let λ be an arbitrary time independent parameter with values in $\mathbb{R}$. The system (1.5) is equivalent to the Lax pair system*

$$(X + \lambda N)' = [NX + XN + \lambda N^2, X + \lambda N] \tag{3.1}$$

Proof The proof is a computation. In particular we can verify that the coefficient of λ on the right hand side of equation (3.1) is given by

$$\begin{aligned} & \lambda \left([NX + XN, N] + [N^2, X] \right) \\ = \; & \lambda (NXN + XN^2 - N^2X - NXN + N^2X - XN^2) \end{aligned}$$

which is zero. □

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

$$M' = [M, \Omega], \quad t \geq 0, \qquad M(0) = M_0 \in \mathfrak{so}(n), \tag{3.2}$$

where $Q \in \mathrm{SO}(n)$ denotes the configuration space variable (the attitude of the body), $\Omega = Q^{-1}Q' \in \mathfrak{so}(n)$ 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, 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 $I_1 = \Lambda_2 + \Lambda_3$, $I_2 = \Lambda_3 + \Lambda_1$, $I_3 = \Lambda_1 + \Lambda_2$.

In this case the generalized rigid body equations as Lax equations with parameter due to Manakov take the form

$$\frac{\mathrm{d}}{\mathrm{d}t}(M + \lambda\Lambda^2) = [M + \lambda\Lambda^2, \Omega + \lambda\Lambda]. \tag{3.3}$$

Note the contrast with our setting: in the Manakov case the system matrix M is in $\mathfrak{so}(n)$ and Λ is a symmetric matrix while in our case X is symmetric and $N \in \mathfrak{so}(n)$.

For the generalized rigid body 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 $\mathrm{SO}(n)$ (identified with the corresponding coadjoint orbit).

Similarly in our case the nontrivial coefficients of λ in the traces of the powers of $X+\lambda N$ yield conserved quantities of motion:

$$\mathrm{tr}\,(X+\lambda N)^k = \mathrm{const}, \qquad k=1,2,\ldots,n-1, \quad \lambda\in\mathbb{R}.$$

Therefore, expanding the power and assembling equal powers of λ,

$$\mathrm{tr}\sum_{|\boldsymbol{i}|=r}\sum_{|\boldsymbol{j}|=k-r} X^{i_1}N^{j_1}X^{i_2}\cdots X^{i_s}N^{j_s} = \mathrm{const}, \qquad r=1,\ldots,k, \quad k=1,\ldots,n-1.$$

Here $\boldsymbol{i}$ is a multi-index and $|\boldsymbol{i}|=\sum_q i_q$. However, since a trace of a matrix equals the trace of its transpose, $X\in\mathrm{Sym}(n)$ and $N\in\mathfrak{so}(n)$, it is true that

$$\mathrm{tr}\,X^{i_1}N^{j_1}X^{i_2}\cdots X^{i_s}N^{j_s} = (-1)^{|\boldsymbol{j}|}\mathrm{tr}\,N^{j_s}X^{j_s}\cdots X^{i_2}N^{j_1}X^{i_1}.$$

Therefore, if $k-s$ is odd then necessarily

$$\mathrm{tr}\sum_{|\boldsymbol{i}|=r}\sum_{|\boldsymbol{j}|=k-r} X^{i_1}N^{j_1}X^{i_2}\cdots X^{i_s}N^{j_s} = 0$$

and only even $k-s$ qualifies as an invariant. Thus, we are left with the invariants

$$\mathrm{tr}\sum_{|\boldsymbol{i}|=k-2r}\sum_{|\boldsymbol{j}|=2r} X^{i_1}N^{j_1}X^{i_2}\cdots X^{i_s}N^{j_s}, \quad k=1,\ldots,n-1, \quad r=0,1,\ldots,\lfloor\tfrac{k-1}{2}\rfloor. \tag{3.4}$$

Altogether, this results in

$$\lfloor\tfrac{n+1}{2}\rfloor\times\lfloor\tfrac{n+2}{2}\rfloor$$

invariants. As things stand, we are not aware whether they are in involution, indeed even if they are independent of each other. Therefore, although the number is right, we cannot deduce integrability of the system (1.1). Independence and involution of these integrals will be addressed in a companion paper to this one.

3.2 A generalized system

The flow (1.1) can be rendered more general by complexification. Generalizing it to evolution in $\mathfrak{su}(n)$ yields an n^2-dimensional flow of generalized rigid body type with two natural Hamiltonian structures: Let $X_0\in\mathfrak{su}(n)$ and $N\in\mathrm{Sym}(n,\mathbb{R})$ and consider

$$X' = [X^2,N] = [X, XN+NX], \quad t\ge 0, \qquad X(0)=X_0. \tag{3.5}$$

Note that $X(t)$ evolves in $\mathfrak{su}(n)$,

$$(X')^* = (X^2N-NX^2)^* = NX^{*2}-X^{*2}N = NX^2-X^2N = -X' \quad\Rightarrow\quad X'\in\mathfrak{su}(n).$$

Moreover, one can generalize this still further and take a complex Hermitian N.

We define

$$H_1(X) = \tfrac{1}{4}\mathrm{tr}\, X(XN+NX),$$
$$H_2(X) = \tfrac{1}{2}\mathrm{tr}\, X^2.$$

Note that both Hamiltonians are real and that H_2 gives us our earlier Hamiltonian in the case that X is symmetric but that H_1 is zero in this case. One can show that they both give raise to the flow (3.5), of course with different Poisson structures.

This dual Hamiltonian structure in this setting will be explored further in a related publication.

3.3 Casimirs

It is natural to attempt an analysis of Casimirs for our system with our given Poisson structure. Recall that given two smooth functions $f, g : \mathbb{R}^m \to \mathbb{R}$ and the skew-symmetric function $S(y)$ from (2.6), the *Poisson bracket* is

$$\{f,g\} = [\boldsymbol{\nabla} f(\boldsymbol{y})]^\top S(\boldsymbol{y})\boldsymbol{\nabla} g(\boldsymbol{y}) = \sum_{i=1}^{m}\sum_{j=1}^{m}\frac{\partial f(\boldsymbol{y})}{\partial y_i}S_{i,j}(\boldsymbol{y})\frac{\partial g(\boldsymbol{y})}{\partial y_j}.$$

A *Casimir* of the Lie–Poisson system (2.1) is a function C which is in involution (with respect to the above Lie bracket) with all other functions on the Poisson manifold (Marsden & Ratiu 1994). According to the Darboux theorem, if $S(\boldsymbol{x})$ is (locally) of constant even rank $m-\alpha$, say, where $\alpha \geq 1$, then there exist s Casimirs, $c_1, c_2, \ldots, c_\alpha$, say, which are themselves in involution. They satisfy

$$\{c_k, g\} = 0 \qquad \text{for all smooth } g. \tag{3.6}$$

Each Casimir is a *first integral* of the Lie–Poisson system.

Extensive experimentation with MATLAB, generating a large number of matrices S using random $N \in \mathfrak{so}(n)$ and $X \in \mathrm{Sym}(n)$, seems to indicate that $\alpha = \lfloor (n+1)/2 \rfloor$. Note that, in that case, $m - \alpha = 2\lfloor n/2 \rfloor \times \lfloor (n+1)/2 \rfloor$ is indeed even.

Lemma 3 *Suppose that $n \geq 3$ and $X_0, N \neq O$, where O is the zero matrix. Then it is true for the system (2.1) with the structure matrix (2.6) that $\alpha \geq 2$.*

Proof We will demonstrate that $\alpha \geq 2$ by singling out two linearly-independent eigenvectors $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^m$ that correspond to a zero eigenvalue, $S\boldsymbol{v} = S\boldsymbol{w} = \boldsymbol{0}$.

We first assume that both X and N are nonsingular. The elements of the eigenvectors will be denoted by the usual double indices. We let first

$$v_{p,q} = \begin{cases} \frac{1}{2}(X^{-1})_{p,p}, & p = q, \\ (X^{-1})_{p,q}, & p < q. \end{cases}$$

Therefore, using (2.6) and exploiting as necessary symmetry of X and V and skew-symmetry of N,

$$(S\boldsymbol{v})_{p,q} = \sum_{(r,s)\in\mathbb{J}_n} S_{(p,q),(r,s)} v_{r,s}$$

$$
\begin{aligned}
&= \tfrac{1}{2}\sum_{r=1}^{n}\sum_{s=r}^{n}(n_{p,r}x_{q,s}+n_{p,s}x_{q,r}+n_{q,r}x_{p,s}+n_{q,s}x_{p,r})v_{r,s}\\
&= \tfrac{1}{4}\sum_{r=1}^{n}\sum_{s=1}^{n}(n_{p,r}x_{q,s}+n_{p,s}x_{q,r}+n_{q,r}x_{p,s}+n_{q,s}x_{p,r})(X^{-1})_{r,s}\\
&= \tfrac{1}{4}\left[\sum_{r=1}^{n}n_{p,r}\sum_{s=1}^{n}x_{q,s}(X^{-1})_{s,r}+\sum_{s=1}^{n}n_{p,s}\sum_{r=1}^{n}x_{q,r}(X^{-1})_{r,s}\right.\\
&\qquad\left.+\sum_{r=1}^{n}n_{q,r}\sum_{s=1}^{n}x_{p,s}(X^{-1})_{s,r}+\sum_{s=1}^{n}n_{q,s}\sum_{r=1}^{n}x_{p,r}(X^{-1})_{r,s}\right]\\
&= \tfrac{1}{4}(n_{p,q}+n_{p,q}+n_{q,p}+n_{q,p})=0.
\end{aligned}
$$

Therefore, indeed, $S\boldsymbol{v}=\mathbf{0}$.

Likewise, we set

$$
w_{p,q}=\begin{cases}\tfrac{1}{2}(N^{-1})_{p,p}, & p=q,\\ (N^{-1})_{p,q}, & p<q.\end{cases}
$$

The proof of $S\boldsymbol{w}=\mathbf{0}$ follows in an identical manner, because each $S_{(p,q),(r,s)}$ is symmetric in X and N.

The requirement that X and N are nonsingular can be easily lifted, replacing the inverse by an adjugate matrix, as necessary: because of our assumptions, $\boldsymbol{v},\boldsymbol{w}\neq\mathbf{0}$. Since $X\in\mathrm{Sym}(n)$ cannot be a multiple of $N\in\mathfrak{so}(n)$, the vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ are linearly independent. This concludes the proof. $\square$

Note that, in a sense, $\boldsymbol{v}$ and $\boldsymbol{w}$ are 'dual' to each other: while $\boldsymbol{v}=\boldsymbol{v}(\boldsymbol{x})$, but is independent of N, $\boldsymbol{w}$ is dependent solely on N.

The following result identifies two Casimirs of the system (1.1).

Theorem 4 *Let $c_1(X)=\det X$. Then for every $(p,q)\in\mathbb{J}_n$ it is true that*

$$
\begin{aligned}
\frac{1}{c_1(X)}\frac{\partial c_1(X)}{\partial x_{p,p}}&=(X^{-1})_{p,p}, \qquad k=1,2,\ldots,n\\
\frac{1}{c_1(X)}\frac{\partial c_1(X)}{\partial x_{p,q}}&=2(X^{-1})_{p,q}
\end{aligned}
$$

and c_1 is a Casimir of the system (1.1). Likewise,

$$
c_2(X)=\tfrac{1}{2}\mathbf{1}^\top(\operatorname{adj}N)X\mathbf{1},
$$

where adj N *is the adjugate matrix of N, is a Casimir of (1.1) for $N\neq O$. (We may replace* adj N *by N^{-1} if N is nonsingular.)*

Proof We denote by $A_{p,q}$ a matrix A from which we have excised the pth column and the qth row. Thus,

$$
(X^{-1})_{p,q}=(-1)^{p+q}\frac{\det X_{q,p}}{\det X}=(-1)^{p+q}\frac{\det X_{p,q}}{c_1(X)}.
$$

Therefore, it is enough to prove that

$$\frac{\partial c_1(X)}{\partial x_{p,p}} = \det X_{p,p}, \qquad \frac{\partial c_1(X)}{\partial x_{p,q}} = 2(-1)^{p+q}\det X_{p,q}, \quad p \neq q.$$

For $p = l$ the element $x_{p,p}$ appears just one in X and taking the derivative of $c_1(X) = \det X$ with respect to $x_{p,p}$ replaces the pth row and the pth column by $\boldsymbol{e}_p^\top$ and $\boldsymbol{e}_p$ respectively. Thus, $\partial \det X / \partial x_{p,p} = \det X_{p,p}$.

For $p \neq q$ the term $x_{p,q}$ appears twice in X and the derivative is thus a sum of two determinants: one when the pth row is replaced by $\boldsymbol{e}_p^\top$ and the qth column by $\boldsymbol{e}_q$ and the other where the pth row and the qth column are replaced by $\boldsymbol{e}_q^\top$ and $\boldsymbol{e}_p$ respectively. Because of symmetry, both determinants equal $(-1)^{p+q}\det X_{p,q}$. This completes the proof.

Consequently, $c_1(X) = \det X$ is a Casimir iff

$$\sum_{r=1}^{n} S_{(p,q),(r,r)}(X^{-1})_{r,r} + 2\sum_{r=1}^{n-1}\sum_{s=r+1}^{n} S_{(p,q),(r,s)}(X^{-1})_{r,s} = 0$$

for every $(p,q) \in \mathbb{J}_n$. This, however, follows at once from the proof of Lemma 3, once we observe that partial derivatives of c_1, scaled by $c_1^{-1}(X)$, are precisely the elements of the vector $\boldsymbol{v}$ therein.

The proof that c_2 is a Casimir follows from the observation that, once we consider the 'reduced' system with $m = \frac{1}{2}n(n+1)$,

$$\frac{\partial c_2(X)}{\partial x_{p,q}} = \begin{cases} \frac{1}{2}(\operatorname{adj} N)_{p,p}, & p = q, \\ (\operatorname{adj} N)_{p,q}, & p < q, \end{cases} \qquad (p,q) \in \mathbb{J}_n,$$

which, according to Lemma 3, is an eigenvector of S: the proof therein was for $N^{-1} = (\det N)^{-1}\operatorname{adj} N$ but, as already remarked, it is equally valid for the adjugate matrix. □

Of course, we already know that $\det X$ is an invariant of (1.1), since the latter's equivalence with the isospectral system (1.2) implies that the product of the eigenvalues of X is constant. This, of course, does not mean that *all* invariants implied by isospectrality are Casimirs. For example, letting $c(X) = \operatorname{tr} X$, we readily have

$$\sum_{(r,s)\in\mathbb{J}_n} S_{(p,q),(r,s)} \frac{\partial c(X)}{\partial x_{r,s}} = \sum_{r=1}^{n} S_{(p,q),(r,r)} = \sum_{r=1}^{n}(n_{p,r}x_{q,r} + n_{q,r}x_{p,r}) = [N, X]_{p,q},$$

which cannot be expected to vanish for all $(p,q) \in \mathbb{J}_n$.

Needless to say, we do not claim that c_1 and c_2 are the *only* Casimirs of (1.1). The full resolution of integrability, Casimirs and other invariants of the system are deferred to a future paper.

4 Lie-algebraic representation and its computation

We proved in Section 2 that $\{c_{(p,q),(r,s)}^{(k,l)} : (p,q),(r,s),(k,l) \in \mathbb{J}_n\}$ obey (2.2) and (2.3), hence are structure constants of an m-dimensional algebra $\mathfrak{g}$. This algebra is defined formally as a *free Lie algebra*

$$\mathfrak{g} = \mathrm{FLA}(E_{p,q}\,:\,(p,q)\in\mathbb{J}_n), \tag{4.1}$$

where $E_{p,q}$ are (for the time being) purely formal letters, equipped with the commutation relation

$$[E_{p,q},E_{r,s}] = \tfrac{1}{2}(n_{p,r}E_{q,s}+n_{p,s}E_{p,s}E_{q,r}+n_{q,r}E_{p,s}+n_{q,s}E_{p,r}), \qquad (p,q),(r,s)\in\mathbb{J}_n \tag{4.2}$$

(cf. (2.6) or (2.7)). In other words, elements of $\mathfrak{g}$ are linear combinations of the $E_{p,q}$s (which are presumed linearly independent) and its closure under commutation is assured by (4.2).

4.1 The algorithm

We recall from Section 1 that our ultimate goal it to construct representations of $\mathfrak{g}$ and of $\mathfrak{g}^*$ which are orthogonal to each other with respect to the Frobenius inner product $\langle Y,Z\rangle = \mathrm{tr}\,(Y^\top Z)$. Note, however, that once we determine a faithful representation of $\mathfrak{g}$ which is *orthogonal* – for every $(p,q),(r,s)\in\mathbb{J}_n$ $\langle E_{p,q},E_{r,s}\rangle$ is zero if $(p,q)\neq(r,s)$ and $\pi_{p,q}>0$ if $(p,q)=(r,s)$ – we may take $F_{p,q}=\pi_{p,q}^{-1}E_{p,q}$. This follows at once from the Riesz representation theorem for linear functionals once we observe that $\langle\,\cdot\,,\,\cdot\,\rangle$ is the standard Euclidean vector inner product, hence defining a Hilbert space.

Although, as mentioned in Section 1, the existence of a faithful representation of $\mathfrak{g}$ is assured by the Ado theorem, the latter does not provide us with specific linearly-independent matrices $E_{p,q}$ (without fear of confusion, we use the same notation for formal words in (4.1) and for their representation). This motivates the work of this section, central to the entire paper. We consider the problem of finding a faithful representation of (4.1) as an exercise in *numerical algebra* and present an algorithm to this end.

Let us assume first that $\|N\|=1$, where $\|\cdot\|$ is the standard Euclidean matrix norm. Clearly, we may exclude the trivial case $N=O$, hence our assumption is merely a time-reparametrisation of the system (1.7): Once we find a representation subject to $\|N\|=1$, we can generalize it immediately for any $N\neq O$, multiplying each $E_{p,q}$ in (4.1) by $\|N\|$.

Proposition 5 *The Hermitian matrix $I+\mathrm{i}N$ is positive semi-definite and singular.*

Proof The proposition follows at once from the observation that $\sigma(I+\mathrm{i}N)=1+\mathrm{i}\sigma(N)$, $\sigma(N)\subset\mathrm{i}\mathbb{R}$ and $\rho(N)=\|N\|=1$, where σ and ρ denote the spectrum and the spectral radius, respectively. □

We seek an $n\times n$ complex *upper-triangular* matrix R such that

$$R^*R = I+\mathrm{i}N,$$

and which is in a *standard form:* For each row k the first nonzero term (if any) is $R_{k,i_k}>0$, where the i_ks form a strictly monotone sequence, and all zero rows are at the bottom. Note that Proposition 5 implies that the bottom row of R is necessarily composed of zeros. (The term "standard form" is borrowed from the usual terminology of QR factorization (Golub & van Loan 1989).) Precise details how to compute R are deferred to the next subsection.

We now remove the bottom row of R and denote the new matrix by $\tilde{R}$. If R has more rows of zeros, we excise them as well but in the sequel we treat merely the generic case when $\tilde{R}$ is $(n-1)\times n$ – an extension to the general case is trivial. Let

$$B = \operatorname{Re}\tilde{R}, \qquad C = \operatorname{Im}\tilde{R}.$$

Moreover, we set

$$A = \begin{bmatrix} B \\ C \end{bmatrix}$$

and denote its columns by $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n \in \mathbb{R}^{2n-2}$.

Proposition 6 *The following is true:*

$$\begin{aligned} B^\top B + C^\top C &= I, & (4.3)\\ B^\top C - C^\top B &= N, & (4.4)\\ \boldsymbol{a}_p^\top \boldsymbol{a}_q &= \delta_{p,q}, \qquad p,q = 1,2,\ldots,n, & (4.5)\\ \boldsymbol{a}_p^\top J \boldsymbol{a}_q &= n_{p,q}, \qquad p,q = 1,2,\ldots,n, & (4.6) \end{aligned}$$

where $J = \begin{bmatrix} O & I \\ -I & O \end{bmatrix}$ *is the standard* $(2n-2)\times(2n-2)$ *symplectic matrix.*

Proof Recalling the definition of R,

$$\tilde{R}^*\tilde{R} = R^*R = I + \mathrm{i}N.$$

On the other hand, $\tilde{R} = B + \mathrm{i}C$ implies that

$$\tilde{R}^*\tilde{R} = (B^\top - \mathrm{i}C^\top)(B + \mathrm{i}C) = (B^\top B + C^\top C) + \mathrm{i}(B^\top C - C^\top B).$$

Taking real and imaginary parts proves (4.3) and (4.4), respectively.

To prove (4.5) and (4.6), we denote the columns of B and C by $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n \in \mathbb{R}^{n-1}$ and $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n \in \mathbb{R}^{n-1}$, respectively, and compute directly while exploiting (4.3) and (4.4):

$$\boldsymbol{a}_p^\top \boldsymbol{a}_q = \begin{bmatrix} \boldsymbol{b}_p^\top & \boldsymbol{c}_p^\top \end{bmatrix} \begin{bmatrix} \boldsymbol{b}_q \\ \boldsymbol{c}_q \end{bmatrix} = (\boldsymbol{b}_p^\top \boldsymbol{b}_q + \boldsymbol{c}_p^\top \boldsymbol{c}_q) = (B^\top B + C^\top C)_{p,q} = \delta_{p,q}$$

and

$$\boldsymbol{a}_p^\top J \boldsymbol{a}_q = \begin{bmatrix} \boldsymbol{b}_p^\top & \boldsymbol{c}_p^\top \end{bmatrix} \begin{bmatrix} O & I \\ -I & O \end{bmatrix} \begin{bmatrix} \boldsymbol{b}_q \\ \boldsymbol{c}_q \end{bmatrix} = \boldsymbol{b}_p^\top \boldsymbol{c}_q - \boldsymbol{c}_p^\top \boldsymbol{b}_q = (B^\top C - C^\top B)_{p,q} = n_{p,q}.$$

□

Theorem 7 *Let*

$$E_{p,q} = (\boldsymbol{a}_p \boldsymbol{a}_q^\top + \boldsymbol{a}_q \boldsymbol{a}_p^\top)J, \qquad (p,q) \in \mathbb{J}_n. \tag{4.7}$$

Then

(a) *The set $\{E_{p,q} : (p,q) \in \mathbb{J}_n\}$ obeys the commutation relations (4.2) and therefore is a representation of the Lie algebra $\mathfrak{g}$.*

(b) *The above representation is faithful and, moreover,*

$$\operatorname{tr} E_{p,q}^\top E_{r,s} = \begin{cases} \frac{1}{2}, & p = q = r = s, \\ 1, & (p,q) = (r,s),\ p < q, \\ 0, & (p,q) \neq (r,s), \end{cases} \qquad (p,q),(r,s) \in \mathbb{J}_n. \tag{4.8}$$

Proof We prove (a) by direct computation, repeatedly using (4.6) as necessary. Let $(p,q),(r,s) \in \mathbb{J}_n$. Then, using the skew-symmetry of N,

$$\begin{aligned} [E_{p,q}, E_{r,s}] &= \tfrac{1}{4}[(\boldsymbol{a}_p\boldsymbol{a}_q^\top + \boldsymbol{a}_q\boldsymbol{a}_p^\top)J, (\boldsymbol{a}_r\boldsymbol{a}_s^\top + \boldsymbol{a}_s\boldsymbol{a}_r^\top)J] \\ &= \tfrac{1}{4}\{(\boldsymbol{a}_q^\top J\boldsymbol{a}_r)\boldsymbol{a}_p\boldsymbol{a}_s^\top + (\boldsymbol{a}_p^\top J\boldsymbol{a}_r)\boldsymbol{a}_q\boldsymbol{a}_s^\top + (\boldsymbol{a}_q^\top J\boldsymbol{a}_s)\boldsymbol{a}_p\boldsymbol{a}_r^\top + (\boldsymbol{a}_p^\top J\boldsymbol{a}_s)\boldsymbol{a}_q\boldsymbol{a}_r^\top \\ &\quad - (\boldsymbol{a}_s^\top J\boldsymbol{a}_p)\boldsymbol{a}_r\boldsymbol{a}_q^\top - (\boldsymbol{a}_r^\top J\boldsymbol{a}_p)\boldsymbol{a}_s\boldsymbol{a}_q^\top - (\boldsymbol{a}_s^\top J\boldsymbol{a}_q)\boldsymbol{a}_r\boldsymbol{a}_p^\top - (\boldsymbol{a}_r^\top J\boldsymbol{a}_q)\boldsymbol{a}_s\boldsymbol{a}_p^\top\}J \\ &= \tfrac{1}{4}\{n_{p,r}(\boldsymbol{a}_q\boldsymbol{a}_s^\top + \boldsymbol{a}_s\boldsymbol{a}_q^\top) + n_{p,s}(\boldsymbol{a}_q\boldsymbol{a}_r^\top + \boldsymbol{a}_r\boldsymbol{a}_q^\top) + n_{q,r}(\boldsymbol{a}_p\boldsymbol{a}_s^\top + \boldsymbol{a}_s\boldsymbol{a}_p^\top) \\ &\quad + n_{q,s}(\boldsymbol{a}_p\boldsymbol{a}_r^\top + \boldsymbol{a}_r\boldsymbol{a}_p^\top)\}J \\ &= \tfrac{1}{2}(n_{p,r}E_{q,s} + n_{p,s}E_{q,r} + n_{q,r}E_{p,s} + n_{q,s}E_{p,r}). \end{aligned}$$

This confirms that the $E_{p,q}$s obey (4.2) and thereby proves (a).

To prove part (b) of the theorem we observe that $J \in \mathrm{O}(n)$ and use (4.5):

$$\begin{aligned} \operatorname{tr} E_{p,q}^\top E_{r,s} = \operatorname{tr} E_{p,q}E_{r,s}^\top &= \tfrac{1}{4}\operatorname{tr}[(\boldsymbol{a}_p\boldsymbol{a}_q^\top + \boldsymbol{a}_q\boldsymbol{a}_p^\top)JJ^\top(\boldsymbol{a}_r\boldsymbol{a}_s^\top + \boldsymbol{a}_s\boldsymbol{a}_r^\top)] \\ &= \tfrac{1}{4}\operatorname{tr}[(\boldsymbol{a}_p\boldsymbol{a}_q^\top + \boldsymbol{a}_q\boldsymbol{a}_p^\top)(\boldsymbol{a}_r\boldsymbol{a}_s^\top + \boldsymbol{a}_s\boldsymbol{a}_r^\top)] \\ &= \tfrac{1}{4}\operatorname{tr}[(\boldsymbol{a}_q^\top\boldsymbol{a}_r)\boldsymbol{a}_p\boldsymbol{a}_s^\top + (\boldsymbol{a}_q^\top\boldsymbol{a}_s)\boldsymbol{a}_p\boldsymbol{a}_r^\top + (\boldsymbol{a}_p^\top\boldsymbol{a}_r)\boldsymbol{a}_q\boldsymbol{a}_s^\top + (\boldsymbol{a}_p^\top\boldsymbol{a}_s)\boldsymbol{a}_q\boldsymbol{a}_r^\top] \\ &= \tfrac{1}{2}(\delta_{p,r}\delta_{q,s} + \delta_{p,s}\delta_{q,r}), \end{aligned}$$

where we have used the identity $\operatorname{tr}\boldsymbol{a}_k\boldsymbol{a}_l^\top = \boldsymbol{a}_k^\top\boldsymbol{a}_l$. This proves (4.8).

Since the $E_{p,q}$s form an orthogonal set (with respect to the Frobenius norm), they are in particular linearly independent. Therefore the representation is faithful. □

4.2 The computation of R

We have used just a single feature of the matrix R, namely that $R^*R = I + \mathrm{i}N$. (The fact that the bottom row is zero and, indeed, that all zero rows are at the bottom of the matrix, is helpful in deriving smaller matrices but not required for the construction of the faithful representation at the first place. Hence, at least in principle, we may abandon the requirement that R is in a standard form, reconciling ourselves to larger matrices.) One procedure that is assured to produce R in a standard form is the *QR factorization* QR of the Hermitian, positive semidefinite, singular matrix $(I + \mathrm{i}N)^{1/2}$, where $Q \in \mathrm{U}(n)$ and R is upper triangular and in a standard form. Once we abandon the requirement that R is upper-triangular, hence reconciling ourselves to somewhat larger matrices, an alternative is a *polar decomposition* $I + \mathrm{i}N = QR$, where $Q \in \mathrm{U}(n)$ and R is Hermitian and positive semidefinite. However, the most promising route is the *Cholesky factorization,* which factorizes a positive definite matrix into the product R^*R, where R is upper triangular with real, positive diagonal.

We note in passing that we can replace R at will by PR, where $P \in \mathrm{U}(n)$. (Indeed, it is trivial that the manifold of all $n \times n$ complex matrices R such that $R^*R = I + \mathrm{i}N$ is acted multiplicatively from the left by $\mathrm{U}(n)$.) Alternatively, we can bypass R altogether and construct a matrix A, with n columns and an even number of columns, such that

$$A^\top A = I, \qquad A^\top JA = N,$$

except that this is not necessarily the easiest route.

There exist perfectly good routines (e.g. LINPAKC's `CCHDC` and NAG's `f01bnc`) that produce the Cholesky factorization of a Hermitian positive-definite matrix, and they are based on well-known comprehensive theory (Wilkinson & Reinsch 1971). Moreover, the LINPAKC (a complex extension of LINPACK) routine works with positive semidefinite matrices. Yet, there is a problem with standard Cholesky factorization in our setting. As long as the underlying Hermitian matrix is positive definite, R is by definition upper-triangular with positive elements along the diagonal. If it is of rank p, $p \le n-1$ and its principal $p \times p$ minor is nonsingular, the algorithm produces a matrix R with $n-p$ zero rows at the bottom and, otherwise, the diagonal elements are positive. However, if the rank is $p \le n-1$ but the principal $p \times p$ minor is singular, a Cholesky factorization requires pivoting and it is no longer true that $R^*R = I + \mathrm{i}N$.

We wish the best of both worlds: *both* R in a standard form *and* a Cholesky-like algorithm, that produces R without any need to compute square roots. This is provided by the following straightforward extension of the Cholesky algorithm. We progress successively forming the kth row of R for $k = 1, 2, \ldots, p$, where $p = \mathrm{rank}\,(I + \mathrm{i}N) \le n - 1$.

Let $M = N$. Supposing that the first $k-1$ rows of R are available, we note that

$$\sum_{j=1}^{k} |r_{j,k}|^2 = \sum_{j=1}^{n} |r_{j,k}|^2 = 1.$$

Thus,

$$\sum_{j=1}^{k-1} |r_{j,k}|^2 \le 1. \tag{4.9}$$

If the inequality above is sharp, we let

$$r_{k,k} = \left(1 - \sum_{j=1}^{k-1} |r_{j,k}|^2\right)^{\frac{1}{2}} > 0$$

and set

$$r_{k,l} = \frac{1}{r_{k,k}} \left(\mathrm{i}m_{k,l} - \sum_{j=1}^{k-1} \bar{r}_{j,k} r_{j,l}\right), \qquad l = k+1, k+2, \ldots, n.$$

It is trivial to confirm that the kth row of R is consistent with $R^*R = I + \mathrm{i}M$.

If there is an equality in (4.9) we replace R with RP_k, where P_k is the *permutation matrix* that cycles the columns $k, k+1, \dots, n$ to $k+1, k+2, \dots, n, k$ and, simultaneously, replace M by $P_k^\top M P_k \in \mathfrak{so}(n)$. Form the new sum $\sum_{j=1}^{k-1} |r_{j,k}|^2$: if it is strictly less than one, continue as above, otherwise continue cycling. We can cycle at most $n-k$ times. If the sum equals one in each case, then necessarily $\operatorname{rank}(I + \mathrm{i}N) = k$ and we can pad the bottom $n-k$ rows of R with zeros.

The outcome of this procedure is an upper-triangular matrix R and a product $\hat{P}$ of permutation matrices, such that $R^* R = I + \mathrm{i}M = \hat{P}^\top (I + \mathrm{i}N)\hat{P}$. We finally replace R by $R\hat{P}^\top$, whereby $R^* R = I + \mathrm{i}N$, as required. Note that, unlike in standard Cholesky factorization, it is entirely possible that elements of R vanish along the diagonal, but that the matrix is in a standard form.

We mention in passing an obvious amendment to the above procedure, full *row pivoting* of R: instead of cycling the $k, k+1, \dots, n$ columns when $r_{k,k}$ is zero, we exchange columns in each step so that the new pivot element $r_{k,k}$ is the largest possible. For large dimensions this procedure has the virtue of better stability.

As an example of our algorithm, consider

$$N = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and note that $\|N\| = 1$. For $k=1$ we obtain $r_{1,1} = 1$, $r_{1,2} = \mathrm{i}$ and $r_{1,3} = r_{1,4} = 0$, but $k = 2$ results in an equality in (4.9). We thus let

$$M \leftarrow P_2^\top M P_2 = \begin{bmatrix} 0 & 0 & 0 & \mathrm{i} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\mathrm{i} & 0 & 0 & 0 \end{bmatrix}, \qquad R = \begin{bmatrix} 1 & 0 & 0 & \mathrm{i} \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \end{bmatrix}.$$

Now (4.9) is sharp and $r_{2,2} = 1$, $r_{2,3} = r_{2,4} = 0$. Likewise, for $k = 3$ we have $r_{3,3} = 1$, $r_{3,4} = 0$. Finally, for $k = 4$ we have a row of zeros. Since $\hat{P} = P_2$, we finally set

$$R \leftarrow R\hat{P} = \begin{bmatrix} 1 & \mathrm{i} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

in a standard form. It is trivial that, indeed, $R^* R = I + \mathrm{i}N$.

4.3 An example

Let

$$N = \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix},$$

where $a,b,c \in \mathbb{R}$, $a^2+b^2+c^2=1$. It is easy to check that $\|N\|=1$. Assuming for simplicity that $|a|<1$, the procedure of Subsection 4.2 yields

$$R = \begin{bmatrix} 1 & \mathrm{i}a & \mathrm{i}b \\ 0 & \sqrt{b^2+c^2} & \frac{-ab+\mathrm{i}c}{\sqrt{b^2+c^2}} \\ 0 & 0 & 0 \end{bmatrix}.$$

Once we excise the bottom row, we obtain

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{b^2+c^2} & -\frac{ab}{\sqrt{b^2+c^2}} \\ 0 & a & b \\ 0 & 0 & \frac{c}{\sqrt{b^2+c^2}} \end{bmatrix},$$

consequently, by (4.7),

$$E_{1,1} = \begin{bmatrix} 0&0&1&0 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{bmatrix},$$

$$E_{1,2} = \begin{bmatrix} -\frac{1}{2}a & 0 & 0 & \frac{1}{2}\sqrt{b^2+c^2} \\ 0 & 0 & \frac{1}{2}\sqrt{b^2+c^2} & 0 \\ 0 & 0 & \frac{1}{2}a & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$E_{1,3} = \begin{bmatrix} -\frac{1}{2}b & -\frac{c}{2\sqrt{b^2+c^2}} & 0 & -\frac{ab}{2\sqrt{b^2+c^2}} \\ 0 & 0 & -\frac{ab}{2\sqrt{b^2+c^2}} & 0 \\ 0 & 0 & \frac{1}{2}b & 0 \\ 0 & 0 & \frac{c}{2\sqrt{b^2+c^2}} & 0 \end{bmatrix},$$

$$E_{2,2} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -a\sqrt{b^2+c^2} & 0 & 0 & b^2+c^2 \\ -a^2 & 0 & 0 & a\sqrt{b^2+c^2} \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$E_{2,3} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ \frac{b(a^2-b^2-c^2)}{2\sqrt{b^2+c^2}} & -\frac{1}{2}c & 0 & -ab \\ -ab & -\frac{ac}{2\sqrt{b^2+c^2}} & 0 & -\frac{b(a^2-b^2-c^2)}{2\sqrt{b^2+c^2}} \\ -\frac{ac}{2\sqrt{b^2+c^2}} & 0 & 0 & \frac{1}{2}c \end{bmatrix},$$

$$E_{3,3} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ \frac{ab^2}{\sqrt{b^2+c^2}} & \frac{abc}{b^2+c^2} & 0 & \frac{a^2b^2}{b^2+c^2} \\ -b^2 & -\frac{bc}{\sqrt{b^2+c^2}} & 0 & -\frac{ab^2}{\sqrt{b^2+c^2}} \\ -\frac{bc}{\sqrt{b^2+c^2}} & -\frac{c^2}{b^2+c^2} & 0 & -\frac{abc}{b^2+c^2} \end{bmatrix}.$$

The above basis does not share the symmetry implicit in reversing the order of rows and columns of N. This is hardly an impediment, except perhaps on æstethic grounds, but we note as a matter of interest that, at least for $n = 3$, we can single out a ‘symmetric’ basis,

$$E_{1,1} = \begin{bmatrix} 0 & \frac{1}{2}(a+c) & \frac{1}{2}(a-c) & b \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$E_{1,2} = \begin{bmatrix} -\frac{1}{2}a & -\frac{1}{2}b & \frac{1}{2}b & \frac{1}{2}c \\ 0 & \frac{1}{4}(a+c) & \frac{1}{4}(a-c) & \frac{1}{2}b \\ 0 & \frac{1}{4}(a+c) & \frac{1}{4}(a-c) & \frac{1}{2}b \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$E_{1,3} = \begin{bmatrix} -\frac{1}{2}b & \frac{1}{4}(a-c) & -\frac{1}{4}(a+c) & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4}(a+c) & \frac{1}{4}(a-c) & \frac{1}{2}b \end{bmatrix},$$

$$E_{2,2} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -a & -b & b & c \\ -a & -b & b & c \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$E_{2,3} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -\frac{1}{2}b & \frac{1}{4}(a-c) & -\frac{1}{4}(a+c) & 0 \\ -\frac{1}{2}b & \frac{1}{4}(a-c) & -\frac{1}{4}(a+c) & 0 \\ -\frac{1}{2}a & -\frac{1}{2}b & \frac{1}{2}b & \frac{1}{2}c \end{bmatrix}$$

$$E_{3,3} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -b & \frac{1}{2}(a-c) & -\frac{1}{2}(a+c) & 0 \end{bmatrix},$$

which is consistent with the commutation relations (4.2).

We mention in passing that it is possible – not by our algorithm, though – to derive unfaithful representations for sufficiently sparse matrices N. For example, for

$$N = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

we might take

$$E_{1,1} = \begin{bmatrix} -\frac{\sqrt{2}}{4} & -\frac{1}{2} - \frac{\sqrt{2}}{4} & 0 \\ \frac{1}{2} - \frac{\sqrt{2}}{4} & \frac{\sqrt{2}}{4} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad E_{1,2} = \begin{bmatrix} -\frac{\sqrt{2}}{4} & \frac{\sqrt{2}}{4} & 0 \\ \frac{\sqrt{2}}{4} & \frac{\sqrt{2}}{4} & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$E_{1,3} = \begin{bmatrix} 0 & 0 & -1 - \sqrt{2} \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \qquad E_{2,2} = \begin{bmatrix} \frac{\sqrt{2}}{4} & -\frac{1}{2} + \frac{\sqrt{2}}{4} & 0 \\ \frac{1}{2} + \frac{\sqrt{2}}{4} & -\frac{\sqrt{2}}{4} & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$E_{2,3} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 + \sqrt{2} \\ 0 & 0 & 0 \end{bmatrix}, \qquad E_{3,3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Clearly, $E_{3,3} = O$ and the representation is not faithful. This, as the very least, demonstrates the care that must be exercised in the construction of representations for general N, since absence of faithfulness might often be less obvious.

5 Concluding remarks

Having commenced from the equation (1.1), we proved that it is endowed with a Poisson structure and investigated invariants. Because of the relationship of Poisson structures with flows along orbits in the dual to the corresponding Lie algebra $\mathfrak{g}$, we have explored the issue of faithful representations and their numerical generation.

The problem of faithful representation of a Lie algebra has been treated in symbolic algebra and de Graff (1997) provided a very interesting algorithm. Yet, the very generality of this algorithm is its downfall in specific applications, since it produces representations of very large size. Ideally, we strive for minimal representations, in practice we are willing to compromise on 'tight' representations. In Section 4 we have presented an algorithm that generates faithful representations for commutation relations (4.2). Clearly, the algorithm is application-specific and it is unlikely that a similar approach can be applied to a wider (or a different) set of structure constants.

The construction of faithful Lie-algebraic representations has never, to the best of our knowledge, been considered in a numerical-algebraic setting. We firmly believe that, insofar as 'tight' representations are concerned, this is the right way forward. In the course of our research into more general isospectral flows and Poisson systems, we have assembled a significant collection of structure constants that obey the Jacobi identity. Only in few cases are their faithful representations known. A common thread running through all the cases when we have managed to identify a faithful representation is that matrices are low-rank (in our case they are of rank 2). It is premature to hypothesize whether this represents a valid general approach.

Acknowledgments

We would like to thank Brad Baxter (Birkbeck College), Matania Ben Artzi (Hebrew University), Roger Brockett (Harvard), John Butcher (University of Auckland), Willem de Graff (University of Linz), Adi Levin (Cadent), David Levin (University of Tel Aviv), Marta Mazzocco (University Manchester), Hans Munthe-Kaas (University of Bergen), Peter Olver (University of Minnesota), Gopal Prasad (University of Michigan), Tudor Ratiu (EPFL Lausanne) and Jeremy Schiff (Bar Ilan University). The work of AMB was supported by National Science Foundation, while the work of AI was performed in part as a visitor at University of Tel Aviv.

References

Bloch, A. M. & Iserles, A. (2003), On the optimality of double-bracket flows, Technical Report NA04, University of Cambridge. to appear in *Intl J. Maths & Mathematical Sc.*

Bloch, A. M., Brockett, R. W. & Crouch, P. E. (1997), 'Double bracket equations and geodesic flows on symmetric spaces', *Comm. Math. Phys.* **187**, 357–373.

Bloch, A. M., Brockett, R. W. & Ratiu, T. S. (1992), 'Completely integrable gradient flows', *Comm. Math. Phys.* **147**, 57–74.

Brockett, R. W. (1991), 'Dynamical systems that sort lists, diagonalize matrices, and solve linear programming problems', *Linear Algebra Appl.* **146**, 79–91.

Chu, M. T. (1992), 'Matrix differential equations: a continuous realization process for linear algebra problems', *Nonlinear Anal.* **18**, 1125–1146.

Chu, M. T. (1994), A list of matrix flows with applications, *in* 'Hamiltonian and Gradient Flows, Algorithms and Control', number 3 *in* 'Fields Inst. Commun.', Amer. Math. Soc., Providence, RI, pp. 87–97.

Chu, M. T. & Driessel, K. R. (1991), 'Constructing symmetric nonnegative matrices with prescribed eigenvalues by differential equations', *SIAM J. Math. Anal.* **22**, 1372–1387.

de Graff, W. A. (1997), Constructing faithful matrix representations of Lie algebras, *in* 'Proceedings of the 1997 International Symposium on Symbolic and Algebraic Computation', (Kihei, HI), pp. 54–59.

Deift, P., Nanda, T. & Tomei, C. (1983), 'Ordinary differential equations and the symmetric eigenvalue problem', *SIAM J. Numer. Anal.* **20**, 1–22.

Engø, K. & Faltinsen, S. (2001), 'Numerical integration of Lie–Poisson systems while preserving coadjoint orbits and energy', *SIAM J. Numer. Anal.* **39**, 128–145.

Flaschka, H. (1974), 'The Toda lattice', *Phys. Rev. B* **9**, 1924–1925.

Golub, G. H. & van Loan, C. F. (1989), *Matrix Computations*, 2nd edn, The Johns Hopkins University Press, Baltimore.

Horn, R. A. & Johnson, C. R. (1991), *Topics in Matrix Analysis*, Cambridge University Press, Cambridge.

Lewis, D. & Simo, J. C. (1994), ‘Conserving algorithms for the dynamics of Hamiltonian systems on Lie groups’, *J. Nonlinear Sci.* **4**, 253–299.

Manakov, S. V. (1976), ‘Note on the integration of Euler’s equations of the dynamics of an n-dimensional rigid body’, *Funct. Anal. and its Appl.* **10**, 328–329.

Marsden, J. E. & Ratiu, T. S. (1994), *Introduction to Mechanics and Symmetry*, Springer-Verlag, New York.

Olver, P. J. (1995), *Equivalence, Invariants, and Symmetry*, Cambridge University Press, Cambridge.

Ratiu, T. (1980), ‘The motion of the free n-dimensional rigid body’, *Indiana U. Math. J.* **29**, 609–627.

Symes, W. W. (1981/82), ‘The QR algorithm and scattering for the finite nonperiodic Toda lattice’, *Phys. D* **4**, 275–280.

Toda, M. (1987), *Theory of Nonlinear Lattices*, 2nd edn, Springer-Verlag, Berlin.

Varadarajan, V. S. (1984), *Lie Groups, Lie Algebras, and Their Representations*, Springer-Verlag, Berlin.

Watkins, D. (1984), ‘Isospectral flows’, *SIAM Rev.* **26**, 379–391.

Wilkinson, J. H. & Reinsch, C., eds (1971), *Linear Algebra*, number II *in* ‘Handbook for Automatic Computation’, Springer Verlag, Berlin.