ADVANCES IN MATHEMATICS 38, 267–317 (1980)

Completely Integrable Systems, Euclidean Lie Algebras, and Curves

M. ADLER*

Massachusetts Institute of Technology, Cambridge, Massachusetts 02139; and University of Minnesota, Minneapolis, Minnesota 55455

AND

P. VAN MOERBEKE†

Brandeis University, Waltham, Massachusetts 02154, and University of Louvain, Louvain-la-neuve, 1348, Belgium

DEDICATED TO PROFESSOR LOUIS BOUCKAERT

1. INTRODUCTION

In this series of two papers, of which this is the first, we discuss in a systematic fashion the relationship between what is classically known as completely integrable Hamiltonian systems, and polynomials in the indeterminate h, h^{-1}, with coefficients in one of the simple Lie algebras. The reason for putting these Hamiltonian systems in a Euclidean Lie algebra setup is that these systems linearize naturally on Jacobeans of curves associated to these Lie algebras; in many cases the linearization occurs on Prym varieties of the Jacobeans, to be more precise. The linearization shall be discussed in paper II [23].

It had previously been realized that complete integrability is strongly related to either Lie algebra [1, 2] or algebraic curve theory [3, 4]. For instance Adler [5] shows that both the Korteweg–deVries equation and the Toda systems[1] can be viewed as Hamiltonian systems on the co-adjoint orbit of a solvable group with the Kostant–Kirillov orbit structure, and moreover, the complete integrability of these systems may be traced to a single abstract Lie algebra theorem. Moreover, van Moerbeke [6] shows that the periodic Toda systems may be linearized on the Jacobean of a curve related to this system in some reasonable fashion. It is thus natural to combine these two approaches and to see how they relate. In fact they relate quite intimately,

* Sponsored by the NSF under Grant MCS79-17385.
† Partially sponsored by the NSF under Grant MCS79-05576.

[1] B. Kostant also realized this fact for the Toda systems [8].

0001-8708/80/120267-51$05.00/0

and also in great generality; and this is the main message of these two papers. The large number of examples illustrates this message. We also point out that our methods have the severe shortcoming that even if given a completely integrable system (with rational integrals) they do not provide a systematic procedure for discovering what the underlying Lie algebra is.

Specifically, we shall study the generalized (in the sense of Lie algebra) periodic-type Toda systems associated to the simple Lie algebras, of which the basic example associated to $sl(n)$ is a Hamiltonian equation of the form

$$\dot{x}_k = \frac{\partial H}{\partial y_k}, \qquad \dot{y}_k = \frac{-\partial H}{\partial x_k}, \qquad k = 1,\ldots, n,$$

for

$$H(x, y) = \frac{1}{2}\sum_{i=1}^{n} y_i^2 + \sum_{i=1}^{n} e^{x_i - x_{i+1}}, \qquad x_{n+1} = x_1 .$$

We shall also make a Lie algebra study of the so-called m-agonal generalizations, both symmetric and non-symmetric [3], of these equations. Next we study the Neumann problem, i.e., harmonic oscillators constrained to lie on a sphere, and geodesic flow on an ellipsoid (a problem of Jacobi) and centrally forced motion on an ellipsoid, two problems of the last century. We also focus on the well-studied spinning top problem (Euler and Lagrange) and a specialization of it related to geodesic flow on conics, also studied during the last century.

The main thrust of the method is to associate with all the above Hamiltonian systems a Lax matrix differential equation which moreover contains a parameter, i.e., we associate with these systems an equation of the form

$$\dot{A} = [A, B] \equiv AB - BA, \tag{1.1}$$

where A, B are square matrices whose entries depend on the phase space variables, and are polynomials in the indeterminate h and h^{-1}. Equation (1.1) thus takes place in a Kac–Moody Lie algebra, and we use an abstract theorem used in [5] (to study the Korteweg–deVries equation) to yield the complete integrability of (1.1) and in addition to concoct a sequence of flows which commute with (1.1). We then form the curve in (h, z) space

$$X\colon \det(A - z) = Q(z, h) = 0, \tag{1.2}$$

whose coefficients are functions of the phase space. From (1.1), the curve X: $Q(z, h) = 0$ (of genus g) is time independent, i.e., its coefficients are integrals of the motion (1.1); we then linearize (1.1) and its associated flows on the Jacobean $\mathrm{Jac}(X)$ of X. This is done by associating uniquely to each A of the above form a fixed curve X ($A \in \mathscr{A}(X)$) and an element in $Y = S^g(X)$ or $S^{g+1}(X)$ which then projects, via abelian sums, into a subvariety J contained in Jac X. The curves (1.2) will, for all "physical systems," turn out to be hyperelliptic, i.e., expressible in the form $y^2 = P(z)$, so that the linearization statement indeed translates into solutions of our systems by quadrature; namely, the problem is to find "enough" v_i's which depend on the phase coordinates of our system so that the motion of our systems may be expressed in J in the form

$$\sum_{i=1}^{i=g \text{ or } g+1} \int^{v_i(t)} \frac{z^k\, dz}{(P(z))^{1/2}} = c_k t, \qquad 0 \leqslant k \leqslant g-1, \tag{1.3}$$

with constants c_k. The set $\sum v_i$ lies in Y.

The fact that algebras other than $sl(n)$ come into play, reflects itself in symmetries of the matrix A, which in turn is reflected in the existence of an involution (other than the hyperelliptic involution τ) on the curve $Q(z, h) = 0$, i.e., $\sigma: (z, h) \to (-z, h)$. These involutions τ and σ constrain our set of $\sum v_i = \mathscr{D}$, $\mathscr{D} \in Y$, to have certain Prym-like symmetries related to the involutions, i.e., symmetries of the form $\mathscr{D}^{\sigma\tau} \equiv \mathscr{D}$, where the $\equiv$ means an identity of abelian sums. The actual Lie algebra under discussion reveals itself not only at the level of the curve, but in the map $\mathscr{A}(X) \to Y$, i.e., two different algebras may give rise to the same curve X, but then not to the same Y; however, the projection of Y into Jac X, may be identical for the two different algebras and so relationships are often subtle.

In addition, since Eqs. (1.1) arise in Lie algebras, it is natural to ask how representation theory interacts with the linearization theory. Specifically two different representations of the Lie algebra, say $sl(n)$, may very well lead to different curves X given by (1.2), but they must of course be intimately related. It turns out that using the theory of curve correspondences, especially an important theorem of Castelnuovo [22], we find that two curves arising in such a fashion X, X' are such that their Jacobeans Jac(X), Jac(X') must contain (up to isogeny) a common abelian subvariety and it is on this piece that in fact the linearization of our systems occurs. In short the linearization theory is representation independent precisely in the above sense.

The plan of paper I is as follows. In the second section we display explicitly the classical examples to be discussed so as to get the reader acquainted with the breadth of examples covered by our method. The third section contains the abstract Lie algebra framework, which will apply to the examples of Section 2. The fourth section explains how the examples of Section 2 are "poured" into our abstract container. The last section then discusses how geodesic flow on an ellipsoid and related systems fit very naturally into the co-adjoint orbit method for the groups $U(n)$, $U(n) ⓢ S$, $GL(n)$, $U(n) ⓢ u(n)$, where in fact their complete integrability can be easily checked. However, this approach does not indicate at all the way to the linearization. Paper II contains all of the machinery of the curve theory, but it is crucial to the conceptual understanding of the reader that the two papers be read as one.

The results containing the geodesic flow and some related systems were originally to appear as a joint paper of the first author with J. Moser, but different directions of approach and other considerations terminated that plan. In Section 4 we point out how Moser's [7] matrices derive naturally from ours as a limiting case, which indeed reflects the course of events. Kostant informed us of his results [8] with regard to the integrability of the Toda systems. Also, Ratiu [25] embarked upon a study of some aspects along his own directions. In addition, I. Frenkel, A. Reiman and M. Semenov-Tian-Shansky [31, 32] informed us by mail, after having seen a reference to our study in [5], that they had also undertaken such a study. The results in these papers have been the subjects of lectures throughout the period 1978–1979 at the University of California, Berkeley (3/79), Brandeis University (3/79), and Northwestern University (12/78, 6/79), Yale University (9/78), The University of Louvain (12/78), and M.I.T. (11/79), in which most of the important theorems of this paper have been explicitly displayed, and hence these results have been announced throughout the above period. We wish to thank J. Wolf, C. Moore, J. Roberts, R. McGehee, B. Kostant, T. Matsusaka, and S. Sternberg for their kind help, and T. Ratiu for assistance at the beginning of this project. Also, there is overlap of some of the results of Section 5 and the last section of the recent preprint of Sternberg and Jacob [24]. We also wish to thank Mary Birnbaum for her excellent drawings.

2. Classical Examples to be Discussed

In this section we give explicitly the classical Hamiltonian systems discussed in this paper. They are (a) the periodic Toda systems and their Lie algebra generalization; (b) the Euler–Arnold spinning top equations and the Lagrange top; (c) the Neumann problem, geodesic motion on an ellipsoid, and centrally forced motion on an ellipsoid, and a special case of the Euler–Arnold spinning top corresponding to geodesic motion on a conic. The periodic Toda systems were studied by van Moerbeke [6], while Kostant [8] studied the Lie algebra generalizations of the nonperiodic Toda systems. The spinning top problem has been studied by Arnold [1], Mischenko and Fomenko [9], and Dikii [10], and, as especially related to our work, by Manakov [11], and implicitly by Dubrovin *et al.* [4]. Case (c) is studied by Moser [12] in related work. That all these cases fit into our framework is part of the essential message of this paper. A notable example still missing is the top of Kowalevski [13].

The n-periodic Toda equations are best introduced by considering, with Toda, the Hamiltonian

$$H = H(x, y) = \frac{1}{2} \sum_{1}^{n} y_k^2 + \sum_{k=1}^{n} e^{x_k - x_{k+1}}, \qquad x_{n+i} = x_i, \tag{2.1}$$

which describes n particles on a ring with an exponential restoring force. The Hamilton equations $\dot{x}_k = H_{y_k}$, $\dot{y}_k = -H_{x_k}$ imply $\ddot{x}_k = e^{x_{k-1} - x_k} - e^{x_k - x_{k+1}}$, which, after the transformation of Flaschka [20], $a_k = \frac{1}{2} e^{(x_k - x_{k+1})/2}$, $b_k = -\frac{1}{2} y_k$, become

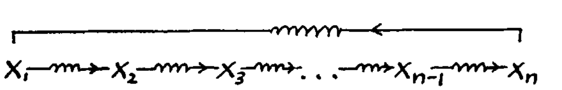
$$\dot{b}_k = 2(a_k^2 - a_{k-1}^2), \qquad \dot{a}_k = a_k(b_{k+1} - b_k), \qquad \text{all } k. \tag{2.2}$$

To describe the Lie algebra generalizations of (2.1), (2.2), of Bogoyavlensky [21], set $V_k = \sum_{i=1}^{k} \exp(x_i - x_{i+1})$; then the Hamiltonians going with the various simple Lie algebras are of the form $H = \frac{1}{2} \sum_{i=1}^{n} y_k^2 + V$, where V equals

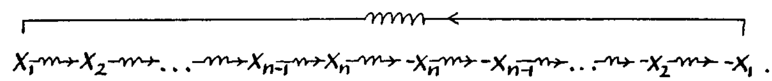
$$\begin{aligned}
V_{A_n} &= V_n + \exp(x_{n+1} - x_1), \qquad n \geqslant 2 \text{ (the } V \text{ of (2.1) with } n \to n+1), \\
V_{B_n} &= V_{n-1} + \exp(x_n) + \exp(-x_1 - x_2), \qquad n \geqslant 2, \\
V_{C_n} &= V_{n-1} + \tfrac{1}{2} \exp(2x_n) + \tfrac{1}{2} \exp(-2x_1), \qquad n \geqslant 3, \\
V_{D_n} &= V_{n-1} + \exp(x_{n-1} + x_n) + \exp(-x_1 - x_2), \qquad n \geqslant 4, \\
V_{G_2} &= \exp(x_1 - x_2) + \exp(-2x_1 + x_2 + x_3) + \exp(x_1 + x_2 - 2x_3);
\end{aligned} \tag{2.3}$$

$$\begin{aligned}
V_{E_7} &= V_5 + \exp(\tfrac{1}{2}(-x_1 + x_2 + \cdots + x_7 - x_8)) + \exp(-x_1 - x_2) \\
&\quad + \exp(-x_7 + x_8), \\
V_{E_8} &= V_6 + \exp(\tfrac{1}{2}(-x_1 + x_2 + \cdots + x_7 - x_8)) + \exp(-x_1 - x_2) \\
&\quad + \exp(x_7 + x_8), \\
V_{F_4} &= \exp(x_1 - x_2) + \exp(x_2 - x_3) + \exp(x_3) + \exp(\tfrac{1}{2}(-x_1 - x_2 - x_3 + x_4)) \\
&\quad + \exp(-x_1 - x_4).
\end{aligned} \tag{2.4}$$

We cast the systems of (2.3) into a form similar to that of (2.2) and give an analogous mechanical interpretation. Cases (2.4) shall be dealt with sketchily throughout this paper. For A_{n-1} we have the following mechanical diagram of n particles located at $x_1, \ldots, x_n$,



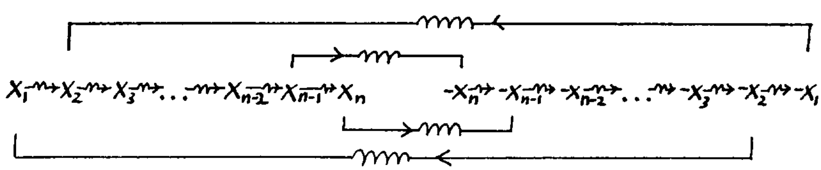
where the *identical* connections indicate exponential springs, and where the arrows indicate whether the connection is of type e^x or type e^{-x}. For V_{C_n} it is obvious from (2.3) that we may interpret the system as a subsystem of the system denoted by $V_{A_{2n-1}}$, where the particles are symmetrically arranged about the origin,

 (2.5)

For this system we again have Eq. (2.2), with $n \to 2n$,

$$
\begin{aligned}
&a_1, \dots, a_{2n}, \; b_1, \dots, b_{2n}, \text{ such that} \\
&a_k + a_{2n-k} = 0, \qquad k = 1, \dots, n-1, \\
&b_k + b_{2n+1-k} = 0, \qquad k = 1, \dots, n, \\
&a_k = \tfrac{1}{2} e^{(x_k - x_{k+1})/2}, \qquad k = 1, \dots, n-1, \; a_n = \tfrac{1}{2} e^{x_n}, \; a_{2n} = \tfrac{1}{2} e^{-x_1}, \\
&b_k = -\tfrac{1}{2} y_k, \qquad k = 1, \dots, n.
\end{aligned}
\tag{2.6}
$$

In a similar fashion the following diagram goes with D_n:

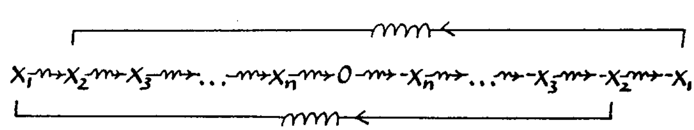


Note that the particles at $-x_n$, x_n are not connected, and this is clearly not a subsystem of the $V_{A_{2n-1}}$ case. The equations of type (2.2) going with V_{D_n} again require the introduction of

$$
\begin{aligned}
&a_1, \dots, a_{2n-1}, \; a_{1.2n-1}, \; a_{n-1.n+1}, \; b_1, \dots, b_{2n}, \text{ where} \\
&a_i + a_{2n-i} = b_i + b_{2n+1-i} = 0, \qquad i = 1, \dots, n, \\
&a_i = \tfrac{1}{2} e^{(x_i - x_{i+1})/2}, \qquad i = 1, \dots, n-1, \; a_{n-1,n+1} = \tfrac{1}{2} e^{(x_{n-1}+x_n)/2} \\
&\qquad a_{1,2n-1} = \tfrac{1}{2} e^{-(x_1+x_2)/2}.
\end{aligned}
\tag{2.7}
$$

The differential equations satisfied by $a_1, \dots, a_{n-1}$, $b_3, \dots, b_{n-2}$ are as in (2.2), and we have that, $b_1 = 2a_1^2 - 2a_{1,2n-1}^2$, $b_2 = 2a_2^2 - 2a_1^2 - 2a_{1,2n-1}^2$, $b_{n-1} = 2a_{n-1}^2 - 2a_{n-2}^2 + 2a_{n-1,n+1}^2$, $b_n = -2a_{n-1}^2 + 2a_{n-1,n+1}^2$.

For V_{B_n} we have the mechanical diagram



so we now have a particle sitting at the origin, $x_{n+1} = 0$. We consider the variables

$$
\begin{aligned}
&a_1, \ldots, a_{2n}, a_{1,2n}, b_1, \ldots, b_{2n+1}, \text{ where} \\
&a_i + a_{2n+1-i} = 0, \qquad i = 1, \ldots, n, \\
&b_i + b_{2n+2-i} = 0, \qquad i = 1, \ldots, n+1, \\
&a_i = \tfrac{1}{2} e^{(x_i - x_{i+1})/2}, \qquad i = 1, \ldots, n, \ x_{n+1} = 0, \ a_{1,2n} = \tfrac{1}{2} e^{-(x_1+x_2)/2}, \\
&b_i = -\tfrac{1}{2} y_i, \qquad i = 1, \ldots, n,
\end{aligned}
\tag{2.8}
$$

and the differential equations governing the motion in these variables are (2.2) except that $\dot{b}_1 = 2a_1^2 - 2a_{1,2n}^2$, $\dot{b}_2 = 2a_2^2 - 2a_1^2 - 2a_{1,2n}^2$, $\dot{b}_{n-1} = 2a_{n-1}^2 - 2a_{n-2}^2 + 2a_{n-1,n+1}^2$, $\dot{b}_n = -2a_{n-1}^2 + 2a_{n-1,n+1}^2$.

For the Hamiltonian associated with V_{G_2} in (2.3), one has three body forces in this three-particle problem, and if in the Hamiltonian equations $\dot{x}_i = H_{y_i}$, $\dot{y}_i = -H_{x_i}$, $i = 1, 2, 3$, $H = \frac{1}{2}\sum_{i=1}^{3} y_i^2 + V_{G_2}$, one sets

$$
\begin{aligned}
b_1 &= (-y_1 + y_3)/2, & b_2 &= (y_1 - y_2)/2, \\
a_1 &= \frac{1}{2} e^{(x_1 - x_2)/2}, & a_2 &= \frac{3^{1/2}}{2} e^{(-2x_1 + x_2 + x_3)/2}, \\
a_3 &= \frac{3^{1/2}}{2} e^{(x_1 + x_2 - 2x_3)/2},
\end{aligned}
\tag{2.9}
$$

then in these variables, the differential equations become

$$
\begin{aligned}
\dot{b}_1 &= 2(a_1^2 - a_2^2 + a_3^2), & \dot{a}_1 &= a_1 b_2, \\
\dot{b}_2 &= 2(a_2^2 - 2a_1^2), & \dot{a}_2 &= a_2(b_1 - b_2), \\
& & \dot{a}_3 &= -a_3(2b_1 + b_2).
\end{aligned}
\tag{2.10}
$$

We shall see (Eq. (4.28)) that (2.10) is just a special case of the B_3 equations, and so the G_2 equations have two mechanical interpretations.

We now discuss the spinning top equations; see also Arnold [1]. We start with the equations of motion of a rigid body B about a fixed point. Let M be the angular momentum of the body with respect to the body coordinates. Let γ be the (Eulerian angles) coordinates of the z-axis (fixed in space) with respect to the body. If the angular velocity of the body is $\Omega = (p, q, r)$, then the *total* derivative with respect to time is given by

$$
\frac{d}{dt}(\cdot)\Big|_{\text{total}} = \frac{d}{dt}(\cdot) + \Omega \times (\cdot).
$$

The torque exerted on the body by the (vertical) force of gravity is $l \times$ (gravity force), where $l =$ center of gravity in the body coordinates, and the $\mu g \gamma =$ downward gravity force, μ being the mass of the top. Thus the rotational version of Newton's law,

$$\frac{d}{dt}(\text{angular momentum})\Big|_{\text{total}} = \text{torque}, \quad \text{and} \quad \frac{d}{dt}(z\text{-axis})\Big|_{\text{total}} = 0,$$

yield the Poisson–Euler equations

$$\frac{dM}{dt} - M \times \Omega = \mu g \gamma \times l, \qquad \frac{d\gamma}{dt} - \gamma \times \Omega = 0. \tag{2.11}$$

If the body frame is principal and I_1, I_2, I_3 denote the principal moments of inertia, then we have $M = (I_1 p, I_2 q, I_3 r)$. As is well known, if one identifies vectors in R^3 with skew-symmetric matrices by the rule

$$a = (a_1, a_2, a_3), \qquad A = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}, \tag{2.12}$$

then $a \times b \mapsto [A, B] = AB - BA$.

Using this isomorphism, (2.12), we write (2.11) as

$$\frac{dM}{dt} = [M, \Omega] + \mu g[\gamma, l], \qquad \frac{d\gamma}{dt} = [\gamma, \Omega], \tag{2.13}$$

which may be regarded as the Lie algebra version of (2.11). In the absence of gravity we have the quadratic differential equation,

$$\frac{dM}{dt} = [M, \Omega], \tag{2.14}$$

where $[M, \beta] = [\Omega, \alpha]$, $[\alpha, \beta] = 0$, $\alpha = \beta^2$, $\beta = \frac{1}{2}(I_1 + I_2 + I_3)I - \operatorname{diag}(I_1, I_2, I_3)$. Equation (2.14) is the Euler–Arnold spinning top equation when $\alpha = \beta^2$; we shall also study (2.14) in general. The Lagrange top corresponds to the case where $I_1 = I_2$, and where the center of gravity and fixed point belong to the principal axis of inertia. Let z_0 be their respective distance and let $l = (0, 0, z_0)$. In that case we adjoin to (2.13) for future reference the relation

$$[M, \hat{\beta}] = [\Omega, \hat{\alpha}], \qquad \hat{\beta} = \mu g l, \qquad \hat{\alpha} = I_1 \hat{\beta}, \qquad \text{hence} \quad [\hat{\alpha}, \hat{\beta}] = 0, \tag{2.15}$$

as we adjoined a similar relation to (2.14); and so we think of (2.13) and (2.15) as the equations of the Lagrange top.

We now discuss geodesic motion on an ellipsoid, for which we refer the reader to Moser [12]. Let $\alpha = \operatorname{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n)$, $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_n$, and we give the ellipsoid[2] by $\langle \alpha^{-1}x, x\rangle = 1$, $x \in R^n$. We introduce the notation

$$Q_z(x, y) = \langle (z - \alpha)^{-1}x, y\rangle, \qquad Q_z(x) = Q_z(x, x),$$

and so the equation of our ellipsoid is

$$Q_0(x) + 1 = 0.$$

Its family of confocal quadrics is given by

$$Q_z(x) + 1 = 0. \tag{2.16}$$

We wish to determine those lines in R^n with distinguished point y and direction x, in other words lines of the form $y + sx$, $x \neq 0$, which are tangent to the surface given by (2.16), i.e., those y, x for which the equations

$$Q_z(y + sx) + 1 = 0, \qquad Q_z(y + sx, x) = 0$$

hold for some s; then $s = -Q_z(y, x)\, Q_z^{-1}(x)$. Eliminating s from the first expression, we find

$$\Phi_z(y, x) \equiv Q_z(x) + Q_z(x)\, Q_z(y) - Q_z^2(x, y) = 0 \tag{2.17}$$

as the equation for tangency to the surface (2.16). This may be thought of as an equation for the tangent bundle of $Q_z(x) + 1 = 0$.

We consider the Hamiltonian system

$$\dot{x} = \partial_y \Phi_z, \qquad \dot{y} = -\partial_x \Phi_z$$

restricted to the energy surface $\Phi_z = 0$. Remembering the expression for s above it is easily verified that $d^2(y + sx)/dt^2$ is proportional to ∇Q_z at the point $y + sx$, i.e., this differential equation governs the motion of the tangents of the hyperquadric $Q_z(x) + 1 = 0$ along geodesics (the parametrization, however, is not of arc length.) We obtain the geodesic flow by just following the motion of the point of tangency, $y + sx$, which amounts to reducing the system by the integral $|x|^2$. If we set $z = 0$ we obtain geodesic flow on the ellipsoid $Q_0(x) + 1 = 0$. If instead of (2.17) we consider

$$\Psi_z(y, x) \equiv -2Q_z(x, y) + (Q_z(x)\, Q_z(y) - Q_z^2(x, y)), \tag{2.18}$$

[2] $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$.

this Hamiltonian leads to the motion on a hyperquadric under the influence of a central force. This is most easily seen by rotating the point (x, y) by $\pi/4$ so that $\Psi_z \to \Phi_z + Q_z(y)$, which amounts to the sum of a kinetic and a potential energy $Q_z(y)$.

It is natural to view z as a parameter and make the developments

$$\Phi_z(y, x) = \sum_{v=1}^{n} \frac{F_v}{z - \alpha_v}, \qquad F_v(y, x) = x_v^2 + H_v\,,$$

$$H_v(y, x) = \sum_{j \neq v} \frac{(x_v y_j - x_j y_v)^2}{(\alpha_v - \alpha_j)}, \tag{2.19}$$

$$\Psi_z(y, x) = \sum_{v=1}^{n} \frac{G_v}{z - \alpha_v}, \qquad G_v(y, x) = -2x_v y_v + H_v\,, \tag{2.20}$$

and we also define

$$\Pi_z(y, x) = \sum_{v=1}^{n} \frac{H_v}{z - \alpha_v}, \tag{2.21}$$

and thus geodesic motion on the ellipsoid $Q_0(x) + 1 = 0$ is given by the Hamiltonian $\Phi_0 = -\sum \alpha_v^{-1} F_v(y, x)$ on $\Phi_0(y, x) = 0$. It is also natural from (2.19) to consider for good functions f Hamiltonians of the form

$$F_\beta = \tfrac{1}{2} \sum \beta_v F_v(y, x) \qquad \text{with} \quad \beta_v = f(\alpha_v). \tag{2.22}$$

A particular simple instance is

$$F_\alpha = \tfrac{1}{2} \sum \alpha_v F_v(x, y) = \tfrac{1}{2}\langle \alpha x, x\rangle + \tfrac{1}{2}(| x |^2 \, | y |^2 - \langle x, y\rangle^2). \tag{2.23}$$

Then Hamilton's equations $\dot{x} = \partial F_\alpha/\partial y$, $\dot{y} = -\partial F_\alpha/\partial x$, lead to

$$\ddot{x}_v = -\alpha_v x_v + \lambda x_v\,, \qquad v = 1,..., n, \quad \lambda = \langle \alpha x, x\rangle - | \dot{x} |^2. \tag{2.24}$$

These are the Neumann equations of forced harmonic motion constrained to lie on the sphere $| x | = 1$, as a result of the constraining force λx; observe $| x |$ is an integral of (2.24). Note that $(I - P_x) y \equiv y - (\langle x, y\rangle / | x |^2) x = \dot{x} / | x |^2$, which suggests reducing the above Hamiltonian system by $| x |^2$. In [12] it is proven by a geometrical argument that the F_v's (and H_v's) form an involutive system of integrals, and so the Neumann and ellipsoid systems belong to the *same* completely integrable family.

From (2.21), it is also natural to study systems of the form

$$H_\beta = \tfrac{1}{2} \sum \beta_v H_v\,, \tag{2.25}$$

where $\beta_v = (\alpha_v)^{1/2} \equiv J_v$, so that upon introducing the $n \times n$ skew-symmetric matrices

$$\Gamma_{xy} = x \otimes y - y \otimes x, \qquad (x \otimes y)_{ij} = x_i y_j \text{, etc.,} \qquad \Gamma_{ij} = (\Gamma_{xy})_{ij}(J_i + J_j)^{-1},$$

we find $H_\beta = H = -\frac{1}{4}\operatorname{tr}(\Gamma_{xy} \cdot \Gamma)$. Then the differential equations $\dot{x} = \partial H/\partial y$, $\dot{y} = -\partial H/\partial x$, are equivalent to the special case of the Arnold–Euler equations [1]

$$\dot{\Gamma}_{xy} = [\Gamma_{xy}, \Gamma], \tag{2.26}$$

and $-2H$ is the first Mischenko integral, J_1 [10]. For $n = 3$, it is easy to see that (2.26) amounts to the classical equation of Euler for the spinning top problem [1]. In fact, the above observation lead us to the tie between the systems discussed in this section and isospectral deformations of a particular nature. This will be discussed in the next section. We remark that the F_v, $v = 1,..., n$, were first discovered by K. Uhlenbeck.

3. Euclidean Lie Algebras and Integrability

In this section we discuss a unified group theoretic framework in view of the examples of Section 2. The next section will be devoted to fitting these examples into that framework. It has the disadvantage of not directly expressing all the geometric data like spatial symmetries of the system in the most immediate fashion, but it has the virtue of picking up all the various hidden symmetries, which are not obvious geometrically. It remains mysterious how to fit an example into this framework. The main tool will be the (modified) affine[3] Lie algebra

$$\mathscr{L} = \left\{ A = \sum_{-\infty < i \leqslant N} A_i h^i \,\middle|\, N \text{ arbitrary, finite, } A_i \in \mathscr{M} \right\}, \tag{3.1}$$

$\mathscr{M}$ a semi-simple Lie algebra, or the algebra of $n \times n$ matrices, and we view elements of $\mathscr{L}$ as Laurent series in the indeterminate h and h^{-1}. We shall show, and this is one of the basic results of this series of papers, that all the differential equations of Section 2 are expressible in the Lax (isospectral) form

$$\dot{A} = [A, B], \qquad A, B \text{ polynomials in } h, h^{-1}, \tag{3.2}$$

and that the complete integrability of these systems follows from a general theorem of Lie algebra relating Lie algebra decompositions to integrability

[3] For simplicity of exposition we work only with the affine Lie algebras, but we could work perfectly well with the Euclidean Lie algebras, i.e., the algebras which arise in the classification of automorphisms of finite type [29].

statements. In paper II it will be shown that these systems (3.2) are linearized on the Jacobeans of the curves

$$X\colon \det(A - zI) = Q(h, z) = 0, \tag{3.3}$$

and the linearization procedure is intimately related to the Lie algebra decompositions to be discussed. Moreover, we shall give an interpretation of the curves X which will prove useful in understanding the decompositions.

We first make some general observations concerning $\mathscr{L}$. The bracket in $\mathscr{L}$ is defined as follows:

$$\left[\sum A_i h^i, \sum B_j h^j\right] = \sum [A_i, B_j]\, h^{i+j}. \tag{3.4}$$

In addition we have the induced (ad-invariant,[4] nondegenerate, symmetric) Killing form on $\mathscr{L}$,

$$\left\langle \sum A_i h^i, \sum B_j h^j \right\rangle = \sum_{i+j=0} (A_i, B_j), \tag{3.5}$$

where $(\cdot, \cdot)$ is the Killing form on $\mathscr{M}$. This form is of course ad-invariant, nondegenerate, and symmetric. We will need to define the following forms, having the same three properties,

$$\langle A, B\rangle_k = \langle A, Bh^k\rangle = \sum_{i+j=-k} (A_i, B_j); \tag{3.6}$$

thus $\langle A, B\rangle_0 = \langle A, B\rangle$. The stated properties of (3.6) follow immediately from those of (3.5). To motivate calling (3.5) the Killing form of $\mathscr{L}$, observe from (3.4)

$$\mathrm{ad}_A \cdot \mathrm{ad}_B(C) = [A, [B, C]] = \sum_{i+j=0,\, l} [A_i[B_j, C_l]]\, h^{i+j+l} + \sum_{i+j\neq 0,\, l} \{\cdot\},$$

hence only the first term contributes to $\mathrm{tr}(\mathrm{ad}_A \cdot \mathrm{ad}_B)$, which is therefore

$$\sum_l \Big(\sum_{i+j=0} \mathrm{tr}\, \mathrm{ad}_{A_i} \cdot \mathrm{ad}_{B_j} \Big) = \sum_{l=0}^{\infty} \langle A, B\rangle = \Big(\sum_{l=0}^{\infty} 1 \Big) \cdot \langle A, B\rangle,$$

and we drop the irrelevant term $\sum_{l=0}^{\infty} 1$, i.e., we "renormalize" the trace.

Given a representation ϕ of $\mathscr{M}$ onto $R^n(C^n)$, it induces a representation of $\mathscr{L}$, also represented by ϕ, via

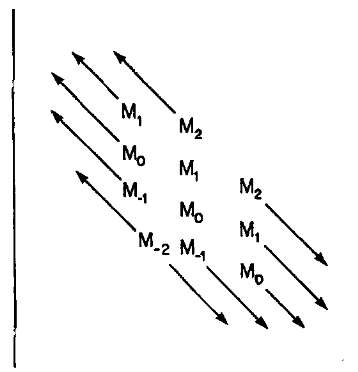
$$\phi : A = \sum A_i h^i \to \sum \phi(A_i)\, h^i \equiv \phi(A). \tag{3.7}$$

We shall think of $\phi(A)$ as a matrix Laurent series over $gl(n)$; however, we

[4] The term "ad-invariance" means $\langle [A, C], B\rangle + \langle C, [A, B]\rangle = 0$, i.e., the operator $[A, \cdot]$ is skew-symmetric with respect to $\langle \cdot, \cdot \rangle$.

also think of it in another way. Consider the map $\phi(A) \to \tilde{\phi}(A)$ of $\phi(A)$ into the infinite n periodic matrices given by

$$\phi(A) = \sum \phi(A_i)\, h^i \equiv \sum M_i h^i \to \tilde{\phi}(A) =$$



(We fix a basepoint.) It is easy to check that the map $\phi(A) \to \tilde{\phi}(A)$ is an algebra isomorphism, and in fact $\phi(A)$ is essentially a Fourier representation of the operator $\tilde{\phi}(A)$. For define the "character" $\chi_h(v) = (\ldots h^{-1}v, v, hv, h^2v, \ldots)^T \in R^\infty$, $v \in R^n$, then $\tilde{\phi}(A)\chi_h(v) = \chi_h(\phi(A)v)$ and so $\chi_h(v)$ is an eigenvector of $\tilde{\phi}(A)$ if and only if $\phi(A)v = zv$, where $z = z(h)$, $v = v(h)$, and so the characteristic equation for the infinite-dimensional operator $\tilde{\phi}(A)$ reduces to the representation dependent curve

$$X : \det(\phi(A) - z) \equiv Q_\phi(h, z) = 0.$$

Thus the infinite dimensionality of $\tilde{\phi}(A)$ is reflected in a characteristic equation containing an indeterminate, i.e., a curve. It is of course natural to consider X in its own right, without discussing $\tilde{\phi}(A)$.

If we stick to the classical Lie algebras $A_n, \ldots, D_n$, and $G_2 \subset B_3$, and if ϕ denotes their classical representation (or for general Lie algebras their adjoint representation), then up to an inessential factor, $(C, D) = \operatorname{tr}(\phi(C) \cdot \phi(D))$, $C, D \in \mathcal{M}$, and so for $A, B \in \mathcal{L}$, $\sum M_i h^i = \phi(A)$, if we define

$$\left\langle \left(\widetilde{\sum_i M_i h^i}\right) \right\rangle \equiv \operatorname{tr} M_0, \text{ then } \langle A, B \rangle = \langle \tilde{\phi}(A) \cdot \tilde{\phi}(B) \rangle, \tag{3.8}$$

where the $\cdot$ indicates matrix multiplication. To see this, note that, by the (algebra) isomorphism, $\langle \tilde{\phi}(A) \cdot \tilde{\phi}(B) \rangle = \langle (\sum_{i,j} \phi(A_i) \cdot \phi(B_j) h^{i+j})^\sim \rangle$,

$$\begin{aligned} \langle \tilde{\varphi}(A) \cdot \tilde{\varphi}(B) \rangle &= \operatorname{tr}\left(\sum_i \phi(A_i) \cdot \phi(B_{-i})\right) = \sum \operatorname{tr}(\phi(A_i) \cdot \phi(B_{-i})) \\ &= \sum (A_i, B_{-i}) \equiv \langle A, B \rangle. \end{aligned}$$

We also note, for future use, the obvious identity

$$\left\langle \sum M_i h^i \cdot \sum N_j h^j \right\rangle = \left\langle \sum N_j h^j \cdot \sum M_i h^i \right\rangle. \tag{3.9}$$

It is also natural to inquire how the characteristic curves X vary as one changes the faithful representation ϕ. To answer this question we need the following fact of Chevalley [14]: Given any Lie group G, and any representation g, faithful at the group level, any other representation faithful at the group level is contained, up to conjugation, in $\sum_{\alpha+\beta\geqslant 1} g^{\otimes\alpha} \otimes (g^*)^{\otimes\beta}$, i.e., in the direct sum of all combinations of the tensor product of g, and the contragredient of g (which over the reals corresponds to taking the transpose). This translates immediately by differentiation at the identity of G to a theorem about representations of the Lie algebra L of G, which come from representations faithful at the group level. Indeed, given a representation l of L, faithful at the group level (upon integration) then any other representation $\hat{l}$ of L, faithful at the group level, is contained, up to conjugation in $\sum_{\alpha+\beta\geqslant 1} l_{\alpha,\beta}$, with

$$l_{\alpha,\beta} \equiv \sum_{\delta=1}^{\alpha} \mathrm{id}^{\otimes(\alpha-\delta)} \otimes l \otimes \mathrm{id}^{\otimes(\beta+\delta-1)} + \sum_{\delta=0}^{\beta-1} (\mathrm{id})^{\otimes(\alpha+\delta)} \otimes l^* \otimes (\mathrm{id})^{\otimes(\beta-\delta-1)}.$$

Furthermore, each $l_{\alpha,\beta}$ splits up into irreducible pieces $l_{\alpha,\beta} = \sum \oplus l^{\gamma}_{\alpha,\beta}$, and so finally we have the identity $u\hat{l}u^{-1} = \sum' \oplus l^{\gamma}_{\alpha,\beta}$, where $\sum'$ indicates summation with repetition, over some set of (α, β, γ), and where u is a nonsingular matrix. The product $\prod'$ is defined similarly. We thus have

$$P_{\hat{l}}(z) \equiv \det(\hat{l} - z) = \prod{}' (\det(l^{\gamma}_{\alpha,\beta} - z) \equiv \prod{}' (P_{l^{\gamma}_{\alpha,\beta}}(z)), \tag{3.10}$$

and so we must analyze $\det(l^{\gamma}_{\alpha,\beta} - z)$, which occurs in the above products. Let us write

$$P_l(z) \equiv \det(l - z) = \prod_{i=1}^{n} (z - \lambda_i), \tag{3.11}$$

i.e., assume $lv_i = \lambda_i v_i$, $i = 1,\dots, n$, where v_i and λ_i are functions of L; then

$$l_{n,0}(v_{i_1} \otimes v_{i_2} \otimes \cdots \otimes v_{i_n}) = \left(\sum_{j=1}^{n} \lambda_{i_j}\right)(v_{i_1} \otimes v_{i_2} \otimes \cdots v_{i_n}),$$

etc., for $l_{\alpha,\beta}$ (remember we work over the reals), hence

$$\det(l_{\alpha,\beta} - z) = \prod \left(z - \sum_{s=1}^{k} \lambda_{i_s}\right) \equiv P_k, \qquad k = \alpha + \beta,\ (i_1, \dots, i_k) \in S_k,$$

S_k the symmetric group on k letters. We wish to find the irreducible factors of P_k, for which we need the following facts of Galois theory, which can be gleaned from Artin [15]:

(1) Given an arbitrary polynomial $P(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0$, over the base field K, with splitting field E; then if $\mathscr{H}$ is the associated Galois group of E/K, and if $\Delta \in E$, then the product $\Pi(z - \tau)$ taken over the orbit $\tau \in \mathcal{O} = \mathcal{O}_\Delta = \{h(\Delta) \mid h \in \mathscr{H}\}$ is an irreducible polynomial in K.

(2) (a) If $K = F(a_0, \ldots, a_{n-1})$, $P(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0 = \prod_{i=1}^n (z - \lambda_i)$, $E = K(\lambda_1, \ldots, \lambda_n)$, $\mathscr{H} = S_n$, which permutes $(\lambda_1, \ldots, \lambda_n)$.

(b) If $P(z) = ((z^2)^n + a_{n-1}(z^2) + \cdots + a_0) = \prod (z - \lambda_i) \cdot \prod (z + \lambda_i)$, $\mathscr{H} = S_n \otimes (Z_2)^n$, where the last factors act by taking $\lambda_i \to \pm\lambda_i$, $i = 1, \ldots, n$.

(c) If $P(z) = z^n + a_{n-2}z^{n-2} + a_{n-3}z^{n-3} + \cdots + a_0$, $K = F(a_0, \ldots, a_{n-2})$, $E = K(\lambda_1, \ldots, \lambda_{n-1})$, $\mathscr{H} = S_{n-1}$ on $(\lambda_1, \ldots, \lambda_{n-1})$.

Let us apply these facts first to the case of $\mathscr{M} = sl(n, R)$ in (3.11), and use for l the classical representation. $P_l(z)$ of (3.11), which depends on L, may be regarded as a generic polynomial of the form $P(z) = z^n + a_{n-2}z^{n-2} + \cdots + a_0$, and so we apply (2c). Now observe that P_k above has factors of the form $(z - \sum_{i=1}^t m_i\lambda_i)$, $t \leqslant k$, $\sum_i m_i = k$, m_i a positive integer, and using $\sum_{i=1}^n \lambda_i = 0$, we rewrite $\sum_{i=1}^t m_i\lambda_i$ to minimize t, allowing $m_i < 0$; concluding that the irreducible factors of P_k are of the form (using (1) and (2c))

$$\prod_{i_1, i_2, \ldots, i_t}{}' \left(z - \sum_{k=1}^t m_k\lambda_{i_k}\right) \equiv P^m(z), \qquad m = (m_1, \ldots, m_t), \tag{3.12}$$

Π' a restricted product if some $m_k = m_j$. Note that $P^m(z)$ is a well-defined polynomial over K, even though the λ_i's are not really well-defined global objects. As an example of the above considerations, for the case $n \geqslant 3$, we would not work with $\prod_{i_1 < \cdots < i_{n-1}}(z - \sum_{j=1}^{n-1} \lambda_{i_j})$, but rather with $\prod_{i=1}^n (z + \lambda_i)$, etc. We observe that as the representation ranges merely over the "Toda-type matrices" of Section 4, the coefficients in $P_l(z)$ of (3.11) may (except for the z^{n-1} term) be regarded as indeterminates. We thus have from (3.10), (3.11), and (3.12),

LEMMA 3.1. *If the curve corresponding to the classical representation is*

$$Q_\phi(h, z) = \sum_{i=1}^n (z - \lambda_i(h)) = 0, \tag{3.13}$$

then the curves $Q_\psi(h, z) = 0$, corresponding to the faithful (on the group level) representation ψ, factor into the generically irreducible curves

$$Q^m(z, h) = \prod{}' \left(z - \sum m_i\lambda_{i_k}\right) = 0. \tag{3.14}$$

Remark 1. Note that if $\psi = \phi^{\gamma}_{\alpha,\beta}$, $\alpha + \beta \geqslant n$, one can expect $Q_\psi(z, h)$ to contain factors where some of the $| m_i | > 1$.

Remark 2. We also note that the above considerations imply that the algebra of coefficients of $Q_\psi(h, z)$ is independent of the particular faithful representation. This is actually true if the representations are faithful only on the Lie algebra level, if by algebra we admit various analytic operations, like taking the square root.

Remark 3. If for the classical Lie algebras B_n, C_n, D_n, $G_2 \subset B_3$, we take for ϕ the classical representations (to be used in Section 4), then we have the generically irreducible curve (after perhaps dividing out by z)

$$Q_\phi(h, z) = \prod_{i=1}^{n} (z^2 - \lambda_i^2(h)) = 0, \tag{3.15}$$

and so from (1) and (2b), we find that $Q_\psi(z, h) = 0$, ψ a faithful representation, factors into irreducible curves of the form (see (3.14) for notation)

$$\prod_{\sigma} (Q^{\sigma(m)}(z, h)) = 0, \qquad \sigma = (\epsilon_1, \ldots, \epsilon_t), \quad \epsilon_i = \pm 1, \quad \sigma(m) = (\epsilon_1 m_1, \ldots, \epsilon_t m_t). \tag{3.16}$$

In paper II we shall show that also at the level of their Jacobeans (3.13)–(3.14), (3.15)–(3.16), respectively, are intimately related, as one would guess from the Torelli theorem.

We now give an appropriate (but by no means the most general) formulation of a general Lie algebra integrability theorem, discussed by Adler [5], and used crucially in a study of the symplectic structure and integrability of the Korteweg–deVries equation and Toda systems. This theorem is a generalization by Symes [16] of an integrability argument of Adler for the Toda systems. It is related to a theorem of Kostant [8].

THEOREM 3.1. *Let L be a Lie algebra paired with itself via a nondegenerate, ad-invariant, bilinear form $\langle\ ,\ \rangle$, L having the* vector space *direct sum decomposition $L = K + N$, for K, N, Lie subalgebras. We identify L with L^* via $\langle\ ,\ \rangle$. Then $L = K^\perp + N^\perp$, $\perp$ being taken with respect to $\langle\ ,\ \rangle$, and so identify $K^\perp \approx N^* \equiv$ the dual of N. Therefore $K^\perp$ inherits, via this identification, the co-adjoint orbit symplectic structure of Kostant and Kirillov* [1]. *Let $\Gamma \subset K^\perp$ be a manifold invariant under the above co-adjoint action of N on $K^\perp \approx N^*$, and let $\mathscr{A}(\Gamma)$ be the algebra of L^* (which is identified with L) functions, or functions at least defined on a neighborhood of Γ, which on Γ are invariant under the co-adjoint L* action *on L^* $(\approx L)$. (The two different co-adjoint actions are not to be confused.) Then the algebra $\mathscr{A}(\Gamma)$ forms a system of commuting integrals on Γ (when thought of as functions on Γ by restriction), and thus on the orbits of Γ themselves via the orbit symplectic structure. Moreover, if $H \in \mathscr{A}(\Gamma)$, the Hamil-*

tonian equations induced by H via the orbit symplectic structure have the isospectral form

$$\dot{a} = [a, b], \qquad b = P_K(\nabla H). \tag{3.17}$$

In the above, $\nabla H \in (L^*)^* = L$ is the gradient[5] *of H when viewed as a function of L^* ($\approx L$), while we have in general that P_K, P_N, $P_{K^\perp}$, $P_{N^\perp}$ are respectively the projections onto K, N, $K^\perp$, $N^\perp$ along N, K, $N^\perp$, $K^\perp$, respectively.*

Proof. We sketch a proof. First observe that if H is a function on L^* ($\approx L$), and if $\nabla_{K^\perp} H$, $\nabla_{N^\perp} H$, ∇H are the gradients of H in the $K^\perp$, $N^\perp$, L^* directions, respectively, then from calculus

$$\nabla_{K^\perp} H = P_N \nabla H, \qquad \nabla_{N^\perp} H = P_K \nabla H.$$

Also note that an L^* function being L invariant on Γ is equivalent[6] to

$$[\nabla H(a), a] = 0,\ a \in \Gamma, \text{ or equivalently,} \qquad [a, \nabla_{K^\perp} H] = -[a, P_K \nabla H], \tag{3.18}$$

by the formula above. If H and F are functions on N^*, then the Kostant–Kirillov Poisson bracket has the form

$$\{H, F\}(a) = \langle\langle a, [\nabla_{N^*} H, \nabla_{N^*} F]\rangle\rangle, \qquad a \in N^*,$$

where $\langle\langle\ ,\ \rangle\rangle$ is the natural pairing between N and N^*, and where $\nabla_{N^*} H \in N$ is the natural gradient of H defined by $dH(X) \equiv \langle\langle dX, \nabla_{N^*} H\rangle\rangle$; so for our case $K^\perp \approx N^*$ and $\langle\langle\ ,\ \rangle\rangle = \langle\ ,\ \rangle|_{K^\perp \times N}$; hence

$$\{H, F\}(a) = \langle a, [\nabla_{K^\perp} H, \nabla_{K^\perp} F]\rangle. \tag{3.19}$$

Suppose $H, F \in \mathscr{A}(\Gamma)$, (satisfying (3.18)); then

$$\begin{aligned}
\{H, F\} &= \langle a, [\nabla_{K^\perp} H, \nabla_{K^\perp} F]\rangle \\
&= \langle [a, \nabla_{K^\perp} H], \nabla_{K^\perp} F\rangle \qquad \text{(by the ad-invariance of } \langle\ ,\ \rangle) \\
&= -\langle [a, P_K \nabla H], \nabla_{K^\perp} F\rangle \text{ (by (3.18))} \\
&= -\langle a, [P_K \nabla H, \nabla_{K^\perp} F] \\
&= \langle a, [P_K \nabla H, P_K \nabla F]\rangle \qquad \text{by repeating the argument for } F \\
&= 0
\end{aligned}$$

since $a \in K^\perp$ and since K is a Lie algebra.

[5] One defines the gradient ∇H of a function H on a vector space V, $v \in V$, by $dH \equiv (\nabla H, dv)_V$, $\nabla H \in V^*$, $(,)_V$ the pairing between V, V^*.

[6] This is immediately inferred from the identity $(d/dt) H(\mathrm{Ad}_g(a))|_{t=0} = 0$, for $g = 1 + tc + \mathcal{O}(t^2)$, any $c \in L$, $a \in \Gamma$.

The Hamilton vector-field equals

$$X_H(F) = \{F, H\} = \langle [\nabla_{K^\perp} H, a], \nabla_{K^\perp} F\rangle,$$

from which it follows that

$$X_H(a) = P_{K^\perp}[\nabla_{K^\perp} H, a]. \tag{3.20}$$

Hence the corresponding Hamiltonian flow

$$\begin{aligned} \dot{a} &= P_{K^\perp}[\nabla_{K^\perp} H, a] && (\text{for } H \in \mathscr{A}(\Gamma)) \\ &= P_{K^\perp}[a, P_K \nabla H] && (\text{by } (3.18)) \\ &= [a, P_K \nabla H], && \text{as } [K^\perp, K] \subset K^\perp, \end{aligned}$$

which proves (3.17) and thus concludes the proof of the theorem.

For future reference, let $(\mathrm{Ad}^*)_N$ denote the co-adjoint action of the group of N on $N^* \approx K^\perp$ and Ad^* the L co-adjoint action; then since $(\mathrm{Ad}^* g)_N = P_{K^\perp} \circ \mathrm{Ad}^* g$,

$$\mathcal{O}_a = \{P_{K^\perp}(\mathrm{Ad}^* g(a)) \mid g \in G_N\}, \tag{3.21}$$

where G_N is the Lie group of N.

4. Examples—Decompositions

In this section we discuss the examples of Section 2 using the machinery of Section 3. This involves studying two different decompositions of $\mathscr{L}$ (Eq. (3.1)), which we shall explore in some detail. In addition, we discuss a decomposition of $L = \mathscr{L} \oplus \mathscr{L}$ which leads to the nonsymmetric generalized Toda matrix systems studied by Mumford and Moerbeke [3], while one of the decomposition cases above leads to the Toda-type systems of Section 2, and also to the generalized symmetric Toda systems studied by Mumford and Moerbeke [3].

Example 1. *Toda-Type Systems and Generalized Symmetrics*

We first need some well-known facts. If $\mathscr{M}$ is a semi-simple Lie algebra, it may be decomposed as follows using the Cartan decomposition.

$$\begin{aligned} &\mathscr{M} = \mathscr{h} \oplus \textstyle\sum \oplus e_\alpha, \quad \alpha \in \Delta, \quad \mathscr{h}^* \approx \mathscr{h} \text{ via } (\ ,\ ), \quad \text{the Killing form, where} \\ &\qquad [e_\alpha, e_{-\alpha}] = h_\alpha, \qquad [h, e_\alpha] = \alpha(h)\, e_\alpha, \quad h_\alpha, h \in \mathscr{h}, \quad \alpha(h) \equiv (\alpha, h), \\ &\qquad [e_\alpha, e_\beta] = 0, \qquad \alpha + \beta \neq 0, \quad \alpha + \beta \notin \Delta, \\ &\qquad [e_\alpha, e_\beta] = N_{\alpha,\beta} e_{\alpha+\beta}, \qquad \alpha + \beta \in \Delta, \quad N_{\alpha,\beta} = -N_{-\alpha,-\beta}, \\ &\qquad \mathscr{h} = \sum_{i=1}^{r} \oplus p_i, \qquad r = \operatorname{rank} \mathscr{M}, \quad (p_i, p_j) = \delta_{ij}. \end{aligned} \tag{4.1}$$

In fact, by a theorem of Chevalley, the $N_{\alpha,\beta}$'s can be taken $\pm(p+1)$, where p is the greatest integer i for which $\beta - i\alpha$ is a root. There exists a set Δ_0, of simple roots, which induce a leveling of $\mathscr{M}$ and Δ,

$$\mathscr{M} = \sum_{-(d-1)\leqslant i\leqslant (d-1)} \oplus \mathscr{M}_i\,,\ [\mathscr{M}_i\,, \mathscr{M}_j] \subset \mathscr{M}_{i+j}\,,\ \dim \mathscr{M}_{\pm(d-1)} = 1,$$

$$\mathscr{M}_0 = \mathscr{h},\ \mathscr{M}_{\pm 1} = \sum_{\pm\alpha\in\Delta_0} \oplus\, e_\alpha\,,\ \mathscr{M}_{\pm(s+1)} = [\mathscr{M}_{\pm s}\,, \mathscr{M}_{\pm 1}].$$

Define $|\alpha| = j$, if $e_\alpha \in \mathscr{M}_j$; if $M = \sum M_i\,,\ M_i \in \mathscr{M}_i$ define (4.2)

$$M \equiv M^- + M^0 + M^+,\ M^\pm = \sum_{i\gtrless 0} M_i\,, M^0 = M_0\,.$$

This decomposition generates an anti-involution on $\mathscr{M}$, the generalization of transpose, which satisfies

$$\begin{aligned} &\left(\sum a_\alpha e_\alpha + p\right)^T = \left(\sum a_\alpha e_{-\alpha} + p\right), \qquad p \in \mathscr{h}, \\ &(A^T, B^T) = (A, B), \qquad [A, B]^T = -[A^T, B^T]. \end{aligned} \tag{4.3}$$

We say an element A is symmetric if $A = A^T$, skew-symmetric if $A = -A^T$. Note that by (4.3)

$$\begin{aligned} &\text{symmetrics} = \left(p + \sum a_\alpha(e_\alpha + e_{-\alpha})\right),\ \text{skew-symmetrics} = \left(\sum a_\alpha(e_\alpha - e_{-\alpha})\right), \\ &(\text{symmetric, skew-symmetric}) = 0. \end{aligned} \tag{4.4}$$

Remembering the definition (3.1) of $\mathscr{L}$, we see that (4.2) induces a leveling of $\mathscr{L}$; namely, let

$$\mathscr{L}_s \equiv \sum_{\substack{k+jd=s\\ |k|\leqslant d-1}} \oplus\, (h^j\mathscr{M}_k), \quad \text{and} \quad \mathscr{A}_{jk} \equiv \sum_{j\leqslant i\leqslant k} \oplus\, \mathscr{L}_i\,. \tag{4.5}$$

From this it follows that

$$\begin{aligned} &\mathscr{L}_0 = \mathscr{M}_0\,,\ \mathscr{L}_1 = \mathscr{M}_1 + h\mathscr{M}_{-(d-1)}\,,\ \mathscr{L}_{-1} = \mathscr{M}_{-1} + h^{-1}\mathscr{M}_{(d-1)}\,, \\ &\mathscr{L} = \sum \mathscr{L}_i\,,\ \mathscr{L}_{\pm(k+1)} = [\mathscr{L}_{\pm k}\,, \mathscr{L}_{\pm 1}],\ [\mathscr{L}_m\,, \mathscr{L}_n] \subset \mathscr{L}_{m+n} \end{aligned} \tag{4.6}$$

and we define $f_\beta \equiv e_\alpha h^k \in \mathscr{L}_s\,,\ \beta \equiv \alpha_k\,,\ |\beta| = s$; therefore we may interpret $\mathscr{L}$ as a Lie algebra whose Cartan subalgebra is $\mathscr{M}_0$, but whose simple eigenvector space is not $\mathscr{M}_1$, but rather $\mathscr{L}_1 = \mathscr{M}_1 + h\mathscr{M}_{-(d-1)}$. We interpret the simple root set of $\mathscr{L}$ to be not Δ_0, but the extended root system, $\Delta_0 \cup \{-\gamma\}$, where γ is the maximal root vector, i.e., $\mathscr{M}_{(d-1)} = R \cdot e_\gamma$.

We now apply Theorem (3.1) to the case $L = \mathscr{L}$ above, $\langle\, , \rangle = \langle\, , \rangle_0$ of (3.5, 6), setting (see (4.3))

$$K = \left\{ \sum A_i h^i \,\middle|\, A_i + A_{-i}^T = 0,\ A_i \in \mathscr{M} \right\},$$

$$N = \sum_{i \leqslant 0} \mathscr{L}_i = \left\{ \sum_{i \leqslant 0} A_i h^i \,\middle|\, A_0 \subset \sum_{j \leqslant 0} \mathscr{M}_j,\ A_i \in \mathscr{M} \right\} = \mathscr{A}_{-\infty,0},$$

$$K^\perp = \left\{ \sum A_i h^i \,\middle|\, A_i - A_{-i}^T = 0,\ A_i \in \mathscr{M} \right\},$$

$$N^\perp = \sum_{i<0} \mathscr{L}_i = \left\{ \sum_{i<0} A_i h^i \,\middle|\, A_0 \subset \sum_{j<0} \mathscr{M}_j,\ A_i \in \mathscr{M} \right\} = \mathscr{A}_{-\infty,-1}, \tag{4.7}$$

and so (see (4.2))

$$P_{K^\perp}\left(\sum A_i h^i\right) = \sum_{i>0} (A_i h^i + A_i^T h^{-i}) + (A_0^+ + (A_0^+)^T + A_0^0),$$

$$P_K\left(\sum A_i h^i\right) = \sum_{i>0} (A_i h^i - A_i^T h^{-i}) + (A_0^+ - (A_0^+)^T).$$

Remark. For $\mathscr{M} = sl(n)$, thinking in terms of infinite n-periodic matrices, we have that K and $K^\perp$ are respectively the (infinite n-periodic) symmetric and skew-symmetric matrices, that N and $N^\perp$ are respectively the lower triangular and strictly lower triangular (infinite n-periodic) matrices; and this is the intuitive content of the decomposition. We shall elaborate in the examples.

For our invariant manifolds take

$$\Gamma = \Gamma_m = K^\perp \cap \mathscr{A}_{-m,m}, \tag{4.8}$$

and for elements of $\mathscr{A}(\Gamma)$ take

$$H(A) = \langle f(\phi(A))\, h^K \rangle \tag{4.9}$$

(see (3.8) for notation), where ϕ is a faithful representation of $\mathscr{M}$ and $f(x)$ is an analytic function. Hence $H(A)$ is a function of the coefficients of $\det(\phi(A) - z) = Q_\phi(h, z)$. It leads to the Lax equations $\dot{A} = X_H = [A, P_K \nabla H]$.

We now elaborate on this example and verify the implicit claims. If one defines $(\sum A_i h^i)^t = \sum A_{-i}^T h^i$, then one calls A symmetric if $A^t = A$, skew-symmetric if $A^t = -A^t$. Clearly K are the symmetrics and $K^\perp$ the skew-symmetrics. One checks from (4.3) that $[A, B]^t = [B^t, A^t]$, hence K is a Lie algebra. The "lower triangulars" $N = \sum_{i \leqslant 0} \mathscr{L}_i$ form a Lie algebra as follows from (4.6). That $K^\perp$ and $N^\perp$ are what they are claimed to be in (4.7) follows from the last statement of (4.4). Note that $\mathscr{L} = K^\perp + N^\perp$ because $\langle , \rangle$ is non-degenerate. The statements concerning $P_{K^\perp}$ and P_K are obvious.

We now study the group G_N whose Lie algebra is N, with a view to computing its orbits in $K \approx N^*$. Pick a faithful representation of $\mathscr{M}$ on $R^k(C^k)$, $\mathscr{M} \to m$, $\mathscr{L} \to l$, etc. We may, after conjugation, assume (see (4.2) $m_i^-(M^- \to m^-)$ are lower triangular matrices, and so, using the isomorphism of Section 3, we may finally visualize $N \to n \to \tilde{n}$, as infinite k-periodic lower triangular matrices. We shall suppress the first map and denote $N \to n \to \tilde{n}$ by $N \to \tilde{n}$. Taking exponentials of such elements, and finite products of such exponentials, we arrive at a representation of the group G_N. We now define the auxiliary (finite-dimensional) Lie algebra $N_s = \sum_{s \leqslant j \leqslant 0} \mathscr{L}_j = \mathscr{A}_{-s,0}$, which is a Lie algebra under the (truncation) rule (see (4.6))

$$\left(\sum a_\alpha f_\alpha, \sum b_{\alpha'} f_{\alpha'}\right) = \sum_{|\alpha|+|\alpha'| \geqslant -s} a_\alpha b_{\alpha'} [f_\alpha, f_{\alpha'}].$$

Under the representation $N_s \to \tilde{N}_s$, $f_\alpha \to \tilde{f}_\alpha$, we construct the corresponding finite-dimensional Lie group G_{N_s}. If $a_i = \sum a_\alpha^{(i)} \tilde{f}_\alpha \in \tilde{N}_s$, define the truncated product in "$\tilde{n}_s{}^j$": $(a_1 \circ a_2 \circ \cdots \circ a_j)|_s = \sum (\prod_{i=1}^j a^i_{\alpha(i)})(\tilde{f}_{\alpha(1)} \cdot \tilde{f}_{\alpha(2)} \cdots \tilde{f}_{\alpha(j)})$, the sum taken over $\sum |\alpha(i)| \geqslant -s$, where $\cdot$ denotes matrix multiplication, and such elements multiply by the (associative) rule

$$(a_1 \circ \cdots \circ a_j)|_s \times (b_1 \circ \cdots \circ b_k)|_s = (a_1 \circ \cdots \circ a_j \circ b_1 \circ \cdots \circ b_k)|_s .$$

Define

$$\exp(a) = (I + a + \tfrac{1}{2}(a \cdot a)_s + \cdots) \in \tilde{n}, \qquad a \in \tilde{N}_s ,$$

whose inverse is $\exp(-a)$. Forming finite products of such elements via the above rules, and closing up this set, we have constructed a Lie group, (or rather its representation) G_{N_s}, whose Lie algebra is clearly $\tilde{N}_s$. We may view G_{N_s} as injected into G_N via this truncation. Hence we have a filtration of N by N_s, G_N by G_{N_s}, and we may view N as the direct limit of N_s, etc., for G, i.e., $N_1 \subsetneq N_2 \subsetneq \cdots \subsetneq N$, $G_{N_1} \subsetneq G_{N_2} \subsetneq \cdots \subsetneq G_N$. We cannot form a group for $\mathscr{L}$ in this fashion, but must impose a norm on $\mathscr{L}$, which is not natural in this context.

We are now able to describe the co-adjoint orbit of G_N through the point $A \in \Gamma_m$. For any $B \in \mathscr{L}$ ($B \to \tilde{B}$ as $\mathscr{L} \to \tilde{\mathscr{L}}$), and $g \in G_N$, we define $\widetilde{\operatorname{Ad} g(B)} = g\tilde{B}g^{-1}$. This definition makes sense, for from the definition of G_N using exponentials it is immediate that $g\tilde{B}g^{-1}$ stays in $\tilde{\mathscr{L}}$, and so by the isomorphism it defines a unique element B in $\mathscr{L}$.

We now require our representation of $\mathscr{M} \to m$ to be either the ad-joint representation, or for the classical cases and $G_2 \subset B_3$, the classical representation; then as follows from (3.8) and (3.9)

$$\langle \operatorname{Ad} gB, C\rangle = \langle g\tilde{B}g^{-1}\tilde{C}\rangle = \langle \tilde{B}g^{-1}\tilde{C}g\rangle = \langle \tilde{B}, \operatorname{Ad} g^{-1}\tilde{C}\rangle,$$

and so we have the desired formula

$$\operatorname{Ad} g^* = \operatorname{Ad} g^{-1}. \tag{4.10}$$

By (3.22) and (4.10) we have for $A \in \Gamma_m = K^\perp \cap \mathscr{A}_{-m,m}$ that $\mathcal{O}_A = \{P_{K^\perp} \operatorname{Ad} g(A) \mid g \in G_N\}$. We elaborate on the above formula. From (4.7), $P_{K^\perp} \operatorname{Ad} g(A) = P_{K^\perp}(\operatorname{Ad} g(A))_+$, where $(B)_+$ (with $B \in \mathscr{L}$) is the projection of B onto $\sum_{i\geqslant 0} \mathscr{L}_i$ along $\sum_{i<0} \mathscr{L}_i$, and so only the "upper triangular" piece of $\operatorname{Ad} g(A)$ matters in the above. But since g and g^{-1} are "lower triangular" matrices (in $\bar{\mathscr{L}}$), $P_{K^\perp} \operatorname{Ad} g(A) = P_{K^\perp}(\operatorname{Ad} g(A_+))_+$. Remembering the injections $G_{N_K} \subset G_N$, we claim that G_N may be replaced by G_{N_m} in (3.21). This is because by the truncation formulas defining G_{N_m} as products of elements of the form $\exp(A) = I + (\tilde{A})_m + \frac{1}{2}(\tilde{A} \cdot \tilde{A})_m + \cdots$, only the G_{N_m} "part" of $g \in G$ above contributes to $(\operatorname{Ad} g(A_+))_+$, since $A_+ \in \mathscr{A}_{0,m}$, and this is true likewise for products of such elements. We conclude

$$\mathcal{O}_A = \{P_{K^\perp}(\operatorname{Ad} g(A_+))_+ \mid g \in G_{N_m}\}, \qquad A \in \Gamma_m. \tag{4.11}$$

We claim $\mathcal{O}_A$ is intrinsically defined in N. For observe that G_{N_m} has Lie algebra N_m, whose dual N_m^* may, through $\langle\ ,\ \rangle$, be identified with $K_m{}^\perp = K^\perp \cap \mathscr{A}_{-m,m}$; this is compatible with the identification of N^* with $K^\perp$.

Therefore, we can think of $\mathcal{O}_A$ as a co-adjoint orbit of the *finite* dimensional group G_{N_m}, and the Kostant–Kirillov structure induced by G_N on A is precisely the one induced by G_{N_m} on A. Upon differentiating (4.11), we have the tangent space (where we implicitly use our isomorphisms)

$$T\mathcal{O}_A(A) = \{P_{K^\perp}([A_+, B])_+ \mid B \in N_m\}, \tag{4.12}$$

and the symplectic structure is independent of the representation we have chosen to construct G_N and G_{N_K}, as is apparent from (4.12). Thus both $\mathcal{O}_A$ and its symplectic structure are intrinsically defined, as it should be. Note that those orbits $\mathcal{O}_A$ are manifolds, since the group G_{N_m} is finite dimensional.

We now show that (4.9) are elements of $\mathscr{A}(\Gamma)$; by Remark 2 of Section 3, it suffices to do that for the representation used in the construction of G_N. There are slight technical difficulties due to the setting which obscure the main point. We take $H(A) = \langle f(\tilde{A})\, h^k\rangle$ (thinking now of having fixed a representation and so $\tilde{A}$ is a polynomial matrix in h, h^{-1}), with $A \in \Gamma_m$, m arbitrary, and for simplicity we require f to be a polynomial.

Having constructed G_N, we could just as well have constructed G_{N^T}, going with the "upper triangular matrices" and so we have (viewing g as a formal series in h or h^{-1}) $\langle f(g\tilde{A}g^{-1})\, h^k\rangle = \langle gf(\tilde{A})\, g^{-1}h^k\rangle = \langle g^{-1}gf(\tilde{A})\, h^k\rangle = \langle f(\tilde{A})\, h^k\rangle$ for $g \in G_N$, G_{N^T} where we have used that f is a polynomial without constant term. That $G_{NT} \not\subset \mathscr{L}$ is irrelevant for the discussion. Now define for *any* good function H of $\mathscr{L}^*$ its gradient by

$$dH = \langle \nabla_{\mathscr{L}} H, dA \rangle, \qquad \nabla_{\mathscr{L}} H \in \mathscr{L}, \text{ if it exists.} \tag{4.13}$$

For the above function,

$$dH = \langle f'(\tilde{A})\, h^k\, dA \rangle \equiv \langle P(f'(\tilde{A})), dA \rangle, \tag{4.14}$$

where $P(f'(\tilde{A}))$ is, by the nondegeneracy of $\langle\,,\rangle$, the unique element in $\mathscr{L}$ satisfying the identity (4.14); hence $\nabla H \equiv \nabla_{\mathscr{L}} H$ is perfectly well defined and equals $P(f'(\tilde{A}))$.

If in the previous formula $\langle f(g\tilde{A}g^{-1})\rangle = \langle f(\tilde{A})\rangle$, we substitute in $g = \exp \epsilon B$, $B \in N$, N^T; regarding it as an identity in ϵ, and if we just look at the ϵ term, we have, upon comparison with (4.13) and (4.14), the identity

$$0 = \langle \nabla H, [A, B] \rangle = \langle [\nabla H, A], B \rangle$$

for $B \in N$, N^T, and so for all $B \in \mathscr{L}$; thus $[\nabla H, A] = 0$, which shows $H(A) \in \mathscr{A}(\Gamma)$. We remove the restriction that f be a polynomial; assume that $f'(A)$ makes sense for all h, and that f can be approximated by polynomials in the C^1 sense; and then take limits in formulas (3.17), and formulas of the form $\{H, F\}(A) = 0$.

We now compute (3.17) and (3.19), the symplectic structures and the differential equations. Recall the definition of ∇H in footnote 5. Since N is paired with $K^\perp \approx N^*$ via $\langle\,,\rangle$, we have that $\nabla_{K^\perp} H \in N$ for a function H on $K^\perp$. Since $dH = \langle \nabla_{K^\perp} H,\ dA \rangle = \langle \sum (\nabla_{K^\perp} H)_i h^i,\ \sum_j h^j dA_j \rangle = \sum_i ((\nabla_{K^\perp} H)_{-i},\ dA_i)$, we have that $(\nabla_{K^\perp} H)_{-i} = \nabla_{A_i} H$.

We conclude

$$\nabla_{K^\perp} H = \sum_{i \geqslant 0} (\nabla_{A_i} H)\, h^{-i} \in N = \mathscr{A}_{-\infty, 0}\,. \tag{4.15}$$

Substituting (4.15) into (3.19), we have

$$\{H, F\} = \sum_{j+k=i} (A_i\,, [\nabla_{A_j} H, \nabla_{A_k} F]). \tag{4.16}$$

To compute (3.21), we consider functions defined on $\mathscr{L}^*(\approx \mathscr{L})$ and compute their gradients (4.13),

$$dH = \langle \nabla_{\mathscr{L}} H, dA \rangle, \qquad \text{implies} \qquad \nabla_{\mathscr{L}} H = \sum (\nabla_{A_i} H)\, h^{-i},$$

where $\nabla_{A_i} H \in \mathscr{M}$ for *all* i; and so by (4.7) and the above we have

$$P_K(\nabla H) = \sum_{i>0} (h^i \nabla_{A_{-i}} H - h^{-i}(\nabla_{A_{-i}} H)^T) + (\nabla_{A_0} H)^+ - ((\nabla_{A_0} H)^+)^T. \tag{4.17}$$

For $H \in \mathscr{A}(\Gamma)$, we substitute (4.17) into (3.17).

Equation (4.17) can be improved for the classical Lie algebras if in the definition of H we use their classical matrix representations with which we shall *identify* them. Except for $sl(n)$, the classical Lie algebras are defined by an equation of the form $M\alpha = -\alpha M$, for some definite α, so $M^k\alpha = (-1)^k \alpha(M^k)$, hence $\{x^{2s+1} \mid x \in \mathscr{M}\} \subset \mathscr{M}$, $\{x^{2s+1} \mid x \in \mathscr{L}\} \subset \mathscr{L}$. Thus for $f'(z)$ an odd polynomial in z, $f'(A)$, $f'(A)h^k \in \mathscr{L}$. We can conclude from the above, $\nabla(f(\tilde{A})h^k) = f'(\tilde{A})h^k$, and so from (4.7),

$$\begin{aligned} P_K(\nabla H) = P_K(f'(\tilde{A})\, h^k) &= \sum_{i>0} ((f'(\tilde{A}))_{i-k}\, h^i - (f'(\tilde{A}))^T_{i-k}\, h^{-i}) \\ &\quad + ((f'(\tilde{A})_{-k})^+) - ((f'(\tilde{A})_{-k})^+)^T. \end{aligned} \tag{4.18}$$

It is useful to note that for the case $H = \frac{1}{2}\langle A, A\rangle = \frac{1}{2} \operatorname{Tr} \tilde{A}^2$,

$$P_K(\nabla H) = P_K(A) = \sum_{i>0} (A_i h^i + A_i{}^T h^{-i}) + A_0{}^+ - (A_0{}^+)^T. \tag{4.19}$$

This gives rise to the Toda systems. Now given any representation ϕ, we may use the above-defined Hamiltonians H in the Lax equation (3.17)

$$(\phi(\dot{A})) = [\phi(A), \phi(P_K(\nabla H))],$$

with $P_K(\nabla H)$ as in (4.18).

Let us return to the level of abstract Lie algebras and for $A \in \Gamma_m = \mathscr{A}_{-m,m} \cap K^\perp$ we display explicitly the Lax flow $\dot{A} = [A, P_K(\nabla H)]$ for $H = \frac{1}{2}\langle A, A\rangle$:

$$\begin{aligned} A &= \sum_{j=1}^{r} \tilde{b}_j p_j + \sum_{\substack{i\geqslant 0 \\ 0<i\mathrm{d}+|\alpha|\leqslant m}} a_\alpha^{(i)}(h^{-i}e_{-\alpha} + e_\alpha h^i), \\ P_K(A) &= \sum_{\substack{i\geqslant 0 \\ 0<i\mathrm{d}+|\alpha|\leqslant m}} a_\alpha^{(i)}(-h^{-i}e_{-\alpha} + e_\alpha h^i), \end{aligned} \tag{4.20}$$

Thus from (4.1), $\dot{A} = [A, P_K(A)]$ is a differential equation in the $a_\alpha^{(i)}$, $\tilde{b}_j$, which depends linearly and only on the roots α and the largest integers r such that: $\beta - r\alpha$ is a root given that α, β are roots.

We consider further the special case $m = 1$; the Poisson bracket, the orbit invariant $\mathscr{A}$, and the Lax equation will be computed:

From Eq. (4.16) the Poisson bracket takes on the form

$$\{H, F\} = (A_0\,, [\nabla_{A_0} H, \nabla_{A_0} F]) + (A_1\,, [\nabla_{A_1} H, \nabla_{A_0} F] + [\nabla_{A_0} H, \nabla_{A_1} F]),$$

where

$$\nabla_{A_0} H = \sum_{\alpha \in \Delta_0} \mu_\alpha H_{a_\alpha} e_{-\alpha} + \sum_{i=1}^{r} H_{\tilde{b}_i} p_i, \quad \nabla_{A_1} H = \mu_\delta H_{a_\delta} e_\delta, \quad \mu_\alpha^{-1} = (e_\alpha, e_{-\alpha}),$$

for $\delta =$ the maximal root (see (4.1)). Written out, we get

$$X_H(F) \equiv \{F, H\} = \sum_{\alpha \in \Delta, j} (p_j, \alpha)(H_{\tilde{b}_j} a_\alpha F_{a_\alpha} - F_{\tilde{b}_j} a_\alpha H_{a_\alpha}), \qquad \Delta = \Delta_0 \cup \{-\delta\},$$

where (,) is the Killing form in (4.1), and so the Hamiltonian vector-field X_H is

$$X_H = -\sum_j \left(\sum_{\alpha \in \Delta} (p_j, \alpha) a_\alpha H_{a_\alpha} \right) \partial_{\tilde{b}_j} + \sum_{\alpha \in \Delta} \left(\sum_j ((p_j, \alpha) H_{\tilde{b}_j}) \right) a_\alpha \partial_{a_\alpha}.$$

If $\delta = \sum_{\alpha \in \Delta_0} l_\alpha \alpha$, then $\mathscr{A} \equiv (\prod_{\alpha \in \Delta_0} a_\alpha^{l_\alpha}) a_\delta$ leads[6] to a null vector-field, and hence $\mathscr{A}$ is an orbit invariant; in fact it is the only such invariant (up to a function of it). For observe that

$$X_{\mathscr{A}} = \sum_j \left(p_j, \left(\sum_{\alpha \in \Delta_0} l_\alpha \alpha - \delta \right) \right) \partial_{\tilde{b}_j} = 0. \tag{4.21}$$

If we define $\ln |a| = z_{j(\alpha)}$, the Poisson[7] bracket has the form $\{F, G\} = (J \nabla F, \nabla G)$, $\nabla F = (\nabla_{\tilde{b}} F, \nabla_z F)^T$, $\tilde{b} = (\tilde{b}_1, \dots, \tilde{b}_r)$, $z = (z_1, \dots, z_{r+1})$, $J = \left[\begin{smallmatrix} 0, & P \\ -P^T, & 0 \end{smallmatrix}\right]$, where $P \equiv (p_j(e_\alpha))$. Since the Killing form is non-degenerate on h, and since $\delta = \sum l_\alpha e_\alpha$, the matrix P has rank r, and hence J has rank $2r$, and so the orbits in Γ_1 are typically $2r$ dimensional, the only generic orbit invariant being $\mathscr{A}$. Equation (4.20) has the following form for the case $m = 1$.

$$\begin{aligned} A &= \sum_{j=1}^{r} \tilde{b}_j p_j + \sum_{|\alpha|=1} a_\alpha (e_\alpha + e_{-\alpha}) + a_\delta (h^{-1} e_\delta + h e_{-\delta}), \\ B &\equiv P_K(A) = \sum_{|\alpha|=1} a_\alpha (e_\alpha - e_{-\alpha}) + a_\delta (-h^{-1} e_\delta + h e_{-\delta}), \end{aligned} \tag{4.22}$$

and we thus conclude

THEOREM 4.1. $\dot{A} = [A, B]$ *is equivalent to the following equations of Bogoyavlensky* [21] (*from which he computed* (2.3) *and* (2.4)):

$$\begin{aligned} \dot{\tilde{b}}_j &= -2 \sum_{|\alpha|=1} a_\alpha^2 \cdot \alpha(p_j) + 2 a_\delta^2 \cdot \delta(p_j), \\ \dot{a}_\alpha &= a_\alpha \sum_j \tilde{b}_j \cdot \alpha(p_j), \qquad \dot{a}_\delta = a_\delta \sum_j \tilde{b}_j \cdot \delta(p_j). \end{aligned} \tag{4.23}$$

[6] This observation is due to B. Kostant, who was also aware of the integrability statement (personal communication).

[7] This is a Lie algebra generalization of a formula which appears in [6, p. 78].

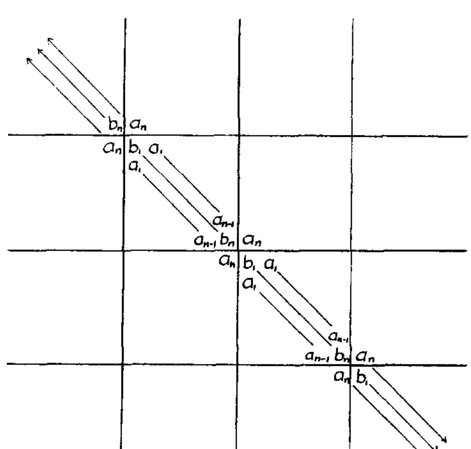
Note that the form of Eq. (4.23) *just depends on the roots.*

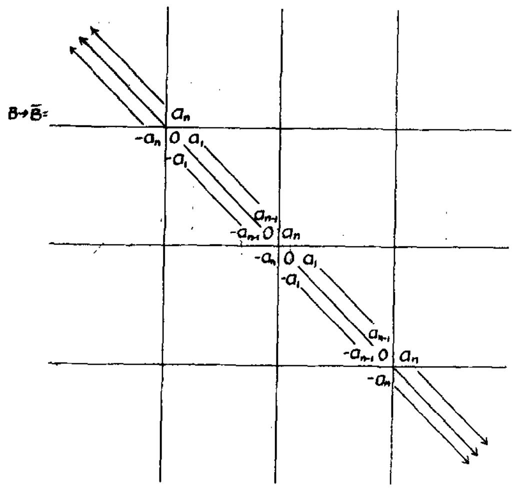
We now exhibit the "symmetric" matrices (4.22) for A_n, B_n, C_n, D_n, G_2, using Gilmore [18, p. 247], and Humphreys [19, p. 103], which contain basic *explicit* information on the Cartan decomposition. For A_{n-1}, the special linear algebra, we have

$$A = \begin{bmatrix} b_1 & a_1 & \cdots & \cdots & a_n h^{-1} \\ a_1 & b_2 & a_2 & & \vdots \\ \vdots & a_2 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & a_{n-1} \\ a_n h & \cdots & \cdots & a_{n-1} & b_n \end{bmatrix} \qquad B = \begin{bmatrix} 0 & a_1 & \cdots & \cdots & -a_n h^{-1} \\ -a_1 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & a_{n-1} \\ a_n h & \cdots & \cdots & -a_{n-1} & 0 \end{bmatrix} \tag{4.24}$$

where $\sum b_i = 0$. Note that these b_i's are linear combinations of the $\tilde{b}_j$'s. Under the isomorphism (see Section 3), $A, B \to \tilde{A}, \tilde{B}$, which are infinite n-periodic Jacobi matrices of the form

$A \to \tilde{A} =$





For C_n , the symplectic algebra, we have

$$A=\left[\begin{array}{cccc|cccc} b_1 & a_1 & & & & & & h^{-1}a_{n+1}\\ a_1 & \ddots & \ddots & & & & & \\ & \ddots & \ddots & a_{n-1} & & & & \\ & & a_{n-1} & b_n & a_n & & & \\ \hline & & & a_n & -b_n & -a_{n-1} & & \\ & & & & -a_{n-1} & \ddots & \ddots & \\ & & & & & \ddots & \ddots & -a_1 \\ h a_{n+1} & & & & & & -a_1 & -b_1 \end{array}\right]$$

$$B=\left[\begin{array}{cccc|cccc} 0 & a_1 & & & & & & -h^{-1}a_{n+1}\\ -a_1 & \ddots & \ddots & & & & & \\ & \ddots & \ddots & a_{n-1} & & & & \\ & & -a_{n-1} & 0 & a_n & & & \\ \hline & & & a_n & 0 & -a_{n-1} & & \\ & & & & a_{n-1} & \ddots & \ddots & \\ & & & & & \ddots & \ddots & -a_1 \\ h a_{n+1} & & & & & & a_1 & 0 \end{array}\right] \qquad (4.25)$$

As in the previous case, it is easy to compute (from Section 3) the symmetric and skew-symmetric matrices $\tilde{A}$, $\tilde{B}$, which are just infinite n-periodic Jacobi matrices having some obvious symmetries. For B_n, the orthogonal algebra $O(n, n+1)$, we have:

$$A = \left[\begin{array}{cccc|c|cccc}
b_1 & a_1 & & & & & & h^{-1}a_{n+1} & 0\\
a_1 & \ddots & \ddots & & & & & 0 & -h^{-1}a_{n+1}\\
 & \ddots & \ddots & a_{n-1} & & & & & \\
 & & a_{n-1} & b_n & a_n & 0 & & & \\
\hline
 & & & a_n & 0 & -a_n & & & \\
\hline
 & & & 0 & -a_n & -b_n & -a_{n-1} & & \\
 & & & & & -a_{n-1} & \ddots & \ddots & \\
h a_{n+1} & 0 & & & & & \ddots & \ddots & -a_1\\
0 & -h a_{n+1} & & & & & & -a_1 & -b_1
\end{array}\right]
\qquad
B = \left[\begin{array}{cccc|c|cccc}
0 & a_1 & & & & & & -h^{-1}a_{n+1} & 0\\
-a_1 & \ddots & \ddots & & & & & 0 & h^{-1}a_{n+1}\\
 & \ddots & \ddots & a_{n-1} & & & & & \\
 & & -a_{n-1} & 0 & a_n & 0 & & & \\
\hline
 & & & -a_n & 0 & -a_n & & & \\
\hline
 & & & 0 & a_n & 0 & -a_{n-1} & & \\
 & & & & & a_{n-1} & \ddots & \ddots & \\
h a_{n+1} & 0 & & & & & \ddots & \ddots & -a_1\\
0 & -h a_{n+1} & & & & & & a_1 & 0
\end{array}\right]
\tag{4.26}$$

For D_n, the orthogonal algebra $O(2n)$, we have:

$$A = \left[\begin{array}{cccc|cccc}
b_1 & a_1 & & & & & h^{-1}a_{n+1} & 0\\
a_1 & \ddots & \ddots & & & & 0 & -h^{-1}a_{n+1}\\
 & \ddots & \ddots & a_{n-1} & -a_n & 0 & & \\
 & & a_{n-1} & b_n & 0 & a_n & & \\
\hline
 & & -a_n & 0 & -b_n & -a_{n-1} & & \\
 & & 0 & a_n & -a_{n-1} & \ddots & \ddots & \\
h a_{n+1} & 0 & & & & \ddots & \ddots & -a_1\\
0 & -h a_{n+1} & & & & & -a_1 & -b_1
\end{array}\right]
\qquad
B = \left[\begin{array}{cccc|cccc}
0 & a_1 & & & & & -h^{-1}a_{n+1} & 0\\
-a_1 & \ddots & \ddots & & & & 0 & h^{-1}a_{n+1}\\
 & \ddots & \ddots & a_{n-1} & -a_n & 0 & & \\
 & & -a_{n-1} & 0 & 0 & a_n & & \\
\hline
 & & a_n & 0 & 0 & -a_{n-1} & & \\
 & & 0 & -a_n & a_{n-1} & \ddots & \ddots & \\
h a_{n+1} & 0 & & & & \ddots & \ddots & -a_1\\
0 & -h a_{n+1} & & & & & a_1 & 0
\end{array}\right]
\tag{4.27}$$

Note for $h = \pm 1$, the A's, B's of these examples are respectively symmetric, skew-symmetric, a reflection of the fact that the corresponding infinite periodic matrix versions are respectively symmetric, and skew-symmetric.

Consider now (4.22) for G_2. From Humphreys [19], it is easy to see, after some fiddling around, that $M \in G_2 \subset B_3$ may be described by a matrix of the form

$$
\left[\begin{array}{c|c|c|c|c|c|c}
T_{11} & T_{12} & T_{13} & -\sqrt{2}x_1 & y_2 & -y_3 & 0 \\ \hline
T_{21} & T_{22} & T_{23} & -\sqrt{2}x_2 & -y_1 & 0 & y_3 \\ \hline
T_{31} & T_{32} & T_{33} & -\sqrt{2}x_3 & 0 & y_1 & -y_2 \\ \hline
\sqrt{2}y_1 & \sqrt{2}y_2 & \sqrt{2}y_3 & 0 & \sqrt{2}x_3 & \sqrt{2}x_2 & \sqrt{2}x_1 \\ \hline
-x_2 & x_1 & 0 & -\sqrt{2}y_3 & -T_{33} & -T_{23} & -T_{13} \\ \hline
x_3 & 0 & -x_1 & -\sqrt{2}y_2 & -T_{21} & -T_{22} & -T_{12} \\ \hline
0 & -x_3 & x_2 & -\sqrt{2}y_1 & -T_{31} & -T_{21} & -T_{11}
\end{array}\right].
$$

From this matrix, after suitably modifying the eigenvector root table given there, one finds A to be of the form

$$
\left[\begin{array}{c|c|c|c|c|c|c}
b_1 & a_2 & h^{-1}a_3 & 0 & a_1 & 0 & 0 \\ \hline
a_2 & b_2 & 0 & \sqrt{2}a_1 & 0 & 0 & 0 \\ \hline
ha_3 & 0 & -b_1-b_2 & 0 & 0 & 0 & -a_1 \\ \hline
0 & \sqrt{2}a_1 & 0 & 0 & 0 & -\sqrt{2}a_1 & 0 \\ \hline
a_1 & 0 & 0 & 0 & b_1+b_2 & 0 & -a_3h^{-1} \\ \hline
0 & 0 & 0 & -\sqrt{2}a_1 & 0 & -b_2 & -a_2 \\ \hline
0 & 0 & -a_1 & 0 & -a_3h & -a_2 & -b_1
\end{array}\right]
$$

and B

$$
\left[\begin{array}{c|c|c|c|c|c|c}
 & a_2 & -a_3h^{-1} & & -a_1 & & \\ \hline
-a_2 & & & \sqrt{2}a_1 & & & \\ \hline
a_3h & & & & & & a_1 \\ \hline
 & -\sqrt{2}a_1 & & & & -\sqrt{2}a_1 & \\ \hline
a_1 & & & & & & a_3h^{-1} \\ \hline
 & & & \sqrt{2}a_1 & & & -a_2 \\ \hline
 & & -a_1 & & -a_3h & a_2 &
\end{array}\right].
$$

After subjecting A, B successively to the four conjugations by row and column permutation matrices summarized by the table, (3, 6) (2, 5) (1, 6) (2, 7), read from left to right, we find

$$A = \begin{bmatrix} -b_1-b_2 & -a_1 & & & & ha_3 & \\ -a_1 & -b_1 & -a_2 & & & & -ha_3 \\ & -a_2 & -b_2 & -\sqrt{2}a_1 & & & \\ & & -\sqrt{2}a_1 & 0 & \sqrt{2}a_1 & & \\ & & & \sqrt{2}a_1 & b_2 & a_2 & \\ h^{-1}a_3 & & & & a_2 & b_1 & a_1 \\ & -h^{-1}a_3 & & & & a_1 & b_1+b_2 \end{bmatrix}$$

$$B = \begin{bmatrix} & a_1 & & & & ha_3 & \\ -a_1 & & a_2 & & & & -a_3h \\ & -a_2 & & \sqrt{2}a_1 & & & \\ & & -\sqrt{2}a_1 & & -\sqrt{2}a_1 & & \\ & & & \sqrt{2}a_1 & & -a_2 & \\ -h^{-1}a_3 & & & & a_2 & & -a_1 \\ & a_3h^{-1} & & & & a_1 & \end{bmatrix} \tag{4.28}$$

Note that this is just a special case of B_3, (see (4.26)), if we make the inessential change $h \to h^{-1}$, $B \to -B$ (the latter reversing time), and so the G_2 system, $\dot{A} = [A, B]$, is a subsystem of the B_3 system. From (4.24)–(4.28), we easily check that $\dot{A} = [A, B]$ is equivalent (after some slight change in variables) to the Toda systems of Section 2, (2.2), (2.7), (2.8), (2.10), with the specified differential equations. Note that we have really shown G_2 sits in B_3 in such a way that the simple root eigenvectors of G_2 are sums of simple root eigenvectors of B_3, and so in general the level m eigenvectors of G_2 are sums of level m

eigenvectors of B_3. Hence all the Toda systems (4.21) of G_2 are subsystems of the B_3 systems.

We conclude this example by computing the orbit Poisson bracket (4.16) for the case of the A_{n-1} orbit of (4.24). We observe

$$A = A_0 + a_n(he_{n,1} + h^{-1}e_{1,n}); \qquad \text{see (4.24) for } A_0 .$$

(In the above, $e_{1,n}$ is the matrix whose only non-zero entry is one in the $(1, n)$ slot.) If $H = H(b_1, \ldots, b_n, a_1, \ldots, a_n)$, we have from (4.15)

$$\begin{aligned}\nabla_{K^\perp} H &= h^{-1}\nabla_{A_1} H + \nabla_{A_0} H, \\ &= h^{-1}H_{a_n}e_{1,n} + \operatorname{diag}(H_{b_1}, \ldots, H_{b_n}) + \sum_{i=1}^{n-1} H_{a_i}e_{i+1,i}\end{aligned}$$

and so

$$\{H, F\} = \sum_{i=1}^{n} a_i F_{a_i}(H_{b_i} - H_{b_{i+1}}) + \sum_{i=1}^{n} F_{b_i}(a_{i-1}H_{a_i} - a_i H_{a_i}),$$

where $H_{b_{n+1}} = H_{b_1}$, $H_{a_0} = H_{a_n}$. This agrees with the standard formula, first observed by Moerbeke [6].

Remark. We note that all the curves $\det(A - z) = Q(z, h) = 0$ have the form $a(h + h^{-1}) = R(z)$, where $a = (\prod_{i=1}^{n} a_i^{v_i})\, a_{n+1}$. We claim that $a = \mathscr{A}$, the orbit invariant computed previously. Remember that all the integrals of the motion of $\dot{A} = [A, B]$ are generated by the coefficients of $Q(z, h)$. We claim that all functions generated by the coefficients of the $R(z)$ are dependent on the b's, and so a is the only b-independent integral, and this must equal some function of $\mathscr{A}$, which by its form must be $\mathscr{A}$ itself. To see the above b independence, observe that for b very large compared to the a_i's, coefficients of $R(z)$ essentially yield the b_i's; hence if the coefficients are labeled $F = (F_1, \ldots, F_n)$, $\det(\partial F/\partial b) \not\equiv 0$, and so given any function $H(F_1, \ldots, F_n)$, for H to be b independent, $\partial F/\partial b \cdot \nabla_F H \equiv 0$, and since $\det(\partial F/\partial b) \not\equiv 0$, we must have $\nabla_F H \equiv 0$, i.e., H cannot depend on the F_i's, thus proving the assertion.

Example 2. The Nonsymmetric Toda Systems

We now discuss the systems associated with nonsymmetric infinite periodic matrices, the nonsymmetric Toda systems. The linearization of these systems was accomplished by Mumford and Moerbeke [3], along with the formula for the symplectic structure. We derive that formula from the point of view of Section 3. We need the following consequence of Theorem 3.1.

THEOREM 4.2. *Let $\hat{L}$ be a Lie algebra with a nondegenerate, ad-invariant,* *bilinear form $\langle\ ,\ \rangle_{\hat{L}}$ through which we identify $\hat{L} \approx \hat{L}^*$. Suppose $\hat{L}$ has the level decomposition*

$$\hat{L} = \oplus \sum_{i \in Z} \hat{L}_i, \qquad [\hat{L}_i, \hat{L}_j] \subset \hat{L}_{i+j}, \qquad [\hat{L}_0, \hat{L}_0] = 0. \tag{4.29}$$

Let $B^+ = \sum_{i \geqslant 0} \oplus \hat{L}_i$, $B^- = \sum_{i \leqslant 0} \hat{L}_i$ and $P^\pm$, P_i be the projections onto $B^\pm$, $\hat{L}_i$, respectively, along $\sum_{i<0} \oplus \hat{L}_i$, $\sum_{i>0} \oplus L_i$, $\sum_{j \neq i} \oplus \hat{L}_j$. We require $\langle \hat{L}_i, (1 - P_{-i})\hat{L}\rangle = 0$, hence $\hat{L}_i$ is paired with $\hat{L}_{-i}$ with respect to $\langle\ ,\ \rangle$ by its nondegeneracy. We define the derived Lie algebra

$$\begin{aligned} &Ł = \{(l^-, l^+) \mid l^+ \in B^+, l^- \in B^-, P_0 l^+ = P_0 l^-\}, \\ \text{where} \quad &[(l_1^-, l_1^+), (l_2^-, l_2^+)] = ([l_1^-, l_2^-], -[l_1^+, l_2^+]). \end{aligned} \tag{4.30}$$

We may identify $Ł^* \approx \hat{L}$ via the nondegenerate pairing between $Ł$, $\hat{L}$

$$\langle (l^-, l^+), l\rangle_{Ł} \equiv \langle (l^+ + l^-), l\rangle_{\hat{L}}, \tag{4.31}$$

and so the connected Lie group $G^{Ł}$ generated by $Ł$ induces through its co-adjoint action the Kostant–Kirillov symplectic structure on $\hat{L} \approx Ł^*$. Then $\Gamma = \Gamma_{kj} = \sum_{k \leqslant i \leqslant j} \oplus L_i$ is an invariant manifold of the co-adjoint action. Let $\hat{\mathscr{A}}(\Gamma)$ be the algebra of smooth functions defined on $\hat{L}$ (identified with $\hat{L}^*$ via $\langle\ ,\ \rangle_{\hat{L}}$), which moreover are invariant on Γ with respect to the co-adjoint L action on $L^* \approx \hat{L}$. Then the $\hat{\mathscr{A}}(\Gamma)$ (upon restriction of its elements to Γ) form a system of involutive functions with respect to the $G^{Ł}$ symplectic structure on Γ. Moreover, the Hamilton vector fields X_{H} and Hamiltonian equations associated with $H \in \hat{\mathscr{A}}(\Gamma)$ are

$$\dot{m} = X_{\mathrm{H}} = [m, n], \qquad n = (P^+ - \tfrac{1}{2}P_0)(\nabla H(m)), \tag{4.32}$$

where $\nabla H(m) \in Ł$ is the gradient of H thought of as a function on $\hat{L}$.

Proof. The proof is most easily accomplished by just mimicking the proof of Theorem 3.1; however, when we showed this theorem to B. Symes, he observed that it could also be thought of as an example of Theorem 3.1 via (4.34) as follows:

In Theorem 3.1 let us take $L = \hat{L} \oplus \hat{L}$, where

$$\begin{aligned} &[(l_1, m_1), (l_2, m_2)] = ([l_1, l_2], -[m_1, m_2]), \\ &\langle\ ,\ \rangle_L = \langle\ ,\ \rangle_{\hat{L}} + \langle\ ,\ \rangle_{\hat{L}}, \end{aligned} \tag{4.33}$$

and take

$$\begin{aligned} &K = \{(m, -m) \mid m \in \hat{L}\} \approx \hat{L}, \qquad N = Ł, \\ &K^\perp = \{(m, m) \mid m \in \hat{L}\} \approx \hat{L}, \qquad N^\perp = \{(l^-, l^+) \mid l^\pm \in B^\pm, P_0(l^+ + l^-) = 0\}. \end{aligned} \tag{4.34}$$

That indeed $N^\perp$ is as stated follows from the orthogonality properties of the L_i. We define $l_\pm = (\sum_{i \gtrless 0} P_i)l$, $l_0 = P_0 l$, hence $l = l_+ + l_0 + l_-$, and also defining $P_\pm = P^\pm - \frac{1}{2}P_0$, we observe $P_+ + P_- = I$. We now compute

$$\begin{aligned} P_K(m, l) &= (P_+m - P_-l, P_-l - P_+m), \\ P_N(m, l) &= (P_-m + P_-l, P_+l + P_+m), \\ P_{K^\perp}(m, l) &= (P_+m + P_-l, P_+m + P_-l), \\ P_{N^\perp}(m, l) &= (P_-m - P_-l, -P_+m + P_+l). \end{aligned} \tag{4.35}$$

From (3.22), (4.30), (4.10), (4.35), it follows that the $G^{\hat{L}}$ orbit $\mathcal{O}_m \subset K^\perp \approx \hat{L}$ through m equals (first without the $P^\pm$ which we may add without any harm)

$$\mathcal{O}_m = \{P_+(g_-^{-1}(P^+m)\, g_-) + P_-(g_+(P^-m)\, g_+^{-1}) \mid (g_-, g_+) \in G^{\hat{L}}\}; \tag{4.36}$$

hence

$$T\mathcal{O}_m\,|_m = \{P_+([P^+m, l_-]) + P_-[l_+, P^-m] \mid (l_-, l_+) \in \hat{L}\}, \tag{4.37}$$

from which it is immediate that $\Gamma = \Gamma_{kj}$ is an invariant manifold of the $G^{\hat{L}}$ action. For $\mathscr{A}(\Gamma)$ of Theorem 3.1, take the special $H(m, l) = f(m)$, where $f(m)$ is $\hat{L}$ invariant as discussed in the statement of the theorem. Then the gradient ∇H of H in $\hat{L} \approx \hat{L}^*$ is $(\nabla H(m), 0)$, where $\nabla H \in L$ is the gradient of f in $\hat{L}$; hence using (4.35): the Lax Hamiltonian vector-field (3.17) on Γ, reads for this special H,

$$\dot{m} = [m, P_+(\nabla H(m))], \tag{4.38}$$

in agreement with (4.32). This concludes the proof.

Remark 3. We compute the Poisson bracket (3.19) on Γ. Since $K^\perp \approx \hat{L}$ (see (4.34)), any $K^\perp$ function is also an $\hat{L}$ function; if $\nabla_{\hat{L}} f$ is its gradient when viewed as a function on $\hat{L}$, then by (4.35), $\nabla_{K^\perp} f = (P_-\nabla_{\hat{L}} f, P_+\nabla_{\hat{L}} f)$. And by (3.19) and (4.34),

$$\begin{aligned} \{f, g\}(m) &= \langle (m, m), [(P_-\nabla_{\hat{L}} f, P_+\nabla_{\hat{L}} f), (P_-\nabla_{\hat{L}} g, P_+\nabla_{\hat{L}} g]\rangle_L \\ &= \langle m, [P_-\nabla_{\hat{L}} f, P_-\nabla_{\hat{L}} g] - [P_+\nabla_{\hat{L}} f, P_+\nabla_{\hat{L}} g]\rangle_L . \end{aligned} \tag{4.39}$$

We apply these considerations to the example $\hat{L} = \mathscr{L} = \sum \oplus \mathscr{L}_i$ of the previous example, where for concreteness we take $\mathscr{M} = sl(n)$. The construction of $G^{\hat{L}}$ is exactly as the construction of G_N previously described, yielding a filtration of $G^{\hat{L}}$ by $G^{\hat{L}_{jk}}$, where $\hat{L}_{jk} = \hat{L} \cap (\mathscr{A}_{-j,0}, \mathscr{A}_{0,k})$. Here we view $\hat{L}_{jk}$ as a Lie algebra with a truncated product rule inherited from N_j and $N_k{}^T$ in the first example; therefore the Γ_{jk} are really invariant manifolds for the co-adjoint action of $G^{\hat{L}_{jk}}$. This is the result analogous to that for the Γ_m of the last example.

The invariant functions have the form, as usual, $H = \langle f(A)\, h^k \rangle$ (see (3.8)) and the Lax equations are of the form (from (4.38))

$$\dot{A} = [A, P_+(f'(A)\, h^k)]. \tag{4.40}$$

Equations (4.39) and (4.40) are in agreement with the formulas of Mumford and Moerbeke [3].

Example 3. *The Spinning Top and Ellipsoid Examples*

In this example we set in Theorem 3.1 $L = \mathscr{L}$, $\mathscr{M} = gl(n, R)$ or $gl(n, \mathbb{C})$ (that $\mathscr{M}$ is not semi-simple is irrelevant), $\langle\, ,\, \rangle_L = \langle\, ,\, \rangle_1$ of (3.6), and[8]

$$K = K^{\perp} = \mathscr{A}_{0,\infty}\,, \qquad N = N^{\perp} = \mathscr{A}_{-\infty,-1}\,,$$

so (4.41)

$$P_K\left(\sum A_i h^i\right) = \sum_{i\geqslant 0} A_i h^i \equiv \left(\sum A_i h^i\right)_+ \,, \; P_N\left(\sum A_i h^i\right) = \sum_{i<0} A_i h^i \equiv \left(\sum A_i h^i\right)_- .$$

For the invariant manifold Γ we take

$$\Gamma = \Gamma_m(\alpha, \gamma) = \alpha h^m + \gamma h^{m-1} + \mathscr{A}^0_{0,m-1}\,, \tag{4.42}$$

with $\alpha = \operatorname{diag}(\alpha_1, \ldots, \alpha_n)$, $\prod_{i<j}(\alpha_i - \alpha_j) \neq 0$, γ a diagonal matrix and $\mathscr{A}^0_{0,m-1} = \{\sum_{0\leqslant i\leqslant m-1} A_i h^i \mid \operatorname{diag}(A_{m-1}) = 0\}$. The terms α, γ are in fact orbit invariants, or parameters.

Actually we may take $\Gamma_m = \Gamma_m(\alpha) = \alpha h^m + \mathscr{A}_{0,m-1}$, where α is some constant diagonalizable matrix, but since $U\alpha U^{-1} = \operatorname{diag}(\alpha_1, \ldots, \alpha_n)$ for some U, we may as well map $\mathscr{L} \to U\mathscr{L}U^{-1}$, and hence we shall just take α to be already in diagonal form. We also make the nondegeneracy assumption that the α_i's are distinct, and thus find ourselves in the situation of (4.42).

For $\mathscr{A}(\Gamma)$ we shall take functions of the form $H = \langle f(Ah^{-j}), h^k \rangle_1$, and so (3.17) becomes

$$\dot{A} = [A, (f'(Ah^{-j})\, h^{k-j})_+], \qquad A = \alpha h^m + \gamma h^{m-1} + \sum_{0\leqslant i\leqslant m-1} A_i h^i. \tag{4.43}$$

For the case $j = m$, $k = m + 1$, one easily computes from (4.43) (first for $f(x) = x^t$, and then in general)

$$\begin{aligned} \dot{A} &= [A, E + \beta h], \qquad \beta = f'(\alpha), \quad E = \alpha d_\beta \alpha d_\alpha^{-1} A_{m-1} + f''(\alpha) \cdot \gamma, \\ E_{ij} &= (1 - \delta_{ij})(\beta_i - \beta_j)(\alpha_i - \alpha_j)^{-1}(A_{m-1})_{ij} + \delta_{ij}\gamma_{ij} f''(\alpha_i), \\ H &= \langle f(Ah^{-m}), h^{m+1} \rangle_1 . \end{aligned} \tag{4.44}$$

[8] In analogy to (4.5), we define $\mathscr{A}_{jk} = \{\sum_{j\leqslant i\leqslant k} A_i h^i \mid A_i \in \mathscr{M}\}$.

We examine the above example. That $K = K^{\perp}$, $N = N^{\perp}$ under $\langle\, ,\rangle_1$, is easily checked, as is the validity of the remainder of (4.41). The group G_N is the affine group $G_N = I + N$, and hence if $I - x \in G_N$, $(I-x)^{-1} = I + x + x^2 + \cdots$. Thus N is filtrated by Lie algebras $N_k = \mathscr{A}_{-k,-1}$, with the bracket $[\sum_{i=-k}^{-1} A_i h^i, \sum_{j=-k}^{-1} B_j h^j] = \sum_{i+j=-k}^{-1} [A_i, B_j]\, h^{i+j}$, and the corresponding Lie group is $G_{N_k} = I + N_k$, with the truncated multiplication rule

$$\sum_{i=-k}^{0} A_i h^i \cdot \sum_{j=-k}^{0} B_j h^j = \sum_{i+j=-k}^{0} A_i B_j h^{i+j},\ A_0 = B_0 = I.$$

This is analogous to the situation in Example 1. The orbits $\mathcal{O}_A$ through $A \in \Gamma_m$ are of the form, by (3.21), (4.10), (4.41),

$$\mathcal{O}_A = \{(g^{-1}Ag)_+ \mid g \in 1 + N\};$$

since only the nonnegative terms in $g^{-1}Ag$ register in $(g^{-1}Ag)_+$, the orbit is easily seen to equal

$$\mathcal{O}_A = \{(g^{-1}Ag)_+ \mid g \in G_{N_{m+1}}\} = \{(g^{-1}Ag)_+ \mid g \in G_{N_m}\},$$

and so, (4.45)

$$T\mathcal{O}_A(A) = \{[A, B]_+ \mid B \in N_m\}.$$

In the first formula since $A \in \mathscr{A}_{0,m} \approx N_{m+1}^*$, under $\langle\, ,\rangle_1$, we think of A being in the dual of N_{m+1}, and so $\mathcal{O}_A$ can really be identified with the co-adjoint orbit of the finite-dimensional Lie group $G_{N_{m+1}}$.

That the given function $H \in \mathscr{A}(\Gamma)$ is proven in the same fashion as in Example 1, as is (4.43). That (4.44) follows from (4.43) is most easily seen for $f(x) = x^{t+1}/(t+1)$, as for a general matrix A_{m-1} we have $(h(\alpha + h^{-1}A_{m-1} + \cdots)^t)_+ = \alpha^t h + [(A_{m-1})_{ij} \sum \alpha_i^s \alpha_j^{t-1-s}] = \alpha^t h + [(1 - \delta_{ij})(A_{m-1})_{ij}(\alpha_i^t - \alpha_j^t)(\alpha_i - \alpha_j)^{-1} + \delta_{ij} t A_{ii} \alpha_j^{t-1}] = E + \beta h$. Subsequently, for a polynomial f we have (4.44) by the above and linearity, while in general we C^1 approximate f by a sequence of polynomials to obtain (4.44). Using $\nabla_{K^{\perp}} H = \sum_{j \geq 0} h^{-j-1} \nabla_{A_j} H$ (compare with (4.15)) the Poisson bracket (3.19) of Theorem 3.1 is easily seen to be:

$$\{H, F\} = \sum_{i=j+k+1} (A_i, [\nabla_{A_j} H, \nabla_{A_k} F]), \qquad \text{with} \qquad (A, B) = \operatorname{trace} A \cdot B.$$

From (4.42) and (4.45), it is clear that $\Gamma = \Gamma_m$ is an invariant manifold; a typical orbit $\mathcal{O}_A$, $A \in \Gamma_m$ will be shown to be of dimension $mn(n-1)$. It suffices to compute $\dim(T\mathcal{O}_A(A))$. Replace $\delta + A_{m-1}$ by A_{m-1}; then first compute $[A, B]_+ = [\alpha h^m + \sum_{i=1}^{m-1} A_i h^i,\ \sum_{k=1}^{m} B_{-k} h^{-k}]_+ = [\alpha, B_{-1}] h^{m-1} + ([\alpha, B_{-2}] + [A_{m-1}, B_{-1}]) h^{m-2} + \cdots + ([\alpha, B_{-k}] + \mathscr{F}(B_{-1}, \ldots, B_{-k-1})) h^{m-k} + \cdots + (\ ) h^0 \in \mathscr{A}_{0,m-1}^0$. So for a fixed A, and given $C = \sum_{i=1}^{m-1} C_i h^i$ we wish to solve the linear system, $[A, B]_+ = C$ for $B \in N_m$. This breaks up into a triangular system of m matrix equations, to be solved by induction:

$$-[\alpha, B_{-k}] = [A_{m-1}, B_{-k+1}] + \cdots + [A_{m-k+1}, B_{-1}] - C_{m-k}, \qquad 1 \leqslant k \leqslant m,$$

where $[\alpha, B_{-k}]_{ij} = (B_{-k})_{ij}(\alpha_i - \alpha_j)$ has a zero diagonal; the diagonal of B_{-k} is therefore irrelevant and may be assumed equal to 0. Consequently this system is solvable uniquely if and only if the diagonal of the right-hand side vanishes, which determines the diagonal of each C_k $(0 \leqslant k \leqslant m-1)$. Therefore the dim $T\mathcal{O}_A(A)$ or, what is the same, the dimension of the range of the linear operator $B \to [A, B]_+$ is $mn(n-1)$.

We now come to the remaining examples of Section 2.

THEOREM 4.3. *If in* (4.44) *we set* $m = 1$, $\gamma = 0$, $\beta = \alpha^{1/2}$, *we arrive at the Euler–Arnold spinning top for the Lie algebra* $gl(n)$, *while if we set* $A_0 = -A_0{}^T$, *we arrive at the Euler–Arnold top for* $u(n)$.

Proof. The proof is a direct consequence of the definitions [1], upon setting $h = 0$ in (4.44).

Now let us define for $x, y \in R^n$ or C^n,

$$\Gamma_{xy} = x \otimes y - y \otimes x, \quad \Gamma_{xx} = x \otimes x, \quad \Gamma_{yy} = y \otimes y, \quad \Delta_{xy} = x \otimes y + y \otimes x, \tag{4.45}$$

where $(x \times y)_{ij} = x_i y_j$. If in Theorem 4.3 we set $A_0 = \Gamma_{xy}$, then as time evolves A_0 remains in this form; but this is part of a more general observation:

THEOREM 4.4. *We have that* (a) *the special Arnold–Euler equations* (2.26), (b) (i) *the geodesic flow on an ellipsoid,* (ii) *the Neumann problem* (2.24), *and* (c) *the central force problem on an ellipsoid associated with the Hamiltonian* (2.18) *correspond respectively to the following three cases of the Lax equation,* $\dot{A} = [A, \Gamma + \beta h]$, $A = A(h)$, $\beta = \mathrm{diag}(\beta_1, \ldots, \beta_n)$, $\Gamma = \mathrm{ad}_\beta\, \mathrm{ad}_\alpha^{-1}\, \Gamma_{xy}$ (see (4.44)),

$$\begin{aligned} &(a) \quad A = \alpha h + \Gamma_{xy}, \\ &(b) \quad A = \alpha h^2 + h\Gamma_{xy} - \Gamma_{xx}, \\ &(c) \quad A = \alpha h^2 + h\Gamma_{xy} + (\Delta_{xy} - \alpha), \end{aligned} \tag{4.46}$$

with the Hamiltonians $H = \langle f(Ah^{-m}), h^{m+1} \rangle$ *respectively of the form*

$$\begin{aligned} &(a) \quad H = \langle \tfrac{2}{3}(Ah^{-1})^{3/2}, h^2 \rangle_1, && f(x) = \tfrac{2}{3}x^{3/2}, \\ &(b) \quad (i) \quad H = \langle \ln(Ah^{-2}), h^3 \rangle_1, && f(x) = \ln x, \\ &\qquad\ (ii) \quad H = \langle \tfrac{1}{2}(Ah^{-2})^2, h^3 \rangle_1, && f(x) = \tfrac{1}{2}x^2, \\ &(c) \quad H = \langle \ln(Ah^{-2}), h^3 \rangle_1, && f(x) = \ln(x), \end{aligned} \tag{4.47}$$

which drives the vector-fields X_H *of our abstract Lie algebra theorem* (3.1) *for the* K, N *decomposition* (4.41). *Moreover,* $\dot{A} = [A, E + \beta h]$, *for the above choices of Hamiltonians are in fact a consequence of the standard Hamiltonian equations*

$$\dot{x}_i = \frac{\partial H}{\partial y_i}, \qquad \dot{y}_i = -\frac{\partial H}{\partial x_i}, \qquad i = 1,\ldots, n, \quad H = H(x, y),$$

with the respective Hamiltonians (distinctly different from (4.47))

$$\begin{aligned}
&\text{(a)} \quad H = -\tfrac{1}{3}\langle (h^{-1}A(h))^{3/2}, h\rangle_1 = \tfrac{1}{2}\textstyle\sum \alpha_i^{1/2} H_i(x, y),\\
&\text{(b)} \quad \text{(i)} \quad H = \tfrac{1}{2}\langle \ln(h^{-2}A(h)), h\rangle_1 = \tfrac{1}{2}\textstyle\sum \alpha_i^{-1} F_i(x, y),\\
&\qquad\;\; \text{(ii)} \quad H = \tfrac{1}{4}\langle \tfrac{1}{2}(h^{-2}A(h))^2, h\rangle_1 = \tfrac{1}{2}\textstyle\sum \alpha_i F_i(x, y),\\
&\text{(c)} \quad H = \tfrac{1}{2}\langle \ln(h^{-2}A(h)), h\rangle_1 = \tfrac{1}{2}\textstyle\sum \alpha_i^{-1} G_i(x, y),\\
&\qquad F_i = x_i^2 + H_i, \quad G_i = -2x_i y_i + H_i, \quad H_i(x, y) = \sum_{j \neq i} \frac{(x_i y_j - x_j y_i)^2}{(\alpha_i - \alpha_j)}.
\end{aligned} \tag{4.48}$$

In fact we have the general formula for the three cases (see (2.25))

$$\begin{aligned}
&\text{(a)} \quad -\tfrac{1}{2}\langle f(h^{-1}A(h)), h\rangle_1 = H_\beta = \tfrac{1}{2}\textstyle\sum \beta_\nu H_\nu, \qquad \beta = f'(\alpha),\\
&\text{(b)} \quad -\tfrac{1}{2}\langle f(h^{-2}A(h)), h\rangle_1 = F_\beta = \tfrac{1}{2}\textstyle\sum \beta_\nu F_\nu,\\
&\text{(c)} \quad -\tfrac{1}{2}\langle f(h^{-2}A(h)), h\rangle_1 = G_\beta = \tfrac{1}{2}\textstyle\sum \beta_\nu G_\nu .
\end{aligned} \tag{4.49}$$

Each of these three sets of functions form a complete set of commuting integrals in the standard (x, y) *Hamiltonian structure. (Actually one must add* $\sum x_i^2$, $\sum y_i^2$, *to the* H_β *to generate a complete set in case* (a)). *These functions and those of* (4.47) *are generated precisely by the coefficients of the hyperelliptic characteristic curves*[9] *of Section* 3: $\det(h^{-1} A - z) = 0$, $\det(h^{-2}A - z) = 0$, $\det((h^2 - 1)^{-1} A - z) = 0$ *for cases* (a), (b), *and* (c), *respectively; they have the form (see* (2.19), (2.20), *and* (2.21))

$$\begin{aligned}
&\text{(a)} \quad h^2 = -\Pi_z(y, x) = -(|x|^2 |y|^2 - \langle x, y\rangle^2) \prod_{i=1}^{n-2} (z - \sigma_i)/a(z),\\
&\text{(b)} \quad h^2 = -\Phi_z(y, x) = -|x|^2 \prod_{i=1}^{n-1} (z - \mu_i)/a(z),\\
&\text{(c)} \quad h^2 - 1 = -\Psi_z = 2\langle x, y\rangle \cdot \prod_{i=1}^{n-1} (z - \nu_i)/a(z),
\end{aligned} \tag{4.50}$$

[9] Note that for $\phi(h)$ a rational function, the curve $\det(\phi(h) - z) = 0$ is merely a reparametrization of the curve $\det(A - z) = 0$. We have picked $\phi(h)$ to desingularize the latter curve.

with $\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i$, $|x|^2 = \langle x, x\rangle$, $a(z) = \det(z - \alpha) = \prod_{i=1}^{n}(z - \alpha_i)$, *and the genera of the curves are* $n - 2$, $n - 1$, $n - 1$, *respectively. Finally the* σ_i*'s and* α_j*'s are independent in* (a), *the* μ_i, α_j*'s in* (b), *and the* ν_i, α_j*'s in* (c).

Proof. The proof is a straightforward computation. Assume for the moment Eqs. (4.48), (4.49). We first do case (b). For $F_\beta = \frac{1}{2}\sum \beta_i F_i$ (see (4.48)), it is easy to check that Hamilton's equations, $\dot{x} = \partial_y F_\beta = -\Gamma x$, $\dot{y} = -\partial_x F_\beta = -\Gamma y - \beta x$, imply, upon using the matrix identity $V(x \otimes y)W = (Vx) \otimes (W^T y)$,

$$\dot{\Gamma}_{xx} = [\Gamma_{xx}, \Gamma], \qquad \dot{\Gamma}_{xy} = [\Gamma_{xy}, \Gamma] + [\Gamma_{xx}, -\beta]; \tag{4.51}$$

this set of equations is equivalent to $\dot{A} = [A, B]$, $A = \alpha h^2 + h\Gamma_{xy} - \Gamma_{xx}$, $B = \Gamma + \beta h$, using $[\Gamma, \alpha] = [\Gamma_{xy}, \beta]$.

Taking $\beta = \alpha^{-1}$, α respectively recovers cases (i), (ii). Comparing (4.44) with the above $\dot{A} = [A, B]$ proves the first statement of the theorem for case (b). To prove case (a), consider $H = \frac{1}{2}\sum \beta_i H_i$ (see (4.48)) and then Hamilton's equations with Hamiltonian H_β imply $\dot{\Gamma}_{xy} = [\Gamma_{xy}, \Gamma]$; hence we have $\dot{A} = [A, B]$, for $A = \alpha h + \Gamma_{xy}$, $B = \Gamma + \beta h$. Now pick $\beta = \alpha^{1/2}$ to recover the first statement of the theorem with regard to case (a). For case (c), if $G = \frac{1}{2}\sum \beta_\nu G_\nu$ (see (4.48)), we have that $\dot{x} = \partial_y G_\beta$, $y = -\partial_x G_\beta$ imply

$$\dot{\Gamma}_{xy} = [\Gamma_{xy}, \Gamma] + [\Delta_{xy}, \beta], \qquad \dot{\Delta}_{xy} = [\Delta_{xy}, \Gamma] + [\Gamma_{xy}, \beta], \tag{4.52}$$

which is equivalent to $\dot{A} = [A, B]$, for $A = \alpha h^2 + h\Gamma_{xy} + (\Delta_{xy} - \alpha)$, $B = \Gamma + \beta h$, and upon taking $\beta = \alpha^{-1}$ we recover case (c). We now verify (4.49), which implies (4.48). It suffices to prove the above for $f(x) = x^k$, by the same reasoning used to check (4.44). Note that for case (a), we need only look at the h^{-2} coefficient of $((h^{-1}A(h)))^k$, while cases (b), (c) are straightforward generalizations of case (a).

We are now facing two distinct symplectic structures, the first deriving from the orbit method applied to Euclidean algebras and the second being the customary Darboux symplectic structure. From the general theory, all expressions of the form $\langle f(Ah^{-m}), h^{m+1}\rangle_1$ are in involution with regard to the first symplectic structure. We now show that they also are with regard to the second one, i.e.,

$$\{H_\beta, H_{\beta'}\} \equiv \sum \frac{\partial(H_\beta, H_{\beta'})}{\partial(x_i, y_i)} = 0.$$

By the preceding arguments, the flow $\dot{x} = \partial_y(H_\beta)$, $\dot{y} = -\partial_x(H_\beta)$ has also the Lax type $\dot{A}(h) = [A(h), B(h)]$; consequently, any polynomial $P(A(h), h)$ satisfies the same Lax flow; hence (d/dt) (trace $(P(A(h), h)) = \sum \dot{c}_\nu h^\nu = 0$ identically in h, implying $\dot{c}_\nu = 0$. But since $H_{\beta'}$ appears among the expressions

c_ν for some appropriate choice of polynomiol P, we have that indeed $0 = \dot{H}_{\beta'} = \{H_{\beta'}, H_\beta\}$. Since the latter holds identically in β and β', we also have $\{H_j, H_k\} = 0$, $1 \leqslant j \leqslant k \leqslant n$, and likewise for the F_i's and G_i's. In the next section we shall see that this Darboux symplectic structure also derives from the orbit method, but for different groups.

That the characteristic curves are given by (4.50) is an easy consequence of the following remark, due to J. Moser.

Remark 1. Consider the rank 2 perturbation of α,

$$\begin{aligned} L = \alpha + P &\equiv \alpha + x \otimes x' + y \otimes y' \\ &\equiv \alpha + ax \otimes x + bx \otimes y + cy \otimes x + dy \otimes y. \end{aligned}$$

Using the above perturbation-type identity $(z - L) = (z - \alpha)(I - R_z P)$, $R_z = (z - \alpha)^{-1}$, one computes

$$\frac{\det(z - L)}{\det(z - \alpha)} = \det(I - W_z),$$

where

$$W_z = \begin{bmatrix} (R_z x, x'), (R_z x, y') \\ (R_z y, x'), (R_z y, y') \end{bmatrix} = \begin{bmatrix} Q_z(x), Q_z(x, y) \\ Q_z(x, y), Q_z(y) \end{bmatrix} \begin{bmatrix} a, c \\ b, d \end{bmatrix}$$

(see Section 2 for notation), and so

$$\begin{aligned} \frac{\det(z - L)}{\det(z - \alpha)} &= 1 - \operatorname{tr} W_z + \det W_z = 1 - \mathscr{J}_z(x, y), \\ \mathscr{J}_z(x, y) &= aQ_z(x) + (b + c)\, Q_z(x, y) + dQ_z(y) \\ &\quad - (ad - bc)(Q_z(x)\, Q_z(y) - Q_z^2(x, y)). \end{aligned} \tag{4.53}$$

Thus, to get (4.50a), set $b = -c = h^{-1}$, $a = d = 0$. To get (4.50b), let $a = -h^{-2}$, $b = -c = h^{-1}$, $d = 0$, and for (4.50c), set $a = d = 0$, $b = (h - 1)^{-1}$, $c = -(h + 1)^{-1}$, We expand upon Theorem 4.4 in a series of remarks and in Section 5.

Remark 2. It is easy to see $\mathscr{J}_z(x, y) = \sum(J_j/(z - \alpha_j))$, where $J_j(y, x) = ax_j^2 + (b + c)\, x_j y_j + dy_j^2 - (ad - bc)\, H_j(y, x)$, and we set

$$2J_\beta = \sum J_j \beta_j = a\langle \beta x, x\rangle + (b + c)\langle \beta x, y\rangle + d\langle \beta y, y\rangle - (ad - bc)\, H_\beta\,.$$

One checks exactly as in Theorem (4.4) that $\dot{x}_i = -(J_\beta)_{x_i}$, $\dot{y}_i = (J_\beta)_{x_i}$, $i = 1, \ldots, n$, implies the Lax equation of (4.44), $\dot{A} = [A, B]$, $m = 2$, with

$$A = \left\{ \left[\frac{(b-c)^2}{4\delta} \right] \cdot \alpha h^2 + \left[\frac{b-c}{2} \right] \cdot \Gamma_{xy} h \right.$$
$$\left. + \left(a\Gamma_{xx} + \left(\frac{b+c}{2} \right) \cdot \Delta_{xy} + d\Gamma_{yy} + \alpha \right) - \frac{(b-c)^2}{4\delta} \cdot \alpha \right\},$$
$$B = \delta\Gamma + \left(\frac{b-c}{2} \right) \beta h, \qquad \delta = ad - bc, \qquad \Gamma = \mathrm{ad}_\beta \mathrm{ad}_\alpha^{-1} \Gamma_{xy}.$$

Putting $\mu = ((b-c)/2^2(h^2-1)+1$ in $\det(A\mu^{-1} - z) = 0$, the corresponding hyperelliptic curve reads

$$-\left(h \left(\frac{b-c}{2\delta} \right) \right)^2 = \left(\frac{d}{\delta} - Q_z(x) \right) \left(\frac{a}{\delta} - Q_z(y) \right) - \left(\frac{b+c}{2\delta} + Q_z(x, y) \right)^2.$$

This clearly comprises the special cases (4.46).

Remark 3. The Lax equations of (4.44)–(4.46), $\dot{A} = [A, B]$, imply Lax equations of the form $\dot{L} = [L, \Gamma]$, with L respectively set equal to

$$\begin{aligned}
&\text{(a)} \quad (I - P_x)(I - P_y)\, \alpha (I - P_y)(I - P_x), && P_x = |x|^{-2} \Gamma_{xx}, \\
&\text{(b)} \quad (I - P_x)(\alpha - \Gamma_{yy})(I - P_x), && \Gamma_{yy} = y \otimes y, \\
&\text{(c)} \quad (I - P_{xy})(\alpha + \tfrac{1}{2}\Gamma_{xy})(I - P_{xy}), && P_{xy} = \langle x, y \rangle^{-1} x \otimes y, \\
&\quad\ \text{and} && \Gamma = \mathrm{ad}_\beta \mathrm{ad}_\alpha^{-1} \Gamma_{xy}.
\end{aligned} \tag{4.54}$$

Case (4.54b) is due to J. Moser and we sketch a proof which generalizes to the other cases. In (4.46b), substituting $h \to ih$, and keeping h real, $A = -\alpha h^2 + i\Gamma_{xx} h - \Gamma_{xx}$ is self-adjoint. We claim that for h small, $-A$ has the orthogonal decomposition

$$-A = \lambda v \otimes \bar{v} / |v|^2 \oplus h^2 \{ L + O(h) \},$$

with $\lambda = |x|^2 + h^2 |y|^2 + O(h^3)$ (the large eigenvalue for $h \to 0$) and $v = x + iyh + O(h^2)$, the corresponding eigenvector. This claim follows immediately from the observation $h^2(-A + \lambda v \otimes \bar{v}/|v|^2) = \alpha - y \otimes y + x \otimes s + s' \otimes x + O(h)$, for some s, s', upon projecting each side of the above formula onto the orthogonal complement of v. By the decomposition for A, when it undergoes an isospectral deformation generated by $\Gamma + \beta h$, so does its summands, in particular $L + O(h)$; and letting $h \to 0$ we have the remark for case (b). Cases (a) and (c) are treated similarly; case (c) involves non-self-adjoint matrices. The conceptual point to bear in mind is that these matrices are descendants of the matrices (4.46). The isospectral flow (4.54b) has a particularly pretty geometrical interpretation: Charsle's theorem on tangents to confocal quadrics and M. Reid's description of hyperelliptic Jacobeans are a consequence of this theory (via a result of Knörrer [30]); see Moser [7] and our subsequent paper.

Remark 4. For the system of Rosochatius discussed in [7], namely the motion of a particle constrained to lie on the sphere $|x| = 1$ in R^n under the influence of a potential

$$U = \frac{1}{2}\left(\langle \alpha x, x\rangle + \sum_{j=1}^{n} c_j^2 x_j^{-2}\right),$$

one has the differential equations

$$\ddot{x} = -U_x - \lambda x, \qquad \lambda = |\dot{x}|^2 - \langle U_x, x\rangle.$$

This implies the quation $\dot{A} = [A, B]$, with

$$\begin{aligned} A &= \alpha h^2 + h(\Gamma_{xy} + i\Delta_{x,c/x}) - \Gamma_{xx}, \qquad i = \sqrt{-1}, \\ B &= \beta h + \mathrm{ad}_\beta \mathrm{ad}_\alpha^{-1}(\Gamma_{xy} + i\Delta_{x,c/x}) - iD, \end{aligned}$$

where c/x is the vector with components c_j/x_j; here $\Delta_{x,c/x} = x \otimes c/x + c/x \otimes x$, the second matrix in B is defined as having no diagonal component, and the third $D = \mathrm{diag}(D_1, \ldots, D_n)$,

$$D_j = \sum_{k\neq j} \left(\frac{\beta_j - \beta_k}{\alpha_j - \alpha_k}\right)\left(\frac{c_j}{x_j^2} + \frac{c_k}{x_k^2}\right) x_k^2.$$

For the particular flow above we take $\beta = \alpha$. But, since in general we have as in (4.49b), $\beta = f'(\alpha)$,

$$\begin{aligned} &-\tfrac{1}{4}\langle f(h^{-2}A(h)), h\rangle_1 \\ &\qquad \equiv \hat{F}_\gamma = \tfrac{1}{2}\left\{\langle \beta x, x\rangle + \sum_{i<j}\left(\frac{\beta_i - \beta_j}{\alpha_i - \alpha_j}\right)\left[(x_i y_j - x_j y_i)^2 - \frac{x_i^2 c_j^2}{x_j^2} - \frac{x_j^2 c_i^2}{x_i^2}\right]\right\} \end{aligned}$$

(modulo a constant factor $-2\sum_{i<j}([(\beta_i - \beta_j)/(\alpha_i - \alpha_j)]\, c_i c_j)$; it makes sense to study the differential equations

$$\dot{x} = (\hat{F}_\beta)_y, \qquad \dot{y} = -(\hat{F}_\beta)_x,$$

which lead to $\dot{A} = [A, B]$ for the general $B = B(\beta)$ given. Since we may write

$$Ah^{-2} = \alpha + h^{-1}x \otimes (-h^{-1}x + y + ic/x) + (-y + ic/x) \otimes h^{-1}x,$$

we can use Remark 1 to compute the hyperelliptic curve

$$h^2(\det(Ah^{-2} - z)) = h^2 - 2sh + t = 0,$$

with

$$t = Q_z(x) + (Q_z(x)\, Q_z(y) - Q_z^2(x, y)) + (Q_z(x)\, Q_z(c/x) - Q_z^2(x, c/x)),$$
$$s = iQ_z(x, c/x),$$

which is thus equivalent to the irrationality $y^2 = (s^2 - t)\, a(z)^2$, i.e.,

$$y^2 = a(z)^2[Q_z^2(x, y) - Q_z(x)\{1 + Q_z(y) + Q_z(c/x)\}].$$

The L matrix of Remark 3 is computed in the usual fashion to be

$$(I - P_x)[\alpha + (y - ic/x) \otimes (y + ic/x)](I - P_x),$$

which is Hermitian for x, y, c, α real. The linearization of these flows will not be discussed in paper II, as it proceeds exactly as in cases (4.46).

Now we consider the Lagrange top, as studied by Ratiu and van Moerbeke [26]; see also Ratiu [25] for n-dimensional generalizations. Formulas (2.13) and (2.15) are equivalent to

$$\frac{d}{dt}(\gamma + hM + h^2(I_1\mu gl)) = [\gamma + hM + h^2(I_1\mu gl), \Omega + (\mu gl)h]. \quad (4.55)$$

Recall from (2.12) that Ω, M, γ and l are the anti-symmetric matrices corresponding in that order to the angular velocity $\Omega = (p, q, r)$, angular momentum $M = (I_1p, I_1q, I_3r)$, coordinates $\gamma = (\gamma_1, \gamma_2, \gamma_3)$ and the spatial z-axis, and the center of gravity $l = z_0(0, 0, 1)$, expressed in the body coordinates. Rescale by replacing $\mu g z_0$ by z_0. As was pointed out before, the fact that the leading term in $\gamma + hM + h^2 I_1 l$ is not diagonal is unimportant. In view of (4.42), we may rewrite

$$A = A_0 + A_1 h + \delta h + \alpha h^2, \quad (4.56)$$

with $A_0 = \gamma$, $A_1 = I_1(p, q, 0)$, $\delta = I_3 r(0, 0, 1)$, $\alpha = I_1 z_0(0, 0, 1)$, and

$$\begin{aligned}\Omega + z_0(0, 0, 1)h &= ((p, q, 0) + z_0(0, 0, 1)h) + r(0, 0, 1)\\ &= I_1^{-1}(Ah^{-1})_+ + [r(1 - I_3 I_1^{-1})(I_1 z_0)^{-1}](Ah^{-2})_+\\ &= (I_1^{-1}/2)[\nabla\langle A^2, h^{-1} + r(1 - I_3 I_1^{-1})\, z_0^{-1} h^{-2}\rangle_1]_+\end{aligned}$$

and thus we have:

THEOREM 4.5. *The Lagrange top equations* (2.13), *and* (2.15) *correspond to a specific flow on the Euclidean orbit* (4.42) *with* $m = 2$, α, δ, A_1 *and* A_0 *defined as* *above and its vector-field is a linear combination of the type occurring in* (4.43):

$$\dot{A} = [A, [\nabla\langle A^2, ah^{-1} + bh^{-2}\rangle_1]_+],$$
$$a = \tfrac{1}{2}I_1^{-1}, \qquad b = \tfrac{1}{2}rI_1^{-1}z_0^{-1}(1 - I_3I_1^{-1}).$$

Note we have implicitly used the fact that r is an orbit invariant.

Remark 5. Theorem 3.1 was employed by Adler [5] in studying the generalized Korteveg–de Vries equation of Gel'fand and Dikii as follows: one takes $L = \{A = \sum_{-\infty<i\leqslant N} a_i\xi^i \mid a_i(x)$ a real periodic $n \times n$, $C^\infty[0, 1]$ matrix function, $N < \infty$, arbitrary$\}$, where the multiplication in L comes from pseudo-differential operator theory, with $\operatorname{tr} A = \int_0^1 sp(a_{-1})\, dx$, $\langle A, B\rangle = \operatorname{tr} A \cdot B$, $K = K^\perp = \{A = \sum_{i\geqslant 0} a_i\xi^i \mid a_i$ as above$\}$, $N = N^\perp = \{A = \sum_{i<0} a_i\xi^i \mid$ etc.$\}$. Therefore K are the differential operators, N the formal Volterra integral operators. Note that $\mathscr{L}$ (Eq. (3.1)) is a subalgebra of L identified with the constant coefficient operators, which gives rise to our application of Theorem 3.1. We also wish to remark that the special cases (4.46) do not seem to relate well to any special feature of $\mathscr{L}$, but just correspond to very great degeneracy in (4.44).

5. Integrability and the Classical Groups

It is natural to view Eqs. (4.44) for the cases (4.46a), (4.46b), and (4.46c) as respectively occurring on the co-adjoint orbits of $U(n)$, $U(n)$ Ⓢ S (where S are the symmetric matrices), and $Gl(n)$, and in fact, the integrability of these systems in the orbit symplectic structure follows from a generalization of a simple argument of Mischenko and Fomenko [9]. The orbit symplectic structures are shown to be the standard Darboux symplectic structures on R^{2n} modulo some simple reduction. In addition, the classical spinning top equations (under gravity) are naturally viewed as occurring in the orbit of $U(n, R)$ Ⓢ $u(n, R)$. We shall first discuss the symplectic structure associated with (4.46).

As a preliminary remark, useful in the future, consider A_0 antisymmetric and the following function $\tilde{H}_\beta(A_0)$ with $\beta = f'(\alpha)$ as in (4.44):

$$\tilde{H}_\beta(A_0) = \langle (f(Ah^{-1}), h\rangle_1 = \tfrac{1}{2}\operatorname{tr}(A_0 \cdot \operatorname{ad}_\beta \operatorname{ad}_\alpha^{-1} A_0)$$
$$= -\sum_{i<j} (A_0)_{ij}^2 \frac{\beta_i - \beta_j}{\alpha_1 - \alpha_j}, \tag{5.1}$$

and define ∇H by $dH = \operatorname{Tr}(\nabla H \cdot dA_0)$, with ∇H skew-symmetric; then $\nabla\tilde{H}_\beta = \operatorname{ad}_\beta \operatorname{ad}_\alpha^{-1} A_0$. Equation (4.44) with $m = 1$, A_0 skew-symmetric is thus equivalent to

$$\dot{A}_0 = [A_0, \nabla\tilde{H}_\beta], \tag{5.2}$$

which may be interpreted as a Hamiltonian equation (with Hamiltonian $\tilde{H}_\beta$) on the co-adjoint orbit of the orthogonal group. This was done by Dikii [10]. If we set $\alpha = J^2$, $\beta = J^k$, then (5.1) is the kth Mischenko integral [10], and (5.2) its associated Hamiltonian flow. This motivates what follows.

We begin by discussing (4.46a). Consider the orthogonal group $U(n)$, its Lie algebra $u(n)$ paired with $u(n)^*$ via $\langle k_1, k_2\rangle = \mathrm{Tr}(k_1 \cdot k_2)$. The co-adjoint action of $U(n)$ through $k = x \wedge y \equiv x \otimes y - y \otimes x$ ($= \Gamma_{xy}$ of Section 4) $\in u(n)^*$ has the form

$$U(x \wedge y)U^{-1} = Ux \wedge Uy = \tilde{x} \wedge \tilde{y}, \qquad Ux = \tilde{x},\ Uy = \tilde{y},\ x, y \in R^n,$$

so that $|\tilde{x}| = |x|$, $|\tilde{y}| = |y|$, and $\langle\tilde{x}, \tilde{y}\rangle = \langle x, y\rangle$. Observe that $x \wedge y = x' \wedge y'$ implies $\mathrm{span}(x, y) = \mathrm{range}\ x \wedge y = \mathrm{range}\ x' \wedge y' = \mathrm{span}(x', y')$; in particular $(x', y') = (ax + by, cx + dy)$ and so $x' \wedge y' = (ad - bc)x \wedge y = x \wedge y$ if and only if $ad - bc = 1$. Hence the set of all $x \wedge y$ is a quotient of $\{(x, y) \mid x, y \in \mathbb{R}^n$ and linearly independent$\}$ by $Sl(2)$. Consequently the dimension of the space of matrices $x \wedge y$ with $x, y \in \mathbb{R}^n$ is $2n - 3$. It is easy to see that, for an appropriate choice of the $Sl(2)$ matrix, $|x| = 1$ and $\langle x, y\rangle = 0$ can always be achieved; $Sl(2)$ acts now on $\mathbb{R}^{2n}$ with $|x| = 1$ and $\langle x, y\rangle = 0$ as follows:

$$(x', y') = (x \cos\varphi - y/|y| \sin\varphi,\ |y|\, x \sin\varphi + y \cos\varphi).$$

Let $\sim$ denote the latter action. Then the orbit $\mathcal{O}_{xy}$ through $x \wedge y$ takes on the form

$$\mathcal{O}_{xy} = \{x' \wedge y' \mid |x'| = 1, \langle x', y'\rangle = 0, |y'| = |y|\}/\sim$$

and is $2n - 4$ dimensional.

On this orbit we have the natural orbit symplectic structure; given a function H defined on this orbit $\mathcal{O}_{xy}$, it leads to the Hamiltonian vector-field and equation

$$(x \wedge y)^{\cdot} = X_H(x \wedge y) \equiv [x \wedge y, \nabla H],$$

or what is the same thing

$$\dot{x} = -\nabla H \cdot x \qquad \text{and} \qquad \dot{y} = -\nabla H \cdot y. \tag{5.3}$$

On the one hand

$$dH = \mathrm{Tr}(\nabla H \cdot d(x \wedge y)) = -2\langle dx, (\nabla H \cdot y)\rangle + 2\langle dy, (\nabla H \cdot x)\rangle,$$

and on the other hand, since H can also be regarded as a function of x and y

$$dH = \langle dx, H_x\rangle + \langle dy, H_y\rangle;$$

by comparison $H_x = -2\nabla H \cdot y$ and $H_y = 2\nabla H \cdot x$ and so the orbit Hamilton flow (5.3) implies the flow

$$\dot{x} = -\tfrac{1}{2}H_y \qquad \text{and} \qquad \dot{y} = \tfrac{1}{2}H_x .$$

We conclude that, up to a factor $-\frac{1}{2}$, the $U(n)$-symplectic structure coincides with the standard Darboux structure on $\mathbb{R}^{2n}$ reduced by the $Sl(2)$-action above. Putting $H = \tilde{H}_\beta(x \wedge y) = -2H_\beta(x, y)$ (of the form (2.25)), we find the flow (2.26) as a consequence of (5.2).

In order to recover the customary Darboux symplectic structure for the case (4.46b), the role played by $U(n)$ in the previous case will now be played by the semi-direct product $G \equiv U(n) \circledS S$, where S denotes the set of symmetric matrices; let K denote the skew-symmetric matrices. Then G is a group with the composition rule $(U_1, S_1) \cdot (U_2, S_2) = (U_1U_2, S_1 + U_1S_2U^{-1})$; its Lie algebra $g \equiv (K, S)$ has for Lie bracket[10]

$$[(k_1, s_1), (k_2, s_2)]_g = ([k_1, k_2], [s_1, k_2] + [k_1, s_2])$$

and for pairing we take

$$\langle (k_1, s_1), (k_2, s_2)\rangle = \operatorname{tr}((k_1 + s_1) \cdot (k_2 + s_2)) = \operatorname{tr}(k_1k_2) + \operatorname{tr}(s_1s_2).$$

The latter implies the bracket between g and g^* ($\approx g$) is

$$[(k_1, s_1), (k_2, s_2)]_{g^*} = ([k_1, k_2] + [s_1, s_2], [s_1, k_2]).$$

Since g^* shall be identified with g via $\langle\, ,\, \rangle$, if H is a function on g^*, its gradient $\nabla H \in g$ admits an orthogonal decomposition $\nabla H = (\nabla_K H, \nabla_S H)$, since $K \perp S$. The Kostant–Kirillov Poisson bracket on g^* is given by

$$\begin{aligned} X_H(F) &= \{F, H\}(k, s) \equiv \langle (k, s), [\nabla H, \nabla F]_g\rangle \\ &= \langle k, [\nabla_K H, \nabla_K F]\rangle + \langle s, ([\nabla_K H, \nabla_S F] + [\nabla_S H, \nabla_K F])\rangle, \end{aligned} \tag{5.4}$$

and therefore, according to the usual recipe, the Hamiltonian vector-field applied to $(k, s) \in g^*$ satisfies the Lax equation

$$(k, s)^{\cdot} = [(k, s), (\nabla_K H, \nabla_S H)]_{g^*} = ([k, \nabla_K H] + [s, \nabla_S H], [s, \nabla_K H])$$

breaking up into the equations

$$\dot{k} = [k, \nabla_K H] + [s, \nabla_S H] \qquad \text{and} \qquad \dot{s} = [s, \nabla_K H]. \tag{5.5}$$

[10] [,] denotes the usual matrix bracket.

Moreover if $(k, s) = (x \wedge y, \ -x \otimes x)$, then these equations preserve their form. To see this, we compute the co-adjoint orbit of G through $(k, s) = (x \wedge y, -x \otimes x)$. The adjoint and co-adjoint actions of G on (k, s) have the following form:

$$\mathrm{Ad}(U, S)(k, s) = (UkU^{-1}, UsU^{-1} + [S, UkU^{-1}])$$

and

$$\mathrm{Ad}^*(U, S)(k, s) = (U^{-1}kU - U^{-1}[S, s]U, U^{-1}sU).$$

In particular

$$\begin{aligned} \mathrm{Ad}^*(U, S)(x \wedge y, -x \otimes x) &= (U^{-1}x \wedge (U^{-1}y - U^{-1}Sx), -U_x^{-1} \otimes U_x^{-1}) \\ &\equiv (\tilde{x}' \wedge \tilde{y}', -\tilde{x}' \otimes \tilde{x}'). \end{aligned}$$

Furthermore observe that $(x \wedge y, -x \otimes x) = (x' \wedge y', -x' \otimes x')$ if and only if $(x', y') = (\pm x, \pm y + tx)$ with $t \in \mathbb{R}$. Hence for an appropriate choice of $t \in \mathbb{R}$, the relation $\langle x' \ y' \rangle = 0$ can always be achieved. This combined with the Ad*-action above shows that the co-adjoint orbit of $(x \wedge y, \ -x \otimes x)$ under G has the form

$$\mathcal{O}_{xy} = \{(x' \wedge y', -x' \otimes x') \quad \text{with} \quad |x'| = |x| \ \text{ and } \ \langle x', y' \rangle = 0\}.$$

Since (5.5) are equations on $\mathcal{O}_{xy}$, they preserve the special form of (k, s). Defining $(V, W) \equiv (\nabla_K H, \nabla_S H) = \nabla H$, these equations reduce to

$$\dot{x} = -Vx \qquad \text{and} \qquad \dot{y} = -Vy - Wx. \tag{5.6}$$

Finally we show that the customary Darboux symplectic structure on the orbit $\mathcal{O}_{xy} \approx T^*S^{n-1}$ is equivalent to the one defined by (5.4). The functions $H = H(k, s)$ defined on this orbit can also be regarded as functions $H = H(x, y)$. To see the relation between their gradients, we observe on the one hand that

$$dH = \langle dx, Hx \rangle + \langle dy, Hy \rangle, \qquad (x, y) \in \mathcal{O}_{xy}, \tag{5.7}$$

and on the other hand that

$$\begin{aligned} dH &= \langle (dk, ds), (\nabla_K H, \nabla_S H) \rangle, \qquad k = x \wedge y, s = -x \otimes x, \\ &= \langle dx, (-2Vy - 2Wx) \rangle + \langle dy, 2Vx \rangle. \end{aligned} \tag{5.8}$$

Comparing (5.7) and (5.8), we find $H_x = -2Wx - 2Vy$ and $H_y = 2Vx$; comparing this with (5.6), we conclude that $\dot{x} = -\frac{1}{2}H_y$ and $\dot{y} = -\frac{1}{2}H_x$, establishing the equivalence between the two symplectic structures, up to a factor $-\frac{1}{2}$. It is interesting to remark that the flow with Hamiltonian $H(k, s) = \tilde{H}_\beta(k) + \mathrm{Tr}(\beta \cdot s)$ (see (5.1)) yields (4.51). If one evaluates (5.5) at a general point (not necessarily of the form $(k, s) = (x \wedge y, -x \otimes x)$) with the same $H(k, s)$ one finds (4.44) with $A = \alpha h^2 + kh + s$. This equation can be named, via Arnold's procedure [1] as the equation of degenerate geodesic motion on T^*G, with an additional linear forcing term $-\mathrm{Tr}(\beta \cdot s)$.

For case (4.46c), the relevant group is the invertible matrices $\hat{G}$, with Lie algebra $\hat{g} = g$ as before with the previous $\langle\, , \rangle$. Here of course the Lie bracket is the usual bracket

$$[(k_1 + s_1), (k_2 + s_2)]_{\hat{g}} = [k_1, k_2] + [k_1, s_2] + [s_1, k_2] + [s_1, s_2];$$

so we may drop the subscript. From the above, the co-adjoint Poisson bracket is

$$\{F, H\} = \langle k, [\nabla_K H, \nabla_K F] + [\nabla_S H, \nabla_S F]\rangle + \langle s, [\nabla_K H, \nabla_S F] + [\nabla_S H, \nabla_K F]\rangle, \quad (5.9)$$

and so Hamilton's equations have the Lax form $(k, s)^\cdot = [(k, s), (\nabla_K H, \nabla_S H)]$ in $\hat{g}^*$ ($\approx \hat{g}$), i.e.,

$$\dot{s} = [s, \nabla_K H] + [k, \nabla_S H], \qquad \dot{k} = [k, \nabla_K H] + [s, \nabla_S H]. \quad (5.10)$$

Note for $H = G(k, s) = \tilde{H}_\beta(K) + \mathrm{Tr}(\beta \cdot s)$, (5.10) reduces to (4.52) at $(k, s) = (x \wedge y, \Delta_{xy})$, and $G(k, s)$ becomes $-4G_\beta(y/2, x)$. Also note for $H = G(k, s)$ in (5.10), we get a system equivalent to (4.44), with $A = (\alpha h^2 + kh + s - \alpha)$.

One shows as before that the $\hat{G}$ co-adjoint orbit through $(\Gamma_{xy}, \Delta_{xy})$ is symplectically equivalent to the standard $(x, y) \in R^{2n}$ structure, reduced by the symplectic group action $(x, y) \mapsto (ax, a^{-1}y)$ (again modulo a fact of $-\frac{1}{2}$). Indeed, observe that since $\frac{1}{2}(\Delta_{xy} \pm x \wedge y) = x \otimes y,\ y \otimes x$, respectively, $(x \wedge y, \Delta_{xy}) = (x' \wedge y', \Delta_{x'y'})$ precisely if (x, y) is related to (x', y') via the action $(x, y) \to (ax, a^{-1}y)$, with $a \in \mathbb{R}\backslash\{0\}$. Thus the pairs of matrices of the above form are represented by pairs of the form (x, y), such that $|x| = 1$. Moreover, since the orbit $\mathcal{O}_{xy}$ through $(x \wedge y, \Delta_{xy})$ are matrices of the form $U(2x \otimes y)U^{-1} = 2(Ux) \otimes ((U^{-1})^T y)$, with $U \in \hat{G}$ only the spectral invariant $\langle x, y\rangle$ is preserved. This is just the determinant of $x \otimes y$ restricted to being an operator with range and domain spanned by the range of $x \otimes y$. Hence the orbit $\mathcal{O}_{xy} = \{(x \wedge y, \Delta_{xy}) \mid |x| = 1, \langle x, y\rangle = \text{constant}\}$, a $(2n - 2)$-dimensional variety. The computation of the symplectic structure proceeds as before, and thus $\mathcal{O}_{xy}$ is the above-describted symplectic manifold.

It is worthwhile to give yet another abstract proof of the integrability of these systems in the setting of the three Lie algebras we have just introduced. We first do the spinning top case.

The co-adjoint orbit symplectic structure on $(gl(n))^*$ is given by the Poisson bracket $\{F, G\}(A_0) = \langle A_0, [\nabla F, \nabla G]\rangle$, for functions F, G on $gl(n)^*$ ($\approx gl(n)$), with A_0 the running variable on $gl(n)$, and as usual, $dH \equiv \langle \nabla H, dA_0\rangle$ defining ∇H. Observe that if $F(A_0)$ and $G(A_0)$ are orbit invariants, i.e., $[\nabla F(A_0), A_0] = 0$, etc., for G, then if $F_h \equiv F(A(h))$, $A(h) \equiv \alpha h + A_0$, etc., for G, $\{F_h, G_t\} = 0$ is any identity in h, t. To see that, observe $[\nabla F(A), A] = 0$, etc., for G, and so using the identity $a(A_0 + \alpha h) + b(A_0 + \alpha t) = A_0$ for $a = t(t-h)^{-1}$, $b = h(h-t)^{-1}$, we compute $\{F_h, G_t\} = \langle aA(h) + bA(t), [\nabla F(A(h)), \nabla G(A(t))]\rangle = a\langle [A(h), \nabla F(A(h))], \nabla G(A(t))\rangle - b\langle [A(t), \nabla G(A(t))], \nabla F(A(h))]\rangle = 0$. Hence,

$$\{F_h, G_t\} = 0; \qquad \text{i.e.,} \qquad F_h, G_t \text{ are in involution.} \tag{5.11}$$

This argument is due to Mischenko and Fomenko [9].

If F, G are of the form $F = F_0^{(j)}(A_0) = \operatorname{tr} A_0{}^j$, etc., for G, $F(A_0 + h\alpha) = \sum c_{jk}h^k$, and since from (5.11), $\{F_h^{(j)}, F_t^{(k)}\} = 0$ is an identity in h, t, we conclude $\{\nabla c_{jk}, \nabla c_{j'k'}\} = 0$. If furthermore we are at a point A_0 such that $A_0 + A_0{}^T = 0$, i.e., A_0 is skew-symmetric, and $j - k$, $j' - k'$ are even, then ∇c_{jk}, $\nabla c_{j'k'}$ are also skew-symmetric. To see this, observe $dF_h^{(j)} = j\langle (A_0 + h\alpha)^{j-1}, dA_0\rangle$, so $\nabla F_h^{(j)} = j(A_0 + h\alpha)^{j-1} = \sum \nabla c_{jk}h^k$, but since for h pure imaginary, $(A_0 + h\alpha)^{j-1}$ is Hermitian, or anti-Hermitian if j is even or odd respectively, identically in h, the observation follows. In fact in that case, $\nabla c_{jk} = \nabla_K c_{jk}$, which is the gradient of c_{jk} regarded as a function on K, the skew-symmetric matrices, by restriction; it being computed via the usual rule $dH = (\nabla H, dA_0)$, $\nabla H + \nabla H^T = 0$. From the above, $0 = \{c_{jk}, c_{j'k'}\}(A_0) = \langle A_0, [\nabla c_{jk}, \nabla c_{j'k'}]\rangle = \langle A_0, [\nabla_K c_{jk}, \nabla_K c_{j'k'}]\rangle = \{c_{jk}, c_{j'k'}\}|_K (A_0)$, and so $\{c_{jk}, c_{j'k'}\}|_K(A_0) = 0$, where the last bracket is just the co-adjoint orbit bracket for $U(n)$. The algebra of functions formed by the coefficients of the F_j's having the property that upon restriction to K, their gradients lie in K, form an involutive system of integrals on the co-adjoint $U(n)$ orbits. The $\tilde{H}_\beta(K)$ of (5.1) are examples of such functions.

This argument in fact generalizes to the other cases, which is computationally surprising. For the case of the semi-direct product, we first consider (5.4), and as a preliminary, as in the previous case, we work with $GL(n) ⓢ gl(n)$, *imagining* in the Poisson bracket (5.4) that $(K, S) = (gl(n), gl(n))$, with running variables (k, s). Then if F, G are the functions considered in the previous case, define $F_h \equiv F(A(h))$, $A(h) \equiv \alpha h^2 + hk + s$, etc., for G. Then $\nabla_K F_h(k, s) = h\nabla F(A(h))$, $\nabla_S F_h(k, s) = \nabla F(A(h))$, with ∇F the previously computed $gl(n)$ gradient, etc., for G. Using (5.4) and the above formula, we find $\{F_h, G_h\}(k, s) = \langle htk + (h+t)s, [\nabla F(A(h)), \nabla G(A(t))]\rangle$, which by the previous argument equals zero, provided we can write $htk + (h+t)s = a(\alpha h^2 + kh + s) + b(\alpha t^2 + tk + s)$ for some a, b. Simply take $a = t^2(t-h)^{-1}$, $b = h^2(h-t)^{-1}$ to satisfy the *three* identities. To get our commuting functions on $U(n)$ ⓢ (symmetric matrices) co-adjoint orbits, take in the algebra formed by the c_{jk}, those functions f such that $(\nabla_K f, \nabla_S f)|_{(K,S)}$ (skew-symmetric,

symmetric) = (skew-symmetric, symmetric). The $F(k, s) = \tilde{H}_\beta(k) + \mathrm{Tr}\langle\beta \cdot s\rangle$ are an example of such functions.

For the last case of the real invertible matrices, we again "imagine" in the Poisson bracket (5.9) that $(K, S) = (gl(n), gl(n))$. Working with the same functions F, G, define $F_h \equiv F(A(h))$, $A(h) = \alpha h^2 + kh + s - \alpha$, etc., for G. To show $\{F_h, G_t\} = 0$, we proceed as in the previous case, requiring that $k(ht + 1) + s(h + t) = aA(h) + bA(t)$, and so set $a = (1 - t^2)(h - t)^{-1}$, $b = (1 - h^2)(t - h)^{-1}$. The arguments then proceed as before. The Lie algebra generalization of this discussion just uses the symmetric decomposition. We note, for $(k, s) = (x \wedge y, \Delta_{xy})$, that previous arguments now show that the G_i's of (2.20) are in involution with respect to a reduced (x, y) Hamiltonian structure, and upon subjecting (x, y) to the previously mentioned $SL(2)$ action, we see the same is true for the J_i's of Remark 2, Section 4.

To sum up, the Hamiltonian structures of (4.46a), (4.46b), and (4.46c) (in that order) in (x, y) coordinates, at the matrix level, are seen to correspond respectively to extremely low dimensional Kostant–Kirillov co-adjoint orbit structures on $U(n)$, the semi-direct product of $U(n)$ with the symmetric matrices, and the invertible matrices, and in fact, from the above, are specializations of integrable systems on more general co-adjoint orbits.

We now discuss the spinning top (2.13) from this point of view. Equations (2.13) have the form

$$\dot{M} = [M, \Omega] + [\gamma, l'], \qquad \dot{\gamma} = [\gamma, \Omega], \qquad M_{ij} = I_{ij}\Omega_{ij}, \tag{5.12}$$

with parameters I_{ij}, $l' = \mu gl$. Glancing at (5.5), we see that if we regard $(M, \gamma) \in u(3) \circledS u(3) \equiv g \approx (u(3) \circledS u(3))^* = g^*$ as an element of the co-adjoint orbit of $U(3) \circledS u(3) = G$, then (5.12) is just an equation on the orbit of G, where (M, γ) replace the previous running parameters (k, s).

To get (5.12) from (5.5), take as the Hamiltonian

$$H = \mathrm{tr}(\tfrac{1}{2}M \cdot \Omega + l' \cdot \gamma);$$

thus $\nabla_M H = \Omega$, $\nabla_\gamma H = l'$. Note the general Hamiltonian equations

$$\dot{M} = [M, \nabla_M H] + [\gamma, \nabla_\gamma H], \qquad \dot{\gamma} = [\gamma, \nabla_M H]$$

have the orbit invariants tr $\gamma \cdot \gamma$, tr $M \cdot \gamma$, which physically correspond respectively to the invariance of the length of the z-axis, and the internal angular momentum along the z-axis—the direction of the gravitional force. Hence since g^* has dimension 6, and since there are typically two orbit invariants, the orbit phase space has dimension 4; therefore, from the above point of view, one has always needed one integral in addition to the Hamiltonian H (and the orbit invariants) to solve the spinning top equations by quadrature, as in the celebrated cases of Lagrange and S. Kovalevsky. These considerations can be generalized to the case $n > 3$. For more details, see Ratiu [25] and Ratiu and van Moerbeke [26].

REFERENCES

1. V. I. ARNOLD, "Mathematical Methods of Classical Mechanics," Springer-Verlag, New York, 1978.
2. D. KAZHDAN, B. KOSTANT, AND S. STERNBERG, "Hamiltonian Group Actions and Dynamical Systems of Calogero Type," CPAM, July 1978.
3. P. VAN MOERBEKE AND D. MUMFORD, The spectrum of difference operators and algebraic curves, *Acta Math.* (1979).
4. B. A. DUBROVIN, V. B. MATVEEV, AND S. P. NOVIKOV, *Uspehi Mat. Nauk.* **31** No. 1 (187) (1976), 55; *Russian Math. Surveys* **31**, No. 1 (1976).
5. M. ADLER, On a trace functional for pseudo-differential operators and the symplectic structure of the Korteweg–deVries Equation, *Invent. Math.* **50**, No. 3 (1979).
6. P. VAN MOERBEKE, The spectrum of Jacobi matrices, *Invent. Math.* **37** (1976).
7. J. MOSER, Geometry of quadrics and spectral theory, preprint.
8. B. KOSTANT, The solution to a generalized Toda lattice and representation theory, *Advances in Math.* **34** (1979), 195–338.
9. A. S. MISHCHENKO AND A. T. FOMENKO, Euler's equations on finite-dimensional Lie groups, *Izv. Akad. Nauk. SSSR Ser. Mat.* **42** (1978).
10. L. A. DIKII, Hamiltional systems connected with the rotation group, *Funksional. Anal. i Prilozen.* **6**, No. 4 (1972), 83–84.
11. S. V. MANAKOV, Remarks on the integrals of the Euler equations of the n-dimensional heavy top, *Functional Anal. Appl.* **10** (4), (1976).
12. J. MOSER, Various aspects of integrable Hamiltonian systems, *in* "Proceedings, CIME Conference, Bressanone, Italy, June 1978," in press.
13. V. V. GOLUBEV, "Lectures on Integration of the Equations of Motion of a Rigid Body About a Fixed Point," Published for the NSF by the Israel Program for Scientific Translations, Haifa, Israel, 1960.
14. C. CHEVALLEY, "Theory of Lie Groups," Princeton Univ. Press, Princeton, N.J., 1946.
15. E ARTIN, "Galois Theory," Notre Dame Mathematical Lectures, No. 2.
16. W. SYMES, Systems of Toda type, inverse spectral problems, and representation theory, *Invent. Math.* **59**, No. 1 (1980), 13–53.
17. R. MOODY, A new class of Lie algebras, *J. Algebra* **10** (1968), 211–230.
18. R. GILMORE, "Lie Groups, Lie Algebras, and Some of Their Applications," Wiley, New York, 1974.
19. J. HUMPHREYS, "Introduction to Lie Algebras and Representation Theory," Springer-Verlag, New York/Berlin, 1972.
20. H. FLASCHKA, On the Toda lattice, I, *Phys. Rev. B* **9** (1974).
21. O. I. BOGOYAVLENSKY, On perturbations of the periodic Toda lattices, *Comm. Math. Phys.*, **51** (1976).
22. J. COOLIDGE, "A Treatise on Algebraic Plane Curves," p. 321, Oxford Univ. Press (Clarendon), London, 1931.
23. M. ADLER AND P. VAN MOERBEKE, Linearization of Hamiltonian systems, Jacobi varieties, and representation theory, *Advances in Math.* **38** (1980), 318–379.
24. S. STERNBERG AND A. JACOB, Co-adjoint structures, solitons and integrability, preprint.
25. T. RATIU, recent preprint.

26. T. RATIU AND P. VAN MOERBEKE, The Lagrange top and Kac–Moody Lie algebras, to appear.
27. D. R. LEBEDEV. AND YU. I., MANIN, preprint ITEP-155 (1978).
28. G. WILSON, Commuting flows and conservation laws for Lax equations, *Math. Proc. Cambridge Philos. Soc.* **86** (1979), 131.
29. S. HELGASON, "Differential Geometry, Lie Groups and Symmetric Spaces," p. 491, Academic Press, New York, 1978.
30. H. KNÖRRER, Geodesics on the ellipsoid, *Invent. Math.* **59**, No. 2 (1980), 119–144.
31. I. FRENKEL, A. REIMAN, AND M. SEMENOV-TIAN-SHANSKY, Graded Lie algebras and completely integrable dynamical systems, *Soviet Math. Dokl.* **20**, No. 4 (1979), 811–814.
32. A. REIMAN AND M. SEMENOV-TIAN-SHANSKY, Reduction of Hamiltonian systems, affine Lie algebras and Lax equations, *Invent. Math.* **54**, No. 1 (1979), 81–101.