Advanced Studies in Pure Mathematics 22, 1993
Progress in Differential Geometry
pp. 213–229

Lax Equations Associated with a Least Squares Problem and Compact Lie Algebras

Yoshimasa Nakamura

Abstract.

The gradient flow in a least squares problem on a Lie group takes a Lax form [8]. We associate the Lax equation with homogeneous spaces and symmetric spaces of compact simple Lie groups. The critical points of the Lax equation lie in the Cartan subalgebras of the simple Lie algebras. A reduction from homogeneous spaces to symmetric spaces is described by a 'coalescence' of roots. For the complex Grassmann manifold, it is shown that an initial value problem of the Lax equation can be uniquely solved. Some applications to a least squares fitting problem and a linear programming problem are discussed.

§1. Introduction

In the recent work [8] Brockett studied a critical point problem for a least squares function defined on the space of orthogonal matrices $\mathbf{SO}(n)$ in terms of a gradient flow. He derived the nonlinear ordinary differential equation of Lax type

$$\frac{dL(t)}{dt} = [L(t),\ [N, L(t)]]$$

by projecting the gradient flow on $\mathbf{SO}(n)$ onto the adjoint orbit of $\mathbf{SO}(n)$ via $L(t) = g^{-1}(t)Qg(t)$. Here $g(t) \in \mathbf{SO}(n)$, Q is symmetric and N is also a symmetric matrix with distinct eigenvalues. The origin of such a least squares problem itself goes back to an old result by von Neumann [19]. It is shown in [8] that the Lax equation provides a method for solving the eigenvalue problem of the symmetric matrix Q. The solution $L(t)$ converges as $t \to \infty$ to a diagonal matrix being some permutation of the eigenvalues of Q. This feature of Lax type equation is very similar to

Received March 11, 1991.

that of finite nonperiodic Toda equation found by Moser [16] and further analyzed by Deift, Nanda and Tomei [10]. See also the expository paper by Watkins [22]. If both Q and N have distinct eigenvalues, then there are $n!$ critical points on $\mathbf{SO}(n)$ and the least squares problem can solve a combinatorial optimization problem [8]. Thus the above property of the Lax equation turns the optimization problem into a problem in the theory of ODEs.

More recently Bloch [4] asserted that the Lax equation, where $L(t)$ and N are skew Hermitian, could be derived from a Hamiltonian flow on the complex Grassmann manifold $\mathbf{SU}(p+q)/\mathbf{S}(\mathbf{U}(p)\times\mathbf{U}(q))$. This makes crucial use of the (almost) complex structure. Here the Hamiltonian is a least squares fitting function in estimation theory [3], which measures the total perpendicular squares distance of m-points sited in $\mathbb{C}^{p+q}$ onto the q-plane. Hence the least squares fitting problem may be solved by investigating critical points of the Lax equation associated with the complex Grassmann manifold. The usual fitting problem is formulated in $\mathbb{R}^{p+q}$ and is related to the real Grassmann manifold, so that the method in [4] does not apply. However, the observation in [4] allows us to introduce ideas from Lie algebras and symmetric spaces and to carry out a generalization and classification of the Lax equation by setting $L(t)$ to lie on some homogeneous and symmetric spaces.

Bloch, Brockett and Ratiu [5] discussed a generalization of the Lax equation to any semisimple Lie algebra $\mathfrak{g}$. They showed that, by a suitable choice of $L(t)$ and N, the Lax equation is reduced to the generalized finite nonperiodic Toda equations studied by many authors (see, for example, [15]). They considered a decomposition of $\mathfrak{g}$ into the direct sum of the centralizer of N and its vector space complement in $\mathfrak{g}$. It is known [12] that the complement is identified with the tangent space of certain homogeneous space. The diagonal matrix N defines the isotropy subgroup of the corresponding Lie group G, say H, which gives rise to the homogeneous space G/H. Here $L(t) = g^{-1}(t)Qg(t)$ can be viewed as an element of an adjoint orbit through Q of G. Thus it has not been clear how we can regard $L(t)$ as representing a point of the homogeneous space G/H. Moreover, since the Lax equation is the projection of a gradient flow on G into its adjoint orbit, it should be checked that one can reconstruct the gradient flow on G from any solution of the Lax equation. It is also an important problem to propose a method for solving the Lax equation itself.

In this paper we consider another type of decomposition of any simple Lie algebra $\mathfrak{g}$, namely, a decomposition into the centralizer $\mathfrak{k}$ of Q (*not* N) and its complement $\mathfrak{m}$ in $\mathfrak{g}$. This enables us not only to associate the Lax equation with homogeneous spaces G/K but to give an explicit description of reduction of the Lax equation from homogeneous spaces to symmetric spaces. Here K is the isotropy subgroup having the Lie algebra $\mathfrak{k}$ and $L(t)$ can be identified with a point of G/K. Note that (irreducible) symmetric spaces are completely classified in terms of simple Lie algebras [12]. The equivalence between the gradient flow on any simple Lie group G and the Lax equation associated with the homogeneous space G/K is proved in §3. We also see that 'coalescing' of eigenvalues of $\operatorname{ad} Q$ (or, roots of Lie algebras) gives rise to a sequence of reductions in §4. The Lax equations in least squares fitting problem are derived without using the complex structure as Lax equations associated with real and complex Grassmann manifolds. A class of linear programming problems which can be regarded as a generalization of that of Brockett [8] is also discussed. The Lax equation describes an interior flow on a convex polytope. For the complex Grassmann manifolds it is shown in §5 that an initial value problem of the Lax equation can be uniquely solved by a decomposition of the exponential of initial value. This provides a new approach to such linear programming problems in terms of nonlinear ODEs.

§2. Preliminary

First we review some of the basic facts concerning simple Lie algebras, reductive homogeneous spaces and Hermitian symmetric spaces. More details can be found in the book [12, 14]. This section is based upon the works by Bogoyavlensky [7], Fordy and Kulish [11] and this author [17] which discuss a generalization and classification of the Toda equation, the nonlinear Schrödinger equation and the Heisenberg model, respectively.

Let G be a simple Lie group and $\mathfrak{g}$ be its Lie algebra. Let M be a homogeneous space of G, namely, M is a differentiable manifold on which G acts transitively. It is known [12, p.121] that there is a homeomorphism of the coset space G/K onto M for some isotropy subgroup K of G at a point of M. Let $\mathfrak{k}$ be the Lie algebra of K. We consider the decomposition

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{m}, \qquad [\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}, \tag{1}$$

where $\mathfrak{m}$ is the vector space complement of $\mathfrak{k}$ in $\mathfrak{k}$. If $\mathfrak{k}$ and $\mathfrak{m}$ satisfy (1) and

$$[\mathfrak{k}, \mathfrak{m}] \subset \mathfrak{m}, \tag{2}$$

then G/K is called a *reductive homogeneous space.* The flag manifold gives a striking example of the space, as we will see in §3. If $\mathfrak{k}$ and $\mathfrak{m}$

satisfy (1), (2) and

$$[\mathfrak{m}, \mathfrak{m}] \subset \mathfrak{k}, \tag{3}$$

then G/K is a *symmetric space.*

Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$ which is the maximal abelian subalgebra of diagonalizable elements of $\mathfrak{g}$. In terms of a Weyl basis [12, p.421] we can see that $\mathfrak{g}$ has the following commutation relations

$$\begin{aligned} [H_i, H_j] &= 0, \qquad [H_i, X_\alpha] = \alpha(H_i)X_\alpha, \\ [X_\alpha, X_\beta] &= N_{\alpha,\beta}X_{\alpha+\beta} \quad (\alpha+\beta \in \Delta), \\ &= \sum_{i=1}^{\operatorname{rank}\mathfrak{h}} C_{\alpha,i}H_i \quad (\alpha+\beta=0), \\ &= 0 \quad (\alpha+\beta \notin \Delta, \alpha+\beta \neq 0) \end{aligned} \tag{4}$$

for any $H_i \in \mathfrak{h}$ and $X_\alpha \in \mathfrak{g} \bmod \mathfrak{h}$, where $N_{\alpha,\beta}$ and $C_{\alpha,i}$ are constants and Δ is a set of nonzero linear functionals $\alpha : \mathfrak{h} \to \mathbb{C}$ called roots. Let Q be an element of $\mathfrak{h}$. In this paper we choose the isotropy K such that its Lie algebra $\mathfrak{k}$ is given by the centralizer $C_\mathfrak{g}(Q)$ of Q in $\mathfrak{g}$,

$$\mathfrak{k} = C_\mathfrak{g}(Q) = \{X \in \mathfrak{g} \mid [X, Q] = 0\}. \tag{5}$$

The existence of such Q is proved in [14, p.261]. Compare $\mathfrak{k}$ in (5) with the subalgebra $C_\mathfrak{g}(N)$ defined in [5]. It is to be noted [12, p.163] that if Q is *regular*, namely, the eigenvalues $\alpha(Q)$ of $\operatorname{ad} Q$ are mutually distinct, then

$$C_\mathfrak{g}(Q) = \mathfrak{h}. \tag{6}$$

From (4) we see $[\mathfrak{h}, \mathfrak{m}] \subset \mathfrak{m}$. Since $\mathfrak{k} = \mathfrak{h}$ in this case, the corresponding coset space G/K is automatically a reductive homogeneous space. The decomposition (1) is essentially a Cartan decomposition of $\mathfrak{g}$. If Q is not regular, then $\mathfrak{k} = C_\mathfrak{g}(Q) \supset \mathfrak{h}$. This implies that as the eigenvalues $\alpha(Q)$ 'coalesce', $C_\mathfrak{g}(Q)$ becomes larger, and consequently, the homogeneous space G/K becomes smaller. Hence coalescing of $\alpha(Q)$ can give rise to a sequence of reductions from homogeneous spaces to symmetric spaces.

Consider a special but important class of symmetric spaces called (irreducible) *Hermitian symmetric spaces.* In this case the eigenvalues $\alpha(Q)$ have only three distinct values $\{0, \pm\alpha\}$. We can decompose $\mathfrak{g}$ into three

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{m}^+ \oplus \mathfrak{m}^-. \tag{7}$$

Here if we set $X = X^0 + X^+ + X^-$ with $X^0 \in \mathfrak{k}$ and $X^\pm \in \mathfrak{m}^\pm$ for any $X \in \mathfrak{g}$, then $[Q, X^0] = 0$, $[Q, X^\pm] = \pm\alpha X^\pm$. Namely, the eigenvalues $\alpha(Q)$ take the same value for all $X^\pm \in \mathfrak{m}^\pm$. The second property of the Hermitian symmetric space to be noted here is the existence of the (almost) complex structure [12, p.391]. By multiplying Q by a suitable nonzero constant we have a linear endomorphism $\operatorname{ad} Q : \mathfrak{m} \to \mathfrak{m}$ satisfying

$$(\operatorname{ad} Q)^2 = -1. \tag{8}$$

There are 6-types of Hermitian symmetric spaces [12, p.518]. One of them denoted by **AIII** is the complex Grassmann manifold which will be discussed in §4 and §5.

§3. Lax equations and reductive homogeneous spaces

Let N be a fixed regular element of the Cartan subalgebra $\mathfrak{h}$ of any simple Lie algebra $\mathfrak{g}$. Define the function $f(g)$ on G

$$f(g) = \operatorname{tr}(g^{-1}QgN), \tag{9}$$

where Q is a fixed element of $\mathfrak{h}$ and $g \in G$. Following Brockett [8] we consider the critical point problem for $f(g)$ by investigating the gradient flow on G. It is not hard to derive the gradient flow

$$\frac{dg}{dt} = gNg^{-1}Qg - QgN, \tag{10}$$

where g is a smooth function of $t \in \mathbb{R}$ with value in G. Define a (matrix) group action on G, $\gamma : K \times G \to G$ by $\gamma(k, g) = kg$. Since $f(g)$ is invariant under γ, the gradient flow on G can be projected onto the coset space G/K. We introduce a point on adjoint orbit of G through Q by

$$L(t) = g^{-1}(t)Qg(t) \in \mathfrak{g}. \tag{11}$$

Brockett [8] found that if $g(t)$ satisfy (10), then $L(t)$ holds

$$\frac{dL(t)}{dt} = [L(t),\ [N, L(t)]]. \tag{12}$$

Following [1, p.59] we say that the gradient flow (10) admits the *Lax representation* (12). Noting $[N, L(t)] \in \mathfrak{g}$ we see that when $L(t)$ changes in time, $L(t)$ remains on the adjoint orbit of $\mathfrak{g}$. Hence the invariants of the orbit (for example, eigenvalues of $L(t)$) are first integrals of (10). It is known [5] that (12) is actually the gradient flow on the adjoint orbit (11) endowed with the standard metric

$$ds^2 =< (\mathrm{ad}\, L)^{-1}dL, (\mathrm{ad}\, L)^{-1}dL >,$$

where the bracket is defined by the Killing form. See also [6] for the Grassmannian case. The following proposition shows that we can reconstruct a solution of (10) from any solution of (12) if G/K is reductive.

Proposition 1. *Let G/K be a reductive homogeneous space. If $L(t) = g^{-1}(t)Qg(t)$ satisfies the Lax equation* (12), *then $kg(t)$ satisfies the gradient flow* (10) *for some $k \in K$.*

Proof. Differentiating $L(t) = g^{-1}(t)Qg(t)$ and using (12) we have

$$[\frac{dg}{dt}g^{-1} - [gNg^{-1}, Q],\ Q] = 0.$$

This implies $\frac{dg}{dt}g^{-1} - [gNg^{-1}, Q] \in \mathfrak{k}$. On the other hand, from $[\mathfrak{k}, \mathfrak{m}] \subset \mathfrak{m}$, we see $[gNg^{-1}, Q] \in \mathfrak{m}$. There is an element k of K such that $\frac{d(kg)}{dt}(kg)^{-1} \in \mathfrak{m}$, and then $\frac{dg'}{dt}g'^{-1} - [g'Ng'^{-1}, Q] \in \mathfrak{m}$, where $g' = kg$. For such k, $g'^{-1}Qg' = L(t)$. Hence we can obtain an element of G from any given $L(t)$ which satisfies the gradient flow (10) on G. Q.E.D.

From the definition, our G/K is automatically a reductive homogeneous space. This fact does not depend on the regularity of Q and enable us to carry out a reduction to symmetric spaces (§4). The matrix N is also permitted to be non-regular in Proposition 1.

It is important to remark the followings. If we consider a decomposition of $\mathfrak{g}$ into the centralizer of N denoted by $C_{\mathfrak{g}}(N)$ and its complement as in [5], then Proposition 1 does not hold unless Q is regular. The Lax equation discussed in [5] is a special case of

$$\frac{dM(t)}{dt} = [M(t),\ [Q, M(t)]] \tag{13}$$

which is derived from (10) by setting

$$M(t) = g(t)Ng^{-1}(t). \tag{14}$$

Here the roles of N and Q are interchanged. From (13) and (14) we have

$$[g^{-1}\frac{dg}{dt} + [g^{-1}Qg, N],\ N] = 0.$$

If both N and Q are regular, then $g^{-1}\frac{dg}{dt} + [g^{-1}Qg, N] \in \mathfrak{k}$. Since G/K is reductive and $N \in \mathfrak{k}$, $[g^{-1}Qg, N] \in \mathfrak{m}$. Note that there is an action $\widetilde{\gamma} : G \times K \to G$, with $\widetilde{\gamma}(g,k) = gk$, which leaves $M(t)$ invariant. We can choose so that $g^{-1}\frac{dg}{dt} \in \mathfrak{m}$ in terms of $\widetilde{\gamma}$. Thus we can derive (10) from (13) providing that both N and Q are regular.

We call (12) the *Lax equation associated with the reductive homogeneous space* G/K. For a while let us suppose that Q is regular as well as N. As was shown in the above (10), (12) and (13) are mutually equivalent. Since $\mathfrak{k}$ and $\mathfrak{m}$ satisfy (1) and (2), the Lax equation (13) is expressed as

$$\text{(15)} \qquad \begin{aligned} \frac{dM_{\mathfrak{k}}}{dt} &= [M_{\mathfrak{m}},\ [Q, M_{\mathfrak{m}}]]_{\mathfrak{k}}, \\ \frac{dM_{\mathfrak{m}}}{dt} &= [M_{\mathfrak{k}},\ [Q, M_{\mathfrak{m}}]] + [M_{\mathfrak{m}},\ [Q, M_{\mathfrak{m}}]]_{\mathfrak{m}}, \end{aligned}$$

where $M(t) = M_{\mathfrak{k}} + M_{\mathfrak{m}}$ and the subscript $\mathfrak{k}$ and $\mathfrak{m}$ refer to the components in the vector subspaces $\mathfrak{k}$ and $\mathfrak{m}$, respectively. The Lax equation in [5] is essentially equivalent to (15). This class of Lax equations is visualized in terms of the flag manifold $G/K = \mathbf{SU}(r)/\mathbf{S}(\mathbf{U}(1) \times \cdots \times \mathbf{U}(1))$, where K is the maximal torus of G. We give a simple example, $\mathbf{SU}(3)/\mathbf{S}(\mathbf{U}(1) \times \mathbf{U}(1) \times \mathbf{U}(1))$. Let us set

$$Q = i \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix},$$

$$M_{\mathfrak{k}} = i \begin{pmatrix} b_1 & 0 & 0 \\ 0 & b_2 & 0 \\ 0 & 0 & b_3 \end{pmatrix}, \qquad M_{\mathfrak{m}} = \begin{pmatrix} 0 & a_1 & a_3 \\ -\bar{a}_1 & 0 & a_2 \\ -\bar{a}_3 & -\bar{a}_2 & 0 \end{pmatrix},$$

where the bar denotes the complex conjugate, $\sum_j \lambda_j = 0$, $\lambda_i \neq \lambda_j$, $\sum_j b_j = 0$ and λ_j, b_j are real. Part of the Lax equation (15) is then

$$\begin{aligned} \frac{da_1}{dt} &= -(\lambda_1 - \lambda_2)a_1(b_1 - b_2) - 3i\lambda_3 a_3 \bar{a}_2, \\ \frac{db_1}{dt} &= 2(\lambda_1 - \lambda_2)|a_1|^2 - 2(\lambda_3 - \lambda_1)|a_3|^2. \end{aligned}$$

If we set $a_3 = 0$ and $\lambda_1 - \lambda_2 = 1$, $\lambda_2 - \lambda_3 = 1$, the Lax equation is just the (complex) finite nonperiodic Toda equation discussed in [10, 16].

In [5] it is also proved that $\lim_{t\to\infty} M(t)$ exists and lies in the Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$. Here the regularity of Q is assumed and then $\mathfrak{h} = \mathfrak{k} = C_{\mathfrak{g}}(Q)$. Here we extend this result to the case where $\operatorname{ad} Q$ is permitted to have repeated eigenvalues, namely, $\mathfrak{h} \subset \mathfrak{k} = C_{\mathfrak{g}}(Q)$.

Proposition 2. *Let* $G = \mathbf{SU}(r)$. *The limits* $\lim_{t\to\infty} L(t)$ *and* $\lim_{t\to\infty} M(t)$ *exist and lie in* $\mathfrak{h}$ *and* $\mathfrak{k}$, *respectively.*

Proof. Since L, $N \in \mathfrak{su}(r)$, $[N, L]^* = -[N, L]$, where $*$ denotes the Hermitian conjugate. Set

$$S(L) = \mathrm{tr}(LN) = f(g). \tag{16}$$

We have

$$\begin{aligned}\frac{dS(L)}{dt} &= \mathrm{tr}([L,[N,L]]N) \\ &= \mathrm{tr}(2LNLN - L^2N^2 - NL^2N) \\ &= \mathrm{tr}([L,N]^2) \\ &= -\,\mathrm{tr}([L,N][L,N]^*) \leq 0.\end{aligned}$$

Thus $S(L)$ is a monotonically decreasing function of t. Since G is compact, $S(L)$ is bounded from below, and then $\lim_{t\to\infty} S(L)$ exists and $\lim_{t\to\infty} \frac{dS(L)}{dt} = 0$. This implies

$$[\lim_{t\to\infty} L(t),\ N] = 0.$$

Noting that the regularity of N leads to $\mathfrak{h} = C_{\mathfrak{g}}(N)$, we see $\lim_{t\to\infty} L(t) \in \mathfrak{h}$. Differentiating $\mathrm{tr}(MQ) = f(g)$, we can derive

$$[\lim_{t\to\infty} M(t),\ Q] = 0.$$

From this it follows that $\lim_{t\to\infty} M(t) \in \mathfrak{k} = C_{\mathfrak{g}}(Q)$. Q.E.D.

Remark. Critical points of the Lax equation (12) which do not lie in $\mathfrak{h}$ are not stable. Hence the integral curve of (12) generally flows from an initial value toward a matrix lies in $\mathfrak{h}$. A similar property of the Lax equation (13) also can be proved.

As was pointed out in [8, 9] in the case $G = \mathbf{SO}(r)$, when Q is regular, this asymptotic property of solutions of the Lax equations (12) has an application to a simple combinatorial optimization problem. Let us set

$$Q = \mathrm{diag}(q_1, q_2, \cdots, q_r), \qquad N = \mathrm{diag}(n_1, n_2, \cdots, n_r), \tag{17}$$

where $q_i \neq q_j$ and $n_i \neq n_j$. From Proposition 2, $\lim_{t\to\infty} L(t) \in \mathfrak{h}$. Set $\mathrm{diag}(l_1^\infty, l_2^\infty, \cdots, l_r^\infty) = \lim_{t\to\infty} L(t)$. Since $\mathrm{Spec}\, L(t) = \mathrm{Spec}\, Q$ for any $t \in \mathbb{R}$, we see that l_j^∞ takes the form

$$l_j^\infty = q_{\pi(j)},$$

$1 \le j \le r$, for some permutation π of r-words. This implies that the limit of $g(t)$ takes the form $\lim_{t\to\infty} g(t) = D\Pi$, where $D = \mathrm{diag}(d_1, d_2, \cdots, d_r)$ with $|d_j| = 1$ and Π is the permutation matrix associated with π. Thus there is an infinite number of critical points of $f(g)$ on $G = \mathbf{SU}(r)$. We note that $\lim_{t\to\infty} L(t) = \Pi^{-1}Q\Pi$, and there are $r!$ critical points of $S(L)$ on the flag manifold $G/K = \mathbf{SU}(r)/\mathbf{S}(\mathbf{U}(1)\times\cdots\times\mathbf{U}(1))$. Consequently,

$$\lim_{t\to\infty} S(L) = \sum_{j=1}^{r} q_{\pi(j)} n_j \in \mathbb{R}.$$

On exactly one of the critical points on G/K, $S(L)$ takes the maxima (minima) of $\sum_{j=1}^{r} q_{\pi(j)} n_j$. Each local maxima (minima) of $S(L)$ is realized according to the choice of the initial value $L(0)$. Thus the flow $L(t)$ of the Lax equation solves the combinatorial optimization problem.

§4. Reduction to Hermitian symmetric spaces

We discuss a reduction of the Lax equations associated with homogeneous spaces to symmetric spaces by coalescing eigenvalues of $\mathrm{ad}\, Q$. When Q is regular, $\mathfrak{k} = \mathfrak{h}$. As the eigenvalues coalesce, $\mathfrak{k}$ grows larger. If $\mathfrak{k}$ and its complement $\mathfrak{m}$ satisfy $[\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}$, $[\mathfrak{k}, \mathfrak{m}] \subset \mathfrak{m}$ and $[\mathfrak{m}, \mathfrak{m}] \subset \mathfrak{k}$, then G/K is called a symmetric space. In this case the Lax equation (13) reads

$$\frac{dM_{\mathfrak{k}}}{dt} = [M_{\mathfrak{m}},\ [Q, M_{\mathfrak{m}}]], \qquad \frac{dM_{\mathfrak{m}}}{dt} = [M_{\mathfrak{k}},\ [Q, M_{\mathfrak{m}}]]. \tag{18}$$

We here restrict ourselves to an interesting class of symmetric spaces, the Hermitian symmetric spaces. Let us set $M(t) = M_{\mathfrak{k}} + M_{\mathfrak{m}}^{+} + M_{\mathfrak{m}}^{-}$, where $M_{\mathfrak{m}}^{\pm} \in \mathfrak{m}^{\pm}$. Since $[Q, M_{\mathfrak{m}}^{\pm}] = \pm\alpha(Q) M_{\mathfrak{m}}^{\pm}$, we derive from (18)

$$\frac{dM_{\mathfrak{k}}}{dt} = 2\alpha(Q)[M_{\mathfrak{m}}^{-},\ M_{\mathfrak{m}}^{+}], \qquad \frac{dM_{\mathfrak{m}}^{\pm}}{dt} = \pm\alpha(Q)[M_{\mathfrak{k}},\ M_{\mathfrak{m}}^{\pm}]. \tag{19}$$

Next we give some examples of the Lax equation (19) associated with the complex projective space $\mathbf{SU}(r)/\mathbf{S}(\mathbf{U}(1)\times\mathbf{U}(r-1))$, type **AIII**, and the Hermitian symmetric space $\mathbf{Sp}(n)/\mathbf{U}(n)$ of type **CI**. When $r = 3$, we may write [14, p.275]

$$Q = i\begin{pmatrix} 2\lambda & 0 & 0 \\ 0 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{pmatrix}, \qquad M_{\mathfrak{k}} = \begin{pmatrix} ib_1 & 0 & 0 \\ 0 & ib_2 & b_4 \\ 0 & -\bar{b}_4 & ib_3 \end{pmatrix},$$

$$M_{\mathfrak{m}}^{+} = \begin{pmatrix} 0 & a_1 & a_2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad M_{\mathfrak{m}}^{-} = \begin{pmatrix} 0 & 0 & 0 \\ -\bar{a}_1 & 0 & 0 \\ -\bar{a}_2 & 0 & 0 \end{pmatrix},$$

where $\sum_{j=1}^{3} b_j = 0$. It is easy to see $\alpha(Q) = 3i\lambda$. Part of components of the Lax equation associated with $\mathbf{SU}(3)/\mathbf{S}(\mathbf{U}(1)\times\mathbf{U}(2))$ are:

$$\frac{da_1}{dt} = -3\lambda a_1(b_1 - b_2) + 3i\lambda a_2\bar{b}_4,$$
$$\frac{db_1}{dt} = -6\lambda(|a_1|^2 + |a_2|^2),$$
$$\frac{db_2}{dt} = 6\lambda|a_1|^2, \qquad \frac{db_4}{dt} = 6\lambda\bar{a}_1 a_2.$$

Suppose we take the Hermitian symmetric space of type **CI**, where $n = 2$ for convenience. In this case we may set [17]

$$Q = i\begin{pmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & -\lambda & 0 \\ 0 & 0 & 0 & -\lambda \end{pmatrix},$$

$$M_{\mathfrak{k}} = \begin{pmatrix} 0 & b_1 + ib_2 & 0 & 0 \\ -b_1 + ib_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & b_1 - ib_2 \\ 0 & 0 & -b_1 - ib_2 & 0 \end{pmatrix},$$

$$M_{\mathfrak{m}}^{+} = \begin{pmatrix} 0 & 0 & a_1 & a_2 \\ 0 & 0 & a_2 & a_3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad M_{\mathfrak{m}}^{-} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\bar{a}_1 & -\bar{a}_2 & 0 & 0 \\ -\bar{a}_2 & -\bar{a}_3 & 0 & 0 \end{pmatrix}.$$

Explicit form of the Lax equation associated with $\mathbf{Sp}(2)/\mathbf{U}(2)$ can be easily derived. We here omit it.

We now consider an application of the Lax equations associated with the Grassmann manifold to a least squares fitting problem. Let $y_j = (y_{j1}, y_{j2}, \cdots, y_{jp+q})$, $1 \le j \le m$, be m-data in $\mathbb{C}^{p+q}$, where each y_{jk} is measured with observation error ε_{jk}. Suppose that ε_{jk} are independent and follow the normal distribution $N(0,1)$. We set

$$Y = \begin{pmatrix} y_{11} & \cdots & y_{1p+q} \\ \vdots & & \vdots \\ y_{m1} & \cdots & y_{mp+q} \end{pmatrix}. \tag{20}$$

Let P be the orthogonal projection of $\mathbb{C}^{p+q}$ onto a p-plane, namely, $P^2 = P$, $\operatorname{rank} P = p$ and $P^* = P$. The matrix $\widetilde{P}$ defined by $\widetilde{P} = I - P$ is the orthogonal projection of $\mathbb{C}^{p+q}$ onto a q-plane. The total perpendicular squares distance of the m-points y_j onto the q-plane is given by

$$\operatorname{tr}(Y(I-\widetilde{P})(Y(I-\widetilde{P}))^*) = \operatorname{tr}(PY^*Y). \tag{21}$$

Let us consider the total least squares fitting problem of finding the matrix P which minimizes $\operatorname{tr}(PY^*Y)$. Since Y^*Y is Hermitian, there is an element g of $\mathbf{U}(p+q)$ such that $g^{-1}Y^*Yg$ is diagonal. We set $D_Y = g^{-1}Y^*Yg$. It is easy to see $\operatorname{tr}(PY^*Y) = \operatorname{tr}(g^{-1}PgD_Y)$. Moreover, since P is diagonalizable by an element of $\mathbf{U}(p+q)$, we can suppose that P is diagonal. Let us consider the critical point problem for $\operatorname{tr}(PY^*Y)$ via the gradient flow on $\mathbf{U}(p+q)$. Set

$$L_P(t) = g^{-1}Pg, \tag{22}$$

where $g = g(t) \in \mathbf{U}(p+q)$ and $t \in \mathbb{R}$. From Proposition 1, we can prove

Proposition 3. *If D_Y has distinct eigenvalues, namely, the singular values of Y are distinct, then the Lax equation*

$$\frac{dL_P(t)}{dt} = [L_P(t),\ [D_Y, L_P(t)]] \tag{23}$$

describes the gradient flow for the least squares fitting function (21).

Proof. Set $N = iD_Y$. This is an regular element of the Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{u}(p+q)$. Note also $iL_P \in \mathfrak{u}(p+q)$. Since ${L_P}^2 = L_P$, $\operatorname{rank} L_P = p$ and ${L_P}^* = L_P$, iL_P can be regarded as an element of an adjoint orbit of $\mathbf{U}(p+q)$ through $iP \in \mathfrak{h}$ and then identified with a point in $\mathbf{U}(p+q)/\mathbf{U}(p) \times \mathbf{U}(q) \approx \mathbf{SU}(p+q)/\mathbf{S}(\mathbf{U}(p) \times \mathbf{U}(q))$, the Grassmann manifold. We set $Q = -iP \in \mathfrak{h}$. From the assumption and Proposition 1, the Lax equation $\frac{dL(t)}{dt} = [L(t), [N, L(t)]]$, where $L(t) = g^{-1}Qg = -iL_P$, is equivalent to the gradient flow for $\operatorname{tr}(PY^*Y)$. Q.E.D.

It follows from Proposition 2 that the critical points of the fitting function $\mathrm{tr}(PY^*Y)$ can be given by those of the flow $L(t)$ on the Grassmann manifold. Consequently, the projection matrix P minimizing $\mathrm{tr}(PY^*Y)$ can be expressed as

$$P_{\min} = \lim_{t\to\infty} L_P(t). \tag{24}$$

The least squares estimate of q-plane is determined by $I - P_{\min}$. This is also the maximum likelihood estimate through the theorem of Gauss and Markov.

There is an identity $[L,[L,[L,N]]] = -[L,N]$ which follows from $P^2 = P$. Substituting $L = g^{-1}Qg$ into this we have

$$[Q,\ [Q,\ [Q,\ gNg^{-1}]]] = -[Q,\ gNg^{-1}]. \tag{25}$$

Let us set $J = \mathrm{ad}\, Q : \mathfrak{m} \to \mathfrak{m}$. Since $[Q, gNg^{-1}] \in \mathfrak{m}$ and $J([Q, gNg^{-1}]) \in \mathfrak{m}$ for any N and g, J defines the (almost) complex structure in the sense of (8). Regarding the fitting function $\mathrm{tr}(PY^*Y) = \mathrm{tr}(LN)$ as a Hamiltonian, Bloch [3] derived the Hamiltonian equation of the Lax form

$$\frac{dL(\tau)}{d\tau} = [L(\tau),\ N]. \tag{26}$$

The Hamiltonian equation has $p+q-1$ independent integrals of motion in involution and can be solved explicitly when $p = 1$. Recently, Bloch [4] showed that the gradient flow for $\mathrm{tr}(LN)$ may be obtained from the Hamiltonian flow (26) by letting a linear transformation $\widetilde{J} = \mathrm{ad}\, L$ act on the Hamiltonian flow. It is easy to check $\widetilde{J}^2([L,N]) = -[L,N]$ and $[L,N] \in \mathfrak{m} = T(G/K)$, where $T(G/K)$ is the tangent space of the complex Grassmann manifold G/K. However, Bloch's mapping $\widetilde{J}$ is not an endomorphism of $\mathfrak{m}$, namely, $\widetilde{J}([L,N]) \in \mathfrak{g} = \mathfrak{k} \oplus \mathfrak{m}$. In Proposition 3, we have derived the Lax equation (23) in the least squares fitting problem without using the complex structure. We should remark here that our derivation naturally fits into the usual situation in least squares estimation in the following sense. A Lax equation similar to (23) can be obtained in the case of real Grassmann manifold, $G/K = \mathbf{SO}(p+q)/\mathbf{S}(\mathbf{O}(p) \times \mathbf{O}(q))$. The resulting Lax equation solves a least squares problem of q-plane fitted to a date in $\mathbb{R}^{p+q}$. Hence the maximum likelihood problem can be turned into a critical point problem of the Lax equation.

Finally in this section we briefly deal with a linear programing problem. Let $G/K = \mathbf{U}(p+q)/\mathbf{U}(p) \times \mathbf{U}(q)$, the Grassmann manifold.

Brockett already pointed out in [8] that when $p = 1$ the Lax equation (12) can solve a linear programming problem. We wish to pursue a somewhat different manner here for any integer p. As in the proof of Proposition 3, we write $f(g) = \mathrm{tr}(g^{-1}QgN) = \mathrm{tr}(L_P D_Y)$, where $N = iD_Y \in \mathfrak{h}$, regular, and $Q = -iP \in \mathfrak{h}$. Here L_P is defined by (22) with $iL_P \in \mathfrak{u}(p+q)$, $L_P{}^2 = L_P$ and $L_P{}^* = L_P$. Set $D_Y = \mathrm{diag}(|\delta_1|^2, |\delta_2|^2, \cdots, |\delta_{p+q}|^2)$. Let $(l_1(t), l_2(t), \cdots, l_{p+q}(t))$ be diagonal elements of $L_P(t)$. Since $\mathrm{tr}(L_P D_Y) = \sum_{j=1}^{p+q} l_j(t)|\delta_j|^2$, the Lax equation (23) is viewed as the gradient flow for $\sum_{j=1}^{p+q} l_j(t)|\delta_j|^2$. It is to be noted from $L_P{}^2 = L_P$, $L_P{}^* = L_P$ and $\mathrm{tr}\, L_P = p$ that the diagonal elements should satisfy the conditions

$$0 \leq l_j(t) \leq 1, \qquad \sum_{j=1}^{p+q} l_j(t) = p. \tag{27}$$

Such a $(l_1(t), l_2(t), \cdots, l_{p+q}(t))$ may then be regarded as a point in a convex polytope in $\mathbb{R}^{p+q}$. The Lax equation (23) describes an interior flow on it which approaches critical points (vertices of the polytope) as t goes to infinity. This assertion follows from Proposition 2, i.e. $\lim_{t\to\infty} iL_P(t) \in \mathfrak{h}$. The critical point problem for $f(g)$ then corresponds to a linear programming problem. Recently, Bayer and Lagarias [2] considered a dynamical picture of Karmarkar's polynomial time algorithm [13] which solves various linear programming problems as an interior flow on polytopes. They observed that the algorithm is deeply related to the nonlinear integrable system of non-Lax type. It would be most interesting to make clear the link between the Lax equation (23) and the nonlinear system in [2].

§5. Solutions of Lax equations associated with the Grassmann manifold

We come in the position to propose a method for solving the Lax equations associated with the complex Grassmann manifold. Let us express the gradient flow (10) on $G = \mathbf{SU}(p+q)$ as

$$\frac{dg}{dt} \cdot g^{-1} = [gNg^{-1}, Q].$$

Decomposing $\mathfrak{g} = \mathfrak{su}(p+q)$ as in (7) we have

$$\left(\frac{dg}{dt} \cdot g^{-1}\right)^{\pm} = \mp\alpha(Q)(gNg^{-1})^{\pm}, \tag{28}$$

where $X^{\pm}$ denote the $\mathfrak{m}^{\pm}$-part of X, respectively. It is to be remarked from $\overline{\alpha(Q)} = -\alpha(Q)$ that the Hermitian conjugate of the $\mathfrak{m}^{+}$-part is equivalent to the $\mathfrak{m}^{-}$-part. Let $g(0) \in \mathbf{SU}(p+q)$ be an initial value for the $\mathfrak{m}^{+}$-part of (28) and $M(0) = g(0)Ng^{-1}(0) \in \mathfrak{su}(p+q)$. The exponential of matrix $\alpha(Q)tM(0)$ is nonsingular for any $t \in \mathbb{R}$ and always admits the unique decomposition

$$\exp(\alpha(Q)tM(0)) = A^{-1}(t)B(t). \tag{29}$$

Here $A(t) \in \mathbf{SU}(p+q)$, $A(0) = I$ and $B(t)$ is an element of the group of lower triangular matrices with positive diagonal entries such that $B(0) = I$. The decomposition (29) is carried out by a Gram-Schmidt orthogonalization (see, for example, [23]). The following proposition is the key for our construction of solutions of the Lax equation.

Proposition 4. *The initial value problem for the* $\mathfrak{m}^{+}$*-part of* (28) *is uniquely solved via the decomposition* (29) *as*

$$g(t) = A(t)g(0). \tag{30}$$

Proof. From (29), $\alpha(Q)AM(0)A^{-1} = -\frac{dA}{dt} \cdot A^{-1} + \frac{dB}{dt} \cdot B^{-1}$. We take the $\mathfrak{m}^{+}$-part of it and derive

$$(\frac{dA}{dt} \cdot A^{-1})^{+} = -\alpha(Q)(AM(0)A^{-1})^{+}.$$

Setting $A(t) = g(t)g^{-1}(0)$, where $g(0)$ is the initial value, we see that $g(t)$ satisfies the $\mathfrak{m}^{+}$-part of (28), $(\frac{dg}{dt} \cdot g^{-1})^{+} = -\alpha(Q)(gNg^{-1})^{+}$. Since the gradient flow satisfies the Lipschitz condition for any $t \in \mathbb{R}$, $g(t)$ defined by (30) gives a unique solution of the initial value problem.

Q.E.D.

With a help of Proposition 4, we can easily construct solutions of the Lax equations (12) and (13). Let $A(t)$ be the factor of the unique decomposition (29) and let $L(0) = g^{-1}(0)Qg(0)$ and $M(0) = g(0)Ng^{-1}(0)$ be initial values for the Lax equations (12) and (13) associated with the complex Grassmann manifold. Then

Proposition 5. *$L(t)$ and $M(t)$ defined by*

$$\begin{aligned} L(t) &= g^{-1}(0)A^{-1}(t)QA(t)g(0), \\ M(t) &= A(t)g(0)Ng^{-1}(0)A^{-1}(t) \end{aligned} \tag{31}$$

uniquely solve the initial value problems for (12) *and* (13), *respectively.*

It is well-known that the generalized Toda equations are solved by the QR decomposition of $\exp(tL(0))$ into a product of unitary and upper triangular matrices, where $L(0)$ is a Jacobi matrix [20]. We have shown that the Lax equations (12) and (13) in the specific case are also solved by a similar but slightly different way. From $M(0) = g(0)Ng^{-1}(0)$ we can compute $\exp(\alpha(Q)tM(0))$ and the factor $A(t)$ in (29) by a polynomial time algorithm. This may provide a suggestive picture of a polynomial time calculation process for the linear programming problem in the previous section.

§6. Discussions

We have established a nontrivial generalization and classification of the Lax equation which appears in a least squares problem. A very explicit way of a reduction of the Lax equation from reductive homogeneous spaces to symmetric spaces is given. We also found a method for solving an initial value problem for the Lax equation associated with the complex Grassmann manifold. The Lax equation has some applications to a least squares fitting problem and a linear programming problem. Any algorithm which finds 'optimal' solution must be iterative, and consequently, it describes a dynamical system. It is to be expected that evidence provided by the results will give an impetus to the further design of efficient algorithms for these problems as dynamical processes.

The Lax equation in the least squares problem induces an interior flow on a convex polytope in $\mathbb{R}^{p+q}$. Here let us recall a result by Tomei [21]. He proved that the level manifold of a generalized Toda equation is homeomorphic to a convex polytope in $\mathbb{R}^n$. The Lax equation (15) obviously includes the Toda equation in [21]. However it is not clear how to relate this Toda equation to the Lax equation (23). The level manifold of the original Toda equation itself is $\mathbb{R}^n$ [16]. Recently the author [18] obtained a different generalization of the Toda equation. The resulting level manifold is diffeomorphic to a certain cylinder. It would be interesting to extend the approach developed in this paper to the generalized Toda equation in [18] and clarify the nature of these connections.

Acknowledgment. The author would like to thank A.M. Bloch, R.W. Brockett and J.C. Lagarias for useful conversations. Thanks are also due to S. Amari, T. Iwai and K. Ueno for stimulating discussions. This research was partially supported by Grant-in-Aid for Scientific Research no. 02854003 from the Japan Ministry of Education, Science, and Culture.

References

[1] V.I. Arnol'd and A.B. Givental', Symplectic geometry, in "Dynamical Systems IV (V.I. Arnol'd and S.P. Novikov eds.)", Encyc. of Math. Sci. Vol. 4, Springer-Verlag, Berlin, 1990.

[2] D.A. Bayer and J.C. Lagarias, The nonlinear geometry of linear programming. II Legendre transform coordinates and central trajectories, Trans. Amer. Math. Soc., **314** (1989), 527–581.

[3] A.M. Bloch, A completely integrable Hamiltonian system associated with line fitting in complex vector spaces, Bull. Amer. Math. Soc. (New Series), **12** (1985), 250–254.

[4] ———, The Kahler structure of the total least squares problem, Brockett's steepest descent equations, and constrained flows, preprint.

[5] A.M. Bloch, R.W. Brockett and T. Ratiu, A new formulation of the generalized Toda lattice equations and their fixed points analysis via the moment map, Bull. Amer. Math. Soc. (New Series), **23** (1990), 477–485.

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

[7] O.I. Bogoyavlensky, On perturbations of the periodic Toda lattice, Commun. Math. Phys., **51** (1976), 201–209.

[8] R.W. Brockett, Dynamical systems that sort lists, diagonalize matrices and solve linear programming problem, in Proc. 27th IEEE Conf. on Decision and Control (1988), 799–803.

[9] ———, Least squares matching problems, Lin. Alg. Appl., **122/123/124** (1989), 761–777.

[10] P. Deift, T. Nanda and C. Tomei, Ordinary differential equations and the symmetric eigenvalue problem, SIAM J. Numer. Anal., **20** (1983), 1–22.

[11] A.P. Fordy and P.P. Kulish, Nonlinear Schrödinger equations and simple Lie algebras, Commun. Math. Phys., **89** (1983), 427–443.

[12] S. Helgason, "Differential Geometry, Lie Groups, and Symmetric Spaces", Academic, New York, 1978.

[13] N. Karmarkar, A new polynomial-time algorithm for linear programming, Combinatorica, **4** (1984), 373–395.

[14] S. Kobayashi and K. Nomizu, "Foundation of Differential Geometry Vol. II", Interscience, New York, 1969.

[15] B. Kostant, The solution to a generalized Toda lattice and representation theory, Adv. Math., **34** (1979), 195–338.

[16] J. Moser, Finitely many points on the line under the influence of an exponential potential – An integrable system, in "Dynamical Systems, Theory and Applications (J. Moser ed.)", Lec. Notes in Phys. Vol. 38, Springer-Verlag, Berlin, 1975, pp. 467–497.

[17] Y. Nakamura, On Heisenberg models with values in Hermitian symmetric spaces, Lett. Math. Phys., **9** (1985), 107–112.

[18] ———, The level manifold of a generalized Toda equation hierarchy, Trans. Amer. Math. Soc., to appear.

[19] J. von Newmann, Some matrix-inequalities and metrization of matric-space, in "John von Neumann: Collected Works Vol. IV (A.H. Taub ed.)", Pergamon, New York, 1962, pp. 205–218. This paper first appeared in Tomsk. Univ. Rev. **1** (1937), 286–300.

[20] A.G. Reyman and M.A. Semenov-Tian-Shansky, Reduction of Hamiltonian systems, affine Lie algebras and Lax equations, Invent. Math., **54** (1979), 81–100.

[21] C. Tomei, The topology of isospectral manifolds of tridiagonal matrices, Duke Math. J., **51** (1984), 981–996.

[22] D.S. Watkins, Isospectral flows, SIAM Rev., **26** (1984), 379–391.

[23] J.H. Wilkinson, "The Algebraic Eigenvalue Problem", Clarendon, Oxford, 1965.

Department of Mathematics
Gifu University
Yanagido, Gifu 501-11
Japan