JOURNAL OF MATHEMATICAL PHYSICS VOLUME 42, NUMBER 8 AUGUST 2001

Solitons and almost-intertwining matrices

Alex Kasman[a)]
Department of Mathematics, College of Charleston, 66 George Street, Charleston, South Carolina 29464

Michael Gekhtman
Department of Mathematics, University of Notre Dame, Notre Dame, Indiana 46556

(Received 13 November 2000; accepted for publication 2 April 2001)

We define the algebraic variety of almost intertwining matrices to be the set of triples (X,Y,Z) of $n\times n$ matrices for which $XZ=YX+T$ for a rank one matrix T. A surprisingly simple formula is given for tau functions of the KP hierarchy in terms of such triples. The tau functions produced in this way include the soliton and vanishing rational solutions. The induced dynamics of the eigenvalues of the matrix X are considered, leading in special cases to the Ruijsenaars–Schneider particle system. © *2001 American Institute of Physics.* [DOI: 10.1063/1.1379313]

I. INTRODUCTION

The KP hierarchy[1,2] is a well-studied system of integrable nonlinear partial differential equations with Lax form

$$\frac{\partial \mathcal{L}}{\partial t_i}=[\mathcal{L},\mathcal{L}_+^i],\quad i=1,2,3,...$$

for a monic, first-order pseudodifferential operator $\mathcal{L}$. In one of its formulations, the KP hierarchy is a set of bilinear equations for the "tau function" $\tau(t_1,t_2,t_3,t_4,\ldots)$ depending upon infinitely many "time variables" t_i $(i\in\mathbb{Z}_+)$. In this paper we will consider τ functions of the form:

$$\tau_M(t_1,t_2,\ldots):=\det(Xe^{g(Z)}+e^{g(Y)}),\tag{1}$$

where $M=(X,Y,Z)$ is a triple of $n\times n$ constant complex matrices and the function g is defined as

$$g(W):=\sum_{i=1}^{\infty} t_iW^i,\quad t_i\in\mathbb{C}.\tag{2}$$

(To avoid issues of convergence, we will here consider only the case in which all but a finite number of the parameters $t_i\in\mathbb{C}$ are nonzero.)

It is *not* true that (1) always gives the formula for a function which satisfies the KP hierarchy. For instance, as we shall see from Remark 2.2 and Theorem 3.1, in the 2×2 case formula (1) is only a tau function if $\det[(XZ-YX)(Y-Z)]=0$. On the other hand, among the solutions one can obviously write this way are the *one-soliton solutions* which are the natural generalizations in this context of the solitary wave from which the term "soliton" was coined by Zabusky and Kruskal.[3] The *standard* τ function for the one-soliton solution takes the form (1) where $M=(X,Y,Z)\in\mathbb{C}^3$ is any triple of scalar constants. (To exclude the degenerate cases we must further assume that X and $Y-Z$ are nonzero.) This τ function describes a single line soliton of the KP equation. More generally, one may be interested in τ functions of n-soliton solutions ("nonlinear superpositions" of n different line solitons) or their rational degenerations. These τ functions are usually written in a form that looks very different than (1).

[a)]Electronic mail: kasman@math.cofc.edu

Of course, when X, Y, and Z are scalar as in the one-soliton case, then the determinant which appears in (1) is unnecessary. However, the main result of this paper is that the n-soliton solutions also take the form (1) and that they arise in the case that X, Y, and Z are three $n\times n$ matrices satisfying the condition $\mathrm{rank}(XZ-YX)=1$. In fact, this same rank one condition provides not only the nondegenerate soliton solutions to the KP hierarchy but also their rational degenerations. Thus, we see that the one-soliton τ function is merely a special case of this much more general formula.

II. ALMOST-INTERTWINING MATRICES

It is common to say that an operator X intertwines the operators Y and Z if one has that

$$XZ=YX. \tag{3}$$

Definition 2.1: Given three $n\times n$ matrices X, Y and Z, we define the rank $\kappa(X,Y,Z)$ *to which* X intertwines Y and Z by the formula

$$\kappa(X,Y,Z)=\mathrm{rank}(XZ-YX)=n-\dim\ker(XZ-YX).$$

For fixed $k,n\in\mathbb{N}$ $(0\leqslant k\leqslant n)$ define

$$\mathcal{M}_n^k=\{(X,Y,Z)|\kappa(X,Y,Z)\leqslant k\}$$

to be the set of all triples of $n\times n$ matrices $M=(X,Y,Z)$ such that $\kappa(M)\leqslant k$.

In most instances, one expects to find that $\kappa(X,Y,Z)=n$, its maximum value. For $\kappa(X,Y,Z)$ to be lower means that X does, in fact, intertwine Y and Z on the positive dimensional subspace $\ker(XZ-YX)$. In particular, when $\kappa(X,Y,Z)=0$, then $XZ=YX$ and so X does actually intertwine the other two matrices. If Y and Z are not intertwined by X, then the best one could ask for would be for $\kappa(X,Y,Z)$ to be equal to one, and so it seems reasonable to say that they are almost intertwined in this case.

Remark 2.1: Note that a triple (X,Y,Z) is in $\mathcal{M}_n^k$ precisely when the $k\times k$ minor determinants of the matrix $XZ-YX$ all vanish. Consequently, $\mathcal{M}_n^k$ has the geometric structure of an affine algebraic variety in the $3n^2$-dimensional vector space of $n\times n$ matrix triples.

The following elementary observations will be used to establish the connection between almost intertwining matrices and solitons:

Lemma 2.1: • There is a natural $GL(n)\times GL(n)$ action on $\mathcal{M}_n^k$ given by

$$(G,H)\in GL(n)\times GL(n):\quad (X,Y,Z)\in\mathcal{M}_n^k\mapsto(GXH^{-1},GYG^{-1},HZH^{-1})\in\mathcal{M}_n^k,$$

which restricts on the diagonal to the natural $GL(n)$ action of simultaneous conjugation

$$G\in GL(n):\quad (X,Y,Z)\in\mathcal{M}_n^k\mapsto(GXG^{-1},GYG^{-1},GZG^{-1})\in\mathcal{M}_n^k.$$

• Let Λ and Ω be $n\times n$ matrices satisfying the commutation relationships

$$[\Lambda,Y]=0,\quad [\Omega,Z]=0,$$

then

$$\kappa(X,Y,Z)=\kappa(\Lambda X\Omega,Y,Z).$$

Proof: Both claims are easily verified by noting that $\kappa(X,Y,Z)\leqslant k$ if and only if

$$XZ=YX+\sum_{i=1}^{k} v_i\otimes w_i \tag{4}$$

for n-vectors $\{v_i\}$ and $\{w_i\}$ and that $\kappa(X,Y,Z)$ is exactly the minimum k for which such an equation exists. □

The main result of this section is the following lemma:

Lemma 2.2: Given three $n\times n$ matrices $\hat{X}$, Y and Z, let $H(a,b,c)\in\mathbb{C}[a,b,c]$ be the polynomial defined by

$$H(a,b,c)=H_1(a)H_2(b,c)-H_1(b)H_2(a,c)+H_1(c)H_2(a,b) \tag{5}$$

with

$$H_1(a)=\det(\hat{X}(aI-Z)+(aI-Y))$$

and

$$H_2(a,b)=(a-b)\det(\hat{X}(aI-Z)(bI-Z)+(aI-Y)(bI-Y)).$$

If $\kappa(\hat{X},Y,Z)\leqslant 1$ then $H(a,b,c)\equiv 0$ is the zero polynomial.

Proof: To say that $\kappa(\hat{X},Y,Z)\leqslant 1$ is equivalent to saying that there exist vectors v and w_1 such that

$$\hat{X}Z-Y\hat{X}=-vw_1^T. \tag{6}$$

[In the case $\kappa(\hat{X},Y,Z)=0$ one of these vectors is the zero vector.] Also, merely for the sake of convenience, we introduce the notation

$$Z_a=(aI-Z),\quad Y_a=(aI-Y)$$

and recall that adj(M) is the classical adjoint matrix [i.e., $\mathrm{adj}(M)=\det(M)M^{-1}$ if M is invertible].

Now, using (6) to eliminate "$\hat{X}Z$," one can rewrite $H_1(a)$, $H_2(a,b)$ as

$$H_1(a)=\det(Y_a(\hat{X}+I)+vw_1^T),\quad H_2(a,b)=(a-b)\det(Y_aY_b(\hat{X}+I)+Y_{a+b}vw_1^T+vw_2^T),$$

where $w_2^T=w_1^TZ$.

Next, since $H(a,b,c)$ depends on $\hat{X}$ polynomially, it is enough to prove that $H(a,b,c)=0$ for almost all $\hat{X}$. Let us assume that $\det(\hat{X}+I)=\gamma\neq 0$. Then we can eliminate reference to $\hat{X}$ by writing

$$H_1(a)=\gamma\det(Y_a+vu_1^T),\quad H_2(a,b)=\gamma(a-b)\det(Y_aY_b+Y_{a+b}vu_1^T+vu_2^T),$$

where

$$u_1^T=w_1\cdot(\hat{X}+I)^{-1},\quad u_2^T=w_2^T\cdot(\hat{X}+I)^{-1}.$$

Let us further rewrite $H_2(a,b)$ as

$$H_2(a,b)=(a-b)\gamma\det(Y_aY_b+Y_avu_1^T+vu_1^TY_b+v(u_2^T+u_1^TY))$$

$$=(a-b)\gamma\det((Y_a+vu_1^T)(Y_b+vu_1^T)+vu_2^T).$$

Finally, denote $Y-vu_1^T$ by M. We obtain

$$H_1(a)=\gamma\det(M_a),\quad H_2(a,b)=\gamma(a-b)\det(M_aM_b+vu_2^T).$$

Note that

$$\det(M_aM_b+vu_2^T)=\det(M_a)\det(M_b)+u_2^T\,\mathrm{adj}(M_aM_b)v=\gamma^{-2}H_1(a)H_1(b)(1+u_2^TM_a^{-1}M_b^{-1}v)$$

and since $[M_a,M_b]=0$ we also have that

$$M_a^{-1}M_b^{-1}=\frac{1}{a-b}(M_b^{-1}-M_a^{-1}).$$

Therefore,

$$\det(M_aM_b+vu_2^T)=\gamma^{-1}H_1(a)H_1(b)+\frac{H_1(a)(u_2^T\,\mathrm{adj}(M_b)v)-H_1(b)(u_2^T\,\mathrm{adj}(M_a)v)}{\gamma(a-b)}.$$

So, using the notation $p(a)=a\gamma^{-2}H_1(a)-\gamma^{-1}u^T\,\mathrm{adj}(M_a)v$, we see that $\det(M_aM_b+vu^T)$ is a Bezoutian of the form

$$\det(M_aM_b+vu^T)=\frac{p(a)H_1(b)-p(b)H_1(a)}{a-b}.$$

Substituting $H_2(a,b)=\gamma(p(a)H_1(b)-p(b)H_1(a))$ into the expression for $H(a,b,c)$ immediately yields $H\equiv 0$. □

Remark 2.2: The special case $n=2$ turns out to be surprisingly simple. A quick calculation verifies that for arbitrary 2×2 matrices $\hat{X}$, Y, and Z the polynomial $H(a,b,c)$ is given by the formula

$$H(a,b,c)=(a-b)(b-c)(c-a)\det[(\hat{X}Z-Y\hat{X})(Y-Z)].$$

III. TAU FUNCTIONS

A. Main theorem

It is easy to check that if $\kappa(M)=0$ then the formula for τ_M defined in (1) is a tau function of the KP hierarchy. In fact, in this case in which (3) is satisfied one has

$$\tau_M(t_1,t_2,t_3,\ldots)=\det(X+I)\exp\left(\sum_{i=1}^{\infty}\sum_{j=1}^{n}(\lambda_j^i)t_i\right),$$

where $\{\lambda_j\}$ are the eigenvalues of Y. Since the function

$$u(x,y,t)=2(\log\tau_M(x,y,t,0,0,...))_{xx}=0$$

is the trivial solution to the KP equation, we say that τ_M is merely a gauge transformation of the trivial tau function.

Moreover, with g defined as in (2) and τ_M defined by (1), we observe that this is still a τ function in the case $\kappa(M)=1$. In fact, it is more interesting in this "almost-intertwining" case since we get nontrivial soliton and rational solutions in this way.

Theorem 3.1: If $\kappa(M)\leqslant 1$ for $M=(X,Y,Z)$ then the function

$$\tau_M(t_1,t_2,\ldots)=\det(Xe^{g(Z)}+e^{g(Y)}),\quad g(W)=\sum_{i=1}^{\infty}t_iW^i$$

is a tau function of the KP hierarchy with corresponding (stationary) Baker–Akhiezer function

$$\psi_M(x,z):=\frac{\det(X(zI-Z)e^{xZ}+(zI-Y)e^{xY})}{z^n\det(Xe^{xZ}+e^{xY})}e^{xz}.$$

Proof: Given the semi-infinite vector $\vec{t}=(t_1,t_2,t_3,\ldots)$, we use the notation $\tau_M(\vec{t})=\tau_M(t_1,t_2,\ldots)$. For an arbitrary constant a, we define the semi-infinite vector $[a]=(a,a^2/2,a^3/3,...)$. Then, it is sufficient to prove that the continuous function $\tau(\vec{t})$ defined in (1) satisfies the Hirota equation in Miwa form[4,5]

$$0=(b-c)\tau(\vec{t}-[a^{-1}])\tau(\vec{t}-[b^{-1}]-[c^{-1}])-(a-c)\tau(\vec{t}-[b^{-1}])\tau(\vec{t}-[a^{-1}]-[c^{-1}])$$
$$+(a-b)\tau(\vec{t}-[c^{-1}])\tau(\vec{t}-[a^{-1}]-[b^{-1}]) \tag{7}$$

uniformly in a, b, and c and for all $\vec{t}$.

However, from the definition we see that

$$\begin{aligned}\tau(\vec{t}-[a^{-1}])&=\det(Xe^{g(Z)}e^{\ln(I-a^{-1}Z)}+e^{g(Y)}e^{\ln(I-a^{-1}Y)})\\&=a^{-n}\det(e^{g(Y)})\det(\hat{X}(aI-Z)+(aI-Y))\\&=a^{-n}\det(e^{g(Y)})H_1(a),\end{aligned}$$

where we have chosen $\hat{X}=e^{-g(Y)}Xe^{g(Z)}$ and used the notation of Lemma 2.2. Similarly,

$$(a-b)\tau(\vec{t}-[a^{-1}]-[b^{-1}])=a^{-n}b^{-n}\det(e^{g(Y)})H_2(a,b).$$

Consequently, (7) is equivalent to demonstrating that the polynomial $H(a,b,c)$ in Lemma 2.2 is zero in the case of this $\hat{X}$, Y, and Z. But, according to the second result in Lemma 2.1 we have that $\kappa(\hat{X},Y,Z)=\kappa(X,Y,Z)\leqslant 1$ and so Lemma 2.2 demonstrates that the Hirota equation is satisfied.

Once we know that τ_M is a tau function, the formula for ψ_M is derived from simply using the "famous Japanese formula,"[2]

$$\psi_M(x,z)=\frac{\tau_M(x-z^{-1},-z^{-2}/2,-z^{-3}/3,...)}{\tau_M(x,0,0,...)}e^{xz}.$$

Note that the numerator is simply $\tau_M(\vec{t}-[z^{-1}])$ with $\vec{t}=(x,0,0,...)$. So, again expanding this in terms of the power series for the logarithm we derive the desired expression for ψ_M. □

Remark 3.1: Technically, although the function $\tau\equiv 0$ solves the bilinear equations of the KP hierarchy, it is not generally considered to be a tau function. [In particular, there is no associated operator $\mathcal{L}$ satisfying the Lax equation or function $u(x,y,t)$ satisfying the KP equation.] In the preceding we have not been careful to make certain that τ is nonzero. In fact, one can certainly choose $M\in\mathcal{M}_n^1$ so that $\tau_M=0$. Consequently, Theorem 3.1 should be understood to say that if τ_M is nonzero (which is generally the case) then it is a KP tau function.

Remark 3.2: Since the Baker–Akhiezer function ψ_M in Theorem 3.1 has the property that $z^n e^{-xz}\psi_M$ is a polynomial in z, it must be that τ_M is the tau function of a rank-one KP solution with a (singular) rational spectral curve. In particular, it must be a soliton solution or one of its rational degenerations. Well-known consequences[6–8] of this fact are the following:

Corollary 3.1: Let $K=K_M(t_1,t_2,t_3,\ldots,\partial_x)$ be the ordinary differential operator determined by simply substituting the formal symbol ∂_x in for z in the polynomial

$$K(t_1,t_2,\ldots,z)=\frac{\det(X(zI-Z)e^{g(Z)}+(zI-Y)e^{g(Y)})}{\det(Xe^{(Z)}+e^{g(Y)})}.$$

Then, equating x and t_1, $\mathcal{L}_M=K\partial_x K^{-1}$ satisfies the Lax equations

$$\frac{\partial}{\partial t_i}\mathcal{L}=[\mathcal{L},(\mathcal{L}^i)_+].$$

Moreover, the function

$$u(x,y,t) := \frac{\partial^2}{\partial x^2}\log \tau_M(x,y,t,0,0,...)$$

satisfies the KP equation

$$\tfrac{3}{4}u_{yy} = (u_t - \tfrac{1}{4}(6uu_x + u_{xxx}))_x .$$

Remark 3.3: It is well known and easily verified (cf. Ref. 2) that multiplication by a function of the form $\exp(\Sigma\gamma_i t_i)$ takes one tau function to another having the same corresponding Lax operator $\mathcal{L}$. Such a change is often referred to as a ''gauge transformation'' in KP theory. Since $\det \exp g(Y)$ is a function of this form with $\gamma_i = \Sigma \lambda_j^i$ (where λ_j are the eigenvalues of Y counted according to multiplicity) it follows that:

Corollary 3.2: For $M=(X,Y,Z)\in\mathcal{M}_n^1$,

$$\hat{\tau}_M(t_1,t_2,\ldots) = \det(Xe^{g(Z)}e^{-g(Y)}+I)$$

is also a KP tau function differing from τ_M by only a gauge transformation.

Remark 3.4: Since the tau function and Baker–Akhiezer function are defined as they are by determinants of X, Y, and Z, simultaneously conjugating all three leaves the corresponding solution unchanged. Consequently, it would be possible to use Lemma 2.1 to take to quotient of $\mathcal{M}_n^1$ by the action of GL and then would be natural to define $\tau_{\bar{M}}$ for $\bar{M}\in\overline{\mathcal{M}_n^1}=\mathcal{M}_n^1/GL(n)$.

B. Special cases

1. Gelfan'd–Dickii hierarchies (N-KdV)

The N-KdV or Gelfan'd–Dickii hierarchies are special classes of KP solutions for which $\mathcal{L}^N$ is an ordinary differential operator and hence is independent of the KP flows whose indices are multiples of N. In particular, we say a tau function is an N-KdV tau function if it factors as $\tau = f\cdot g$ where

$$\frac{\partial}{\partial t_{iN}}g=0 \quad \forall i\in\mathbb{N}, \quad \frac{\partial}{\partial t_1}f=0.$$

In other words, except for a factor independent t_1, τ is independent of t_j for all j that are multiples of N.

Let $\mathcal{M}_n^1(N)$ be the subset of $\mathcal{M}_n^1$,

$$\mathcal{M}_n^1(N)=\{(X,Y,Z)\in\mathcal{M}_n^1 : Y^N=Z^N\}.$$

Theorem 3.2: For $M\in\mathcal{M}_n^1(N)$, the corresponding tau function τ_M is a solution of the N-KdV hierarchy.

Proof: If we consider only the dependence upon t_1 and t_j (j a multiple of N) then

$$\begin{aligned}\tau_M &= \det(Xe^{t_1Z+t_jZ^j}+e^{t_1Y+t_jY^j})\\ &= \det(Xe^{t_1Z+t_jZ^j}+e^{t_1Y+t_jZ^j})\\ &= \det(Xe^{t_1Z}+e^{t_1Y})\det(e^{t_jZ^j}).\end{aligned}$$

□

For example, if we consider the restriction $Y=-Z$, then we are looking for matrix pairs (X,Z) satisfying

$$\mathrm{rank}(XZ+ZX)=1.$$

In this case, the formula (1) will produce a tau-function solution to the KdV hierarchy (independent of all even time flows). (Note that special cases have been considered elsewhere in the literature in the context of integrable particle systems.[9,10])

2. Solitons

The n-soliton solutions to the KP hierarchy are identified by these properties:

(1) The BA function $\psi(x,z)$ when multiplied by a degree n polynomial $q(z)=z^n+\cdots$ has the form

$$\bar{\psi}(x,z)=q(z)\psi(x,z)=\left(\sum_{i=1}^{n} a_i(x)z^i\right)e^{xz}.$$

(2) There are n independent linear "conditions" satisfied by $\overline{\psi}(x,z)$ of the form

$$\hat{\alpha}_i\overline{\psi}(x,\lambda_i)+\hat{\beta}_i\overline{\psi}(x,\mu_i)=0, \quad 1\leqslant i\leqslant n$$

(with $\lambda_i\neq\mu_i$).

These solutions can be constructed from $\mathcal{M}_n^1$ by choosing the point $M=(X,Y,Z)$ with

$$X_{ij}=\frac{\alpha_i}{\beta_j(\lambda_j-\mu_i)}, \quad Y_{ij}=\mu_i\delta_{ij}, \quad Z_{ij}=\lambda_i\delta_{ij}.$$

This can be verified, for instance, by noting that because $[Y,Z]=0$, the tau function $\hat{\tau}_M$ takes the form (cf. Corollary 3.2)

$$\hat{\tau}=\det(Xe^{g(Z)-g(Y)}+I).$$

For any index set $J\subset\{1,\ldots,n\}$, the principal minor of $Xe^{g(Z)-g(Y)}$ can be written as

$$\left(\prod_{i\in J}\frac{\alpha_i}{\beta_i}e^{g(\lambda_i)-g(\mu_i)}\right)\det\left(\frac{1}{\lambda_i-\mu_{i'}}\right)_{i,i'\in J}.$$

The latter determinant is a Cauchy determinant and is equal to

$$\prod_{i,i'\in J;i<i'}\frac{(\lambda_i-\lambda_{i'})(\mu_i-\mu_{i'})}{(\lambda_i-\mu_{i'})(\mu_i-\lambda_{i'})}\prod_{i\in J}\frac{1}{\lambda_i-\mu_i}.$$

Setting

$$c_i=\frac{\alpha_i}{\beta_i(\lambda_i-\mu_i)},$$

we obtain

$$\hat{\tau}=\sum_{J\subset\{1,\ldots,n\}}\prod_{i\in J}c_ie^{g(\lambda_i)-g(\mu_i)}\prod_{i,i'\in J;i<i'}\frac{(\lambda_i-\lambda_{i'})(\mu_i-\mu_{i'})}{(\lambda_i-\mu_{i'})(\mu_i-\lambda_{i'})},$$

which coincides with the known formula for this n-soliton solution of the KP hierarchy.[11]

3. Polynomial τ functions and rational solutions

Clearly, in the case that Y and Z are chosen to be nilpotent, the definition of τ_M produces a *polynomial* in the time variables t_i. It is perhaps of greater interest to note that one may also get tau functions that are—up to a gauge transformation—polynomial in t_1 but an infinite series if all t_i are considered.

For example, choosing

$$X=\begin{pmatrix}1&1&0\\1&0&0\\1&0&0\end{pmatrix},\quad Y=\begin{pmatrix}\lambda&1&0\\0&\lambda&1\\0&0&\lambda\end{pmatrix},\quad Z=\begin{pmatrix}\lambda&0&0\\1&\lambda&0\\0&1&\lambda\end{pmatrix}$$

leads to (after a gauge transformation to remove an exponential factor):

$$\tau(x,y,t,0,0,...)$$
$$=1+(-3\lambda+3\lambda^2)t+\frac{9}{2}\lambda^4t^2+\frac{x^2}{2}+(6\lambda^3t+2\lambda-1)y+2\lambda^2y+(1+3\lambda^2t+2\lambda y)x.$$

Such solutions are well known and have been studied in previous papers.[6,8,12–14] However, one should especially compare the present approach with that of Wilson,[15] where these ''vanishing rational KP solutions'' are produced from matrix pairs (X,Z) satisfying $\mathrm{rank}(XZ-ZX+I)=1$. The main results in that paper concern the induced dynamics of the eigenvalues which behave as particles in a Calogero–Moser particle system. So, it may be of interest to similarly investigate the dynamics of the eigenvalues associated with almost-intertwining matrices.

IV. EIGENVALUE DYNAMICS

One of the most interesting things about the Ruijsenaars–Schneider (RS) particle system[9,16] is its connection to soliton tau functions. Specifically, certain KP tau functions can be written as

$$\tau(t_1,t_2,\ldots)=\det(X+I),$$

where $X=X(t_1,t_2,t_3,\ldots)$ is a matrix whose eigenvalues move according to the Ruijsenaars–Schneider Hamiltonian.

In this section we similarly study the dynamics of eigenvalues of time-dependent matrices in the context of almost-intertwining matrices to both reproduce and extend known results about the RS system and its connection to solitons.

A. Solitons and a matrix flow

Theorem 4.1: The vector fields V_i on the space of $n\times n$ matrix triples defined by

$$V_i(X_0,Y,Z)=(X_0Z^i-Y^iX_0,0,0) \tag{8}$$

are tangent to the manifold $\mathcal{M}_n^1$ and induce the flows in the variables t_i parametrized as

$$M_t=(X_t,Y,Z)=(e^{-g(Y)}X_0e^{g(Z)},Y,Z). \tag{9}$$

Proof: Note that the flows specified have the stated vector fields and that

$$X_tZ-YX_t=e^{-g(Y)}(X_0Z-YX_0)e^{g(Z)}$$

is a rank one matrix if X_0Z-YX_0 is. □

Remark 4.1: Given a parametrized flow $(X_t,Y,Z)\in\mathcal{M}_n^1$ as previously, the function $\hat{\tau}_M=(X_t+I)$ is another way to write the gauge transformed tau function from Corollary 3.2 with $M=(X_0,Y,Z)$.

B. General equations for eigenvalue dynamics

Given any matrices $X_0, \; Y, \; Z$ such that $\kappa(X_0,Y,Z)=1$ let us define $X=X_t$ according to (9). If we denote the eigenvalues of X_t by $\{Q_i(t)\}$ $(1\leqslant i\leqslant n)$, to what extent can we describe their dynamics by intrinsic equations (depending only on Q_i and their derivatives)?

In what follows we will only be considering the flow under the first time parameter t_1, but will write simply t in order to simplify exposition and will use a ''dot'' to indicate differentiation with respect to this parameter.

We define vectors v and w by the formula

$$(X_0 Z - Y X_0) = v w^T \tag{10}$$

and so we have the equations of motion

$$\dot{X} = v w^T, \quad \dot{Y} = 0, \quad \dot{Z} = 0. \tag{11}$$

For convenience we introduce the (time-dependent) matrix U which diagonalizes X and the logarithms of the eigenvalues q_i

$$Q = UXU^{-1} = \begin{pmatrix} Q_1 & 0 & 0 & \cdots & \\ 0 & Q_2 & 0 & 0 & \cdots \\ 0 & 0 & Q_3 & 0 & \cdots \\ \vdots & & & \ddots & \end{pmatrix}, \quad q_i = \ln(Q_i) \tag{12}$$

and define in analogy to (10) the matrices and vectors

$$\hat{Y} = UYU^{-1}, \quad \hat{Z} = UZU^{-1}, \quad \hat{v} = Uv, \quad \hat{w} = wU^{-1} \tag{13}$$

so that

$$Q\hat{Z} - \hat{Y}Q = \hat{v}\hat{w}^T. \tag{14}$$

Note, in particular, that looking at an individual element of Eq. (14) yields

$$Q_i \hat{Z}_{ij} - Q_j \hat{Y}_{ij} = \hat{v}_j \hat{w}_i . \tag{15}$$

Now, defining $M = \dot{U}U^{-1}$ we have in analogy to (11)

$$\dot{Q} = [M,Q] + \hat{v}\hat{w}^T, \quad \dot{\hat{Y}} = [M,\hat{Y}], \quad \dot{\hat{Z}} = [M,\hat{Z}]. \tag{16}$$

Since Q and $\dot{Q}$ have no off diagonal elements, we get from (16) that

$$\dot{Q}_i = \hat{v}_i \hat{w}_i = \dot{q}_i e^{q_i}, \tag{17}$$

and

$$M_{ij} = \frac{\hat{v}_j \hat{w}_i}{Q_i - Q_j} \quad (i \neq j). \tag{18}$$

It turns out to be especially useful to write the equations of motion in terms of q_i rather than Q_i because then we find by multiplying (16) by Q^{-1} that

$$\begin{pmatrix} \dot{q}_1 & 0 & \cdots \\ \vdots & \ddots & \\ 0 & \cdots & \dot{q}_n \end{pmatrix} = \dot{Q}Q^{-1} = M - QMQ^{-1} + Q\hat{Z}Q^{-1} - \hat{Y}. \tag{19}$$

Since Q is diagonal, $M - QMQ^{-1}$ has no diagonal and $Q\hat{Z}Q^{-1} - \hat{Y}$ has the same diagonal as $\hat{Z} - \hat{Y}$ and so

$$\dot{q}_i = (\hat{Z} - \hat{Y})_{ii}. \tag{20}$$

Finally, we can differentiate (20) and use (15), (17), and (18) to find the equation of motion

$$\ddot{q}_i = ([M, \hat{Z} - \hat{Y}])_{ii} \tag{21}$$

$$= \sum_{k \neq i} (M_{ik}(\hat{Z}_{ki} - \hat{Y}_{ki}) - M_{ki}(\hat{Z}_{ik} - \hat{Y}_{ik})) \tag{22}$$

$$= \sum_{k \neq i} \left(\frac{\dot{Q}_i \dot{Q}_k}{Q_i(Q_i - Q_k)} + \frac{\hat{v}_k \hat{w}_i \hat{Z}_{k,i}}{Q_i} + \frac{\dot{Q}_i \dot{Q}_k}{Q_k(Q_k - Q_i)} + \frac{\hat{v}_i \hat{w}_k \hat{Z}_{ik}}{Q_k} \right) \tag{23}$$

$$= \sum_{k \neq i} \frac{\dot{Q}_i \dot{Q}_k (Q_i + Q_k) - (Q_i - Q_k)(Q_i \hat{v}_i \hat{w}_k \hat{Z}_{ik} - Q_k \hat{v}_k \hat{w}_i \hat{Z}_{ki})}{Q_i Q_k (Q_i - Q_k)}. \tag{24}$$

C. A special case

We can further simplify (24) assuming that $\hat{w}$ has no zero component. In that case, we can utilize additional freedom of conjugation by a diagonal matrix to leave Q unchanged but modify U.

In particular, if $\hat{w}$ is a vector with no zero component, then we can put it in a form where $w = (1,1,1,...,1)$ by multiplying U by the diagonal matrix with w_i's along its diagonal. Now, in this "gauge," we know that $\hat{w}_i = 1$ and so by (17) we know that $\hat{v}_i = \dot{Q}_i$. This then gives us that

$$\ddot{q}_i = \sum_{k \neq i} \frac{\dot{Q}_i \dot{Q}_k (Q_i + Q_k) - (Q_i - Q_k)(Q_i \dot{Q}_i \hat{Z}_{ik} - Q_k \dot{Q}_k \hat{Z}_{ki})}{Q_i Q_k (Q_i - Q_k)}.$$

Ideally, we would like to be able to completely eliminate $\hat{Z}_{ki}$ from this equation and have an "intrinsic" equation for the eigenvalues. It seems that this can only be done when certain additional simplifying assumptions are made.

Suppose that we are in the case that

$$-\lambda \hat{Y} + \hat{Z} = \gamma I \Rightarrow \hat{Z}_{ij} = \lambda \hat{Y}_{ij} \quad (i \neq j). \tag{25}$$

Combining Eqs. (15) and (25) we find that

$$\hat{Z}_{ij} = \frac{\lambda \hat{v}_j \hat{w}_i}{\lambda Q_i + Q_j} \quad (i \neq j).$$

Substituting this into (24) and again using (17) one finds the intrinsic equations of motion

$$\ddot{q}_i = (\lambda - 1)^2 \dot{Q}_i \sum_{k \neq i} \frac{\dot{Q}_k (Q_i + Q_k)}{(Q_i - Q_k)(\lambda Q_i - Q_k)(\lambda Q_k - Q_i)}. \tag{26}$$

Note that the equations are independent of γ. In the case $\lambda=-1$ the dynamics of (26) *is* the Ruijsenaars–Schneider model.[9]

V. COMMENTS AND CONCLUSIONS

It is interesting to note that restrictions on $\kappa(X,Y,Z)$ for triples of square matrices have arisen before in the context of integrable systems. For example, though the notations are different, the key operator identity used by Sakhnovich[17] is such a restriction. Perhaps there is a deep connection between the results of that work and this one, though the relationship is not immediately apparent to us. A more relevant result was obtained by Nijhoff and Chalykh,[18] who used the condition rank $(XZ-qZX)=1$ for invertible X and Z and scalar q to construct solutions to the q-difference KP hierarchy. It is reasonable to suppose that their result could now also be obtained as a discretization of the results in the present work in the special case $Y=qZ$. [Another matrix approach[19] to q-KP made use of the condition rank $(XY-qYX+I)=1$.]

The suggestive appearance of these spaces of matrices in such different contexts within the study of integrable systems might indicate that we should look more carefully at the manifolds $\mathcal{M}_n^k$. For instance, we have implicitly constructed a map from $\mathcal{M}_n^1$ to the infinite dimensional Grassmannian[2] Gr_1, and $\mathcal{M}_n^k$ naturally has the structure of an algebraic variety, but so far we have little understanding of the geometry.

Wilson[15] constructs an adelic Grassmannian and a Hilbert scheme from the set of matrices satisfying rank $([X,Z]+I)=1$. Moreover, the natural symmetry of this set which is manifested as the involution $(X,Z)\mapsto(Z^T,X^T)$ has significance both for the KP hierarchy (bispectrality) and the Calogero–Moser particle system (self-duality). So, it is reasonable to wonder how the obvious symmetries of $\mathcal{M}_n^k$ are reflected in the soliton solutions to the KP hierarchy. We have already noted that multiplying X by a function of Y on the left and a function of Z on the right corresponds to the KP flows. Note also that if $\kappa(X,Y,Z)=1$ and X is invertible then $\kappa(X^{-1},Z,Y)=1$ as well and that this triple corresponds to the same KP solution. (In particular, these two points in $\mathcal{M}_n^1$ get mapped to the same point in $\mathrm{Gr}^{\mathrm{rat}}$.) Similarly, if Y is invertible then $\kappa(Y,X,XZY^{-1})=1$, but it is not immediately apparent what symmetry of KP is analogous.

One alternative characterization of $\mathrm{Gr}^{\mathrm{rat}}$ is as the Grassmannian of finite dimensional subspaces of finitely supported distributions.[6] Specifically, to identify a point $W\in\mathrm{Gr}^{\mathrm{rat}}$ it is sufficient to identify the finitely supported distributions in z which annihilate the normalized Baker–Akhiezer function. We showed in Sec. III B 2 that in the case of nondegenerate solitons, the eigenvalues of Y and Z determine the support of the distributions and X determines the coefficients. We conjecture that this situation holds in general:

Conjecture 5.1: The support of the distributions annihilating $z^n\psi_M$ for $M=(X,Y,Z)\in\mathcal{M}_n^1$ is the set of eigenvalues of the matrices Y and Z with the highest derivative taken at a particular eigenvalue being bounded by the size of the corresponding Jordan blocks.

ACKNOWLEDGMENTS

We are grateful to Annalisa Calini, Harry Braden, John Harnad, and Tom Ivey for helpful discussions.

[1] M. Sato and Y. Sato, Lecture Notes Num. Appl. Anal. **5**, 259 (1982).
[2] G. Segal and G. Wilson, Pub. Math. **61**, 5 (1985).
[3] N. J. Zabusky and M. D. Kruskal, Phys. Rev. Lett. **15**, 240 (1965).
[4] I. Krichever, P. Wiegmann, and A. Zabrodin, Commun. Math. Phys. **188**, 267 (1997).
[5] A. V. Zabrodin, Teor. Mat. Fiz. **113**, 179 (1997); translation in Theor. Math. Phys. **113**, 1347 (1998).
[6] A. Kasman, Commun. Math. Phys. **172**, 427 (1995).
[7] A. Kasman, Inverse Probl. **14**, 1481 (1998).
[8] G. Wilson and J. Reine Agnew, Math. **442**, 177 (1993).
[9] H. W. Braden and R. Sasaki, Prog. Theor. Phys. **97**, 1003 (1997).
[10] J. F. van Diejen and H. Puschmann, Duke Math. J. **104**, 269 (2000).
[11] V. G. Kac, *Infinite Dimensional Lie Algebras* (Cambridge University Press, Cambridge, 1985).
[12] H. Airault, H. P. McKean, and J. Moser, Commun. Pure Appl. Math. **30**, 95 (1977).
[13] I. M. Krichever, Funct. Anal. Appl. **12**, 59 (1978).

[14] J. Zubelli, Lett. Math. Phys. **24**, 41 (1992).
[15] G. Wilson, Invent. Math. **133**, 1 (1998).
[16] S. N. M. Ruijsenaars and H. Schneider, Ann. Phys. (Leipzig) **170**, 370 (1986).
[17] A. L. Sakhnovich, Inverse Probl. **10**, 699 (1994).
[18] F. V. Nijhoff and O. A. Chalykh, Usp. Mat. Nauk **54**, 173 (1999).
[19] P. Iliev, C. R. Acad. Sci., Ser. I: Math. **329**, 877 (1999).