Orthogonal polynomials and the finite Toda lattice

Alex Kasman
Department of Mathematics, University of Georgia

(Received 11 March 1996; accepted for publication 12 August 1996)

The choice of a finitely supported distribution is viewed as a degenerate bilinear form on the polynomials in the spectral parameter z and the matrix representing multiplication by z in terms of an orthogonal basis is constructed. It is then shown that the same induced time dependence for finitely supported distributions which gives the ith KP flow under the dual isomorphism induces the ith flow of the Toda hierarchy on the matrix. The corresponding solution is an $\mathcal{N}$ particle, finite, non-periodic Toda solution where $\mathcal{N}$ is the cardinality of the support of c plus the sum of the orders of the highest derivative taken at each point. © *1997 American Institute of Physics.* [S0022-2488(97)00501-X]

I. INTRODUCTION

Recent interest in the Toda lattice[1] has stemmed from its role in relating theories of quantum gravity to soliton theory.[2] This correspondence is given by a measure $d\rho$ determined by the partition function (i.e., the ''specific heat'') of matrix models which is interpreted as an inner product on time-dependent polynomials in the spectral parameter.[3] In that construction, the polynomials are written in terms of an orthogonal basis with respect to this nondegenerate inner product and the Toda lattice is determined as the matrix representing multiplication in the spectral parameter.

The present paper replaces integration with respect to the measure $d\rho$ by an arbitrary finitely supported distribution. It is then shown that the same correspondence between orthogonal polynomials and integrable systems continues to hold in the case of the induced degenerate bilinear form. This relates the Toda hierarchy to techniques for the construction of τ-functions of the KP hierarchy[4,5] using finitely supported distributions.

It should be noted that there is an intersection of the construction developed below and that discussed in the opening paragraph. In particular, finitely supported distributions which are linear combinations of Dirac delta functions can be represented as Stieltjes integrals with respect to Heaviside functions.[6] The solutions constructed from such distributions by the method below are known[7] and are the same as those which would be given by the corresponding measure. However, finitely supported distributions involving differentiation (i.e., $\mu_i > 0$) and the Toda lattices which they generate are discussed here for the first time.

II. ASSOCIATING A JACOBI MATRIX TO A DISTRIBUTION

Let c be the finitely supported distribution

$$c = \sum_{i=1}^{m} \delta_{\lambda_i} \circ \sum_{j=0}^{\mu_i} \alpha_{ij} \partial_z^j, \tag{2.1}$$

where δ_λ is the delta function evaluating its argument at $z=\lambda$, the constants $\lambda_i \in \mathbb{C}$ are distinct, ∂_z is the differential operator ∂/∂_z, and $\alpha_{ij} \in \mathbb{C}$ with $\alpha_{i\mu_i} \neq 0$. (In fact, the discussion to follow only depends upon c as determined up to a nonzero constant multiple, and so the coefficients α_{ij} can be viewed as elements of $\mathbb{P}^{\mathcal{N}-1}\mathbb{C}$.) Then let $\mathcal{N}$ be the integer

$$\mathcal{N}=m+\sum_{i=1}^{m}\mu_i,$$

where m and μ_i are as in (II.1).

Associated to c we have the symmetric bilinear form on $\mathbb{C}[z]$ defined by

$$\langle p,q\rangle=c(pq).$$

Note that given two polynomials

$$p=\sum_{i=1}^{n-1}\alpha_i z^i,\quad q=\sum_{i=1}^{n-1}\beta_i z^i$$

of degree less than n,

$$\langle p,q\rangle=(\alpha_0,\ldots,\alpha_{n-1})\cdot T_n\cdot\begin{pmatrix}\beta_0\\ \vdots\\ \beta_{n-1}\end{pmatrix}$$

where

$$T_n=\begin{pmatrix}c(1) & \ldots & c(z^{n-1})\\ \vdots & \ddots & \vdots\\ c(z^{n-1}) & \ldots & c(z^{2n-1})\end{pmatrix}. \tag{2.2}$$

A. The annihilator of *c*

Any function sufficiently differentiable on the support of c acts on the right by composition:

$$c\circ p(f)=c(pf).$$

In particular, we may associate to c its annihilator in $\mathbb{C}[z]$.

Definition II.1: For any distribution c, let $I_c\subset\mathbb{C}[z]$ denote the ideal

$$I_c=\{p\in\mathbb{C}[z]\,|\,c\circ p\equiv 0\}.$$

Lemma II.1: Let c be written in the form (II.1) *and let*

$$\sigma_c(z)=\prod_{i=1}^{m}(z-\lambda_i)^{\mu_i+1}.$$

Then I_c is the ideal generated by $\sigma_c(z)$:

$$I_c=\sigma_c\mathbb{C}[z].$$

Proof: Since $c\circ\sigma_c\equiv 0$, it is clear that $\sigma_c(z)\mathbb{C}[z]\subset I_c$. Then, let $p(z)\in I_c$ be written in the form

$$p(z)=q(z)r(z),\quad r(z)=\prod_{i=1}^{m}(z-\lambda_i)^{\gamma_i}, \tag{2.3}$$

where $q(z) \in \mathbb{C}[z]$ is such that $q(\lambda_i) \neq 0$. Suppose that $\gamma_j < \mu_j + 1$ for some particular $1 \leqslant j \leqslant m$. Then for the polynomial

$$s(z) = (z - \lambda_j)^{\mu_j - \gamma_j} \prod_{i \neq j} (z - \lambda_i)^{\mu_i + 1}$$

we have that the distribution

$$c \circ r \circ s = k \delta_{\lambda_j}, \quad k = (\mu_j!) \alpha_{j \mu_j} \neq 0,$$

is a nonzero distribution evaluating its argument at λ_j without differentiation. But then, since $c \circ p = c \circ qr \equiv 0$ we have that

$$0 = c \circ qr(s(z)) = c \circ rs(q(z)) = kq(\lambda_j),$$

which implies that $q(\lambda_j) = 0$, contradicting the assumption. Consequently, each element of I_c written in form (II.3) has $\gamma_i \geqslant \mu_i + 1$ and $I_c \subset \sigma_c(z) \mathbb{C}[z]$.

Since deg $\sigma_c = \Sigma_{i=1}^{m} (\mu_i + 1) = \mathscr{N}$, we then have the following.

Corollary II.1: There exists a polynomial $p \in \mathbb{C}[z]$ with deg $p = n$ such that $c \circ p \equiv 0$ if and only if $n \geqslant \mathscr{N}$.

B. A basis for $\mathbb{C}[z]$

The choice of a generic distribution c uniquely specifies a basis for $\mathbb{C}[z]$ as follows.

Definition II.2: For any positive integer i, let τ_i denote the determinant

$$\tau_i = |T_i|,$$

where T_i is the symmetric matrix described in (II.2). We say that c is *regular* if $\tau_i \neq 0$ for $i = 1, \ldots, \mathscr{N}$. Let $\mathscr{P}$ denote the vector space of polynomials of degree less than $\mathscr{N}$. If c is regular, then the Gram–Schmidt orthogonalization specifies a unique basis $\{p_0, \ldots, p_{\mathscr{N}-1}\}$ of $\mathscr{P}$ such that

$$p_i(z) = z^i + O(z^{i-1})$$

and which is orthogonal with respect to the form $\langle \cdot, \cdot \rangle$. Furthermore, since $\tau_{\mathscr{N}-1} \neq 0$ the form is nondegenerate on $\mathscr{P}$ and so

$$\langle p_i, p_i \rangle \neq 0, \quad i = 0, \ldots, \mathscr{N} - 1.$$

It will now be supposed that c is in fact regular and that the polynomials p_i for $i = 0, \ldots, \mathscr{N} - 1$ have been fixed by the Gram–Schmidt orthogonalization. We may then define

$$p_{\mathscr{N}+i}(z) = z^i \sigma_c(z), \quad i = 0,1, \ldots .$$

By Lemma II.1, $p_{\mathscr{N}+i} \in I_c$, and so it is in the kernel of the form. Therefore, the basis of monic polynomials $\{p_i | i \geqslant 0\}$ for $\mathbb{C}[z]$ is orthogonal with respect to the form, but the form is degenerate.

C. The tri-diagonal matrix

The significance of the basis specified in the preceding section is that multiplication by z is represented as a tri-diagonal Jacobi matrix in terms of this basis.

Proposition II.1: There exist numbers a_i and b_i in $\mathbb{C}$ such that

$$zp_i = p_{i+1} + b_i p_i + a_i p_{i-1}$$

for all $i>0$.

Proof: Since each polynomial p_i is monic of degree i, we certainly have

$$zp_n=p_{n+1}+\sum_{j=0}^{n} \alpha_j p_j$$

for some constants α_j. But then applying the functional $\langle p_i,\cdot\rangle$ to zp_n yields

$$\langle p_i,zp_n\rangle=\langle zp_i,p_n\rangle=\langle p_{i+1}+b_ip_i+a_ip_{i-1},p_n\rangle,$$

which is zero if $i<n-2$. On the other hand, one could also compute this as

$$\langle p_i,zp_n\rangle=\langle p_i,\alpha_ip_i\rangle=\alpha_i\langle p_i,p_i\rangle.$$

If $i<\mathscr{N}$, then $\langle p_i,p_i\rangle \neq 0$ and so $\alpha_i=0$. Finally, for $i\geqslant\mathscr{N}$, the claim is true by construction since $zp_i=p_{i+1}$.

Proposition II.2: Denote by A_n *the constant* $\langle p_n,p_n\rangle$. *Then*
(i) $A_n=a_nA_{n-1}$,
(i) $a_n \neq 0$ for $n=0,\dots,\mathscr{N}-1$,
(iii) $A_n/A_k=a_n\cdots a_{k+1}$ for $k<n<\mathscr{N}$.

Proof: The first relationship can be found by using the fact that $\langle zp,q\rangle=\langle p,zq\rangle$ and so

$$\langle zp_n,p_{n-1}\rangle=\langle p_{n+1},p_{n-1}\rangle+b_n\langle p_n,p_{n-1}\rangle+a_n\langle p_{n-1},p_{n-1}\rangle=a_n\langle p_{n-1},p_{n-1}\rangle$$

is also equal to

$$\langle p_n,zp_{n-1}\rangle=\langle p_n,p_n\rangle,$$

producing the desired result.

Then, by the nondegeneracy of the bilinear form on $\mathscr{P}$, we have that $a_nA_{n-1}=A_n=\langle p_n,p_n\rangle \neq 0$ for $0\leqslant n\leqslant\mathscr{N}-1$. The final claim clearly follows from the first by an inductive argument.

Associate to c the $\mathbb{N}\times\mathbb{N}$ tri-diagonal matrix

$$L=\begin{pmatrix} b_0 & 1 & 0 & 0 & 0 & \dots \\ a_1 & b_1 & 1 & 0 & 0 & \\ 0 & a_2 & b_2 & 1 & 0 & \\ \vdots & & \ddots & \ddots & \ddots & \end{pmatrix}.$$

Outside of the principal $\mathscr{N}\times\mathscr{N}$ minor, this matrix is simply the shift matrix with 1s along the super-diagonal and zeroes elsewhere. Note that L corresponds to multiplication by z in $\mathbb{C}[z]$ with basis $\{p_i\}$. This is particularly important in the next result.

Notation: Denote by $L^i_{j,k}$ the element in the jth column and kth row of the matrix L^i. Note that since L is indexed by $\mathbb{N}\times\mathbb{N}$, the top left corner is $L^i_{0,0}$ and not $L^i_{1,1}$ as one might expect.

Proposition II.3: $\langle z^ip_k,p_n\rangle=L^i_{n,k}A_n$.

Proof: By orthogonality, the only significant term in z^ip_k is the p_n term in its expansion in the orthogonal basis. However, this is simply $L^i_{n,k}p_n$. So $\langle zp_k,p_n\rangle=\langle L^i_{n,k}p_n,p_n\rangle=L^i_{n,k}A_n$.

By the symmetry of the form used in Proposition II.3, we then also have the following.

Corollary II.2: $L^i_{n,k}A_n=L^i_{k,n}A_k$.

III. TIME DEPENDENCE

Now suppose that c is an arbitrary, i.e., not necessarily regular, finitely supported distribution. To it we associate the time-dependent distribution

$$\hat{c}=c\circ \exp\sum_{i=1}^{\infty} t_i z^i.$$

Note that $\mathcal{N}_{\hat{c}}=\mathcal{N}_c$ and, moreover, $\sigma_{\hat{c}}=\sigma_c$ since neither the support nor the highest derivative taken at each point is affected by this composition. Whenever $\underline{\mathbf{t}}=(t_1,t_2,\ldots)$ is chosen such that $\hat{c}$ is regular, we may associate to it a basis of polynomials and a tri-diagonal matrix by the method of the preceding section. Thus, one is led to consider a basis $\{p_i(z,\underline{\mathbf{t}})\}$ of polynomials and a time-dependent matrix $L(\underline{\mathbf{t}})$ which are defined whenever $\hat{c}$ is regular.

Note: This time dependence for distributions was introduced in Ref. 8 because it induces the KP flow on the Sato Grassmannian under the dual isomorphism. In fact, this is a convenient way to prove the next claim:

Proposition III.1: The distribution $\hat{c}$ is regular for almost every value of $\underline{\mathbf{t}}=(t_1,t_2,\cdots)$.

Proof: By Corollary II.1 the distributions $c\circ z^i$ are linearly independent for $n=0,\ldots,\mathcal{N}-1$. Then the determinants τ_n are nonzero, time-dependent functions

$$\tau_n(\underline{\mathbf{t}})=\begin{vmatrix} \hat{c}(1) & \cdots & \hat{c}(z^n) \\ & \ddots & \\ \hat{c}(z^n) & & \hat{c}(z^{2n-1}) \end{vmatrix}$$

$$=\begin{vmatrix} c(\exp \Sigma t_i z^i) & \cdots & \partial^{n-1}/\partial t_1^{n-1} c(\exp \Sigma t_i z^i) \\ & \ddots & \\ \partial^{n-1}/\partial t_1^{n-1} c(\exp \Sigma t_i z^i) & \cdots & \partial^{2n-1}/\partial t_1^{2n-1} c(\exp \Sigma t_i z^i) \end{vmatrix}.$$

In fact, if we let $V_{\hat{c},n}$ denote the set of polynomials in the kernel of the distributions $\hat{c}\circ z^i$ for $i=0,\ldots,n$, then the Hilbert closure of $z^{-n}V_{\hat{c},n}$ is the a point $W_n(\underline{\mathbf{t}})$ in the Sato Grassmannian *Gr* and the Wronskian determinant above gives the corresponding tau function for the KP hierarchy.[4] So, we can cite Ref. 9 to show that these functions have isolated zeros. The distribution $\hat{c}$ is then regular on the complement of the zeros of the τ-functions τ_i for $i=0,\ldots,\mathcal{N}-1$.

Note: Tau functions determined from symmetric Wronskian matrices or Hankel determinants of the form above are known to be associated with finite Toda lattices.[10–12]

IV. DIFFERENTIAL EQUATIONS

This section will determine differential equations satisfied by the matrix $L(\underline{\mathbf{t}})$ in the temporal variable t_i. Throughout the remainder, prime ($'$) will be used to denote the derivative with respect to this variable. Since the form $\langle\cdot,\cdot\rangle$ is now taken to be the time-dependent form specified by $\hat{c}$, its derivative is given by the following lemma.

Lemma IV.1: $\langle p,q\rangle'=\langle z^i p,q\rangle+\langle p',q\rangle+\langle p,q'\rangle$.

Proof:

$$\langle p,q\rangle'=(c(e^{\Sigma t_j z^j}pq))'=c(e^{\Sigma t_j z^j}(z^i pq+p'q+pq'))=\langle z^i p,q\rangle+\langle p',q\rangle+\langle p,q'\rangle.$$

The leading coefficients of the polynomials p_n are constant, and so they satisfy differential equations of the form

$$p_n'=\sum_{k=0}^{n-1} C_k^n p_k. \tag{4.1}$$

Define the time-dependent functions C_k^n by this formula. In fact, since $\sigma_{\hat{c}}$ is constant in time,

$$p_n'=0 \quad \text{for} \quad n\geqslant\mathcal{N}.$$

Thus, it is clear that $C_k^n=0$ for $n\geqslant\mathcal{N}$.

Proposition IV.1: The coefficients C_k^n for $k<n<\mathcal{N}$ in (IV.1) *can be written either as*

$$C_k^n=-\frac{A_n}{A_k}L_{n,k}^i \tag{4.2}$$

or

$$C_k^n=-L_{k,n}^i. \tag{4.3}$$

In particular, $C_k^n=0$ if $i<n-k$.

Proof: This can be seen by differentiating the equation

$$\langle p_n,p_k\rangle=0$$

because then you get

$$\langle z^i p_n,p_k\rangle+\langle p_n',p_k\rangle+\langle p_k',p_n\rangle=0$$

which (using Proposition II.3) implies that

$$C_k^n A_k=-L_{n,k}^i A_n.$$

Since $k<\mathcal{N}$, $A_k\neq 0$ and we may solve for C_k^n yielding (IV.2). Then, substituting for A_n by the formula in Corollary II.2,

$$C_k^n A_k=-L_{k,n}^i A_k$$

which leads to the equivalent form (IV.3). Furthermore, it is elementary to determine that $L_{n,k}^i=0$ if $i<n-k$ merely from the tri-diagonal form of the matrix.

The main result of the present paper is the equations of motion satisfied by a_i and b_i.

Theorem IV.1: *The dependence of the distribution $\hat{c}$ on the time variable t_i induces the equations of motion*

$$b_n'=a_{n+1}L_{n+1,n}^i-a_nL_{n,n-1}^i \tag{4.4}$$

and

$$a_n'=(b_n-b_{n-1})L_{n-1,n}^i+L_{n-1,n+1}^i-L_{n-2,n}^i. \tag{4.5}$$

Proof: Since the actions of $\partial/\partial t_i$ and multiplication by z commute, we can equate the coefficients of p_n in $z(p_n')$ and $(\partial/\partial t_i)(zp_n)$.

$$zp_n'=z\sum_{j=0}^{n-1}C_j^n p_j=\sum_{j=0}^{n-1}C_j^n(p_{j+1}+b_jp_j+a_jp_{j-1})$$

and so the coefficient of p_n is just C_{n-1}^n. Alternatively,

$$\frac{\partial}{\partial t_i}(p_{n+1}+b_np_n+a_np_{n-1})=p_{n+1}'+b_n'p_n+b_n\left(\sum_{j=0}^{n-1}C_j^np_j\right)+a_n'p_{n-1}+p_{n-1}'a_n$$

and the coefficient of p_n is just $C_n^{n+1}+b_n'$. Equating these and making use of (IV.2) yields the equation of motion (IV.4). (Here we take $L_{0,-1}^i=0$ to handle the boundary case $n=0$.)

Similarly, equating the coefficients of p_{n-1} in these same expressions we get that

$$b_{n-1}C_{n-1}^n+C_{n-2}^n=b_nC_{n-1}^n+a_n'+C_{n-1}^{n+1}.$$

Using the substitution (IV.3) and solving for a_n' gives the desired form (IV.5). (Again, $L_{1,-1}^i=0$ to handle the case $n=1$.)

The equations (IV.4) and (IV.5) are one form of the Toda hierarchy and can be written in the Lax form

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

where the minus subscript indicates the projection to the lower triangular part. Since the super-diagonal elements are the only nonzero elements outside the principal $\mathcal{N}\times\mathcal{N}$ minor, this is in fact an $\mathcal{N}$-particle finite nonperiodic Toda lattice.

Theorem IV.2: *Let c be any finitely supported distribution and* $\hat{c}=c\circ\exp\Sigma t_iz^i$. *Then the corresponding matrix L is an* $\mathcal{N}$ *particle finite nonperiodic Toda lattice.*

V. REMARKS

As usual,[3] one may write the functions $a_i(\underline{\mathbf{t}})$ and $b_i(\underline{\mathbf{t}})$ in terms of the τ-functions $\tau_i(\underline{\mathbf{t}})$:

$$a_i=\frac{\tau_i\tau_{i+2}}{\tau_{i+1}^2},\quad b_i=\frac{\partial}{\partial t_1}\log\frac{\tau_{i+1}}{\tau_i},$$

for $i=0,\dots,\mathcal{N}-1$ where $\tau_0\equiv 1$. This is an easier way to construct the solution corresponding to a distribution c than determining the orthogonal basis of polynomials as above.

The points $W_i\in Gr$ described in Proposition III.1 are clearly seen to be related by the formula

$$zW_{i+1}\subset W_i$$

and are therefore related by Darboux transformations. As shown in Ref. 10, these are precisely the Darboux transformations which preserve the N-boson form of the corresponding KP solutions. The geometric spectral data is a line bundle over a rational curve with one singularity introduced by bringing together the points on a desingularization with coordinates λ_i and multiplicity μ_i+i+1. The inclusion of the coordinate rings induces covering maps from the more singular to the less singular curves.

One may wish to consider the moduli space of all distributions c with some given value of $\mathcal{N}$ so as to have a moduli of $\mathcal{N}$-particle nonperiodic Toda solutions. The different forms of c leading to an $\mathcal{N}$-particle system are indexed by the Young diagrams of with $\mathcal{N}$ blocks. Given such a Young diagram, a distribution may be specified by attaching a distinct value $\lambda_i\in\mathbb{C}$ to each column and a constant $\alpha_{ij}\in\mathbb{C}$ to the $j+1$st block in the column. The different diagrams lead to qualitatively different behaviors in the corresponding solutions. In particular, the τ functions give KP solitons when the Young diagrams consists entirely of columns of length one and, alternatively, they give rational KP solutions when the Young diagram has only one column.

ACKNOWLEDGMENTS

The author appreciates the advice and assistance of M. Rothstein, M. Bergvelt, and M. Adams.

[1] M. Toda, J. Phys. Soc. Jpn. **22** (1975).
[2] C. Ahn and K. Shigemoto, Phys. Lett. B **263**, 44 (1991).
[3] K. Sogo, J. Phys. Soc. Jpn. **62**, 1887 (1993).
[4] G. Wilson, J. reine angew. Math. **442**, 177 (1993).
[5] A. Kasman, ''Darboux Transformations from *n*-KdV to KP,'' preprint (1996).
[6] R. Kanwal, *Generalized Functions: Theory and Technique* (Academic, New York, 1983).
[7] H. Widom, solv-int/9602001.
[8] A. Kasman, Commun. Math. Phys. **172**, 427 (1995).
[9] G. Segal and G. Wilson, Pub. Math. Instit. Hautes Etudes Sci. **61**, 5 (1985).
[10] H. Aratyn, E. Nissimov, and S. Pacheva, Phys. Lett. A **201**, 293 (1995).
[11] R. Hirota in *Algebraic Analysis* (Academic, New York, 1988), p. 203.
[12] Y. Nakamura, Phys. Lett. A **195**, 346 (1994).