Solutions of the bigraded Toda hierarchy

Chuanzhong Li

Department of Mathematics, USTC, Hefei, 230026 Anhui, People’s Republic of China
and
Department of Mathematics, NBU, Ningbo, 315211 Zhejiang, People’s Republic of China

E-mail: czli@mail.ustc.edu.cn

Received 21 November 2010, in final form 9 March 2011
Published 16 May 2011
Online at stacks.iop.org/JPhysA/44/255201

Abstract
The \((N, M)\)-bigraded Toda hierarchy is an extension of the original Toda lattice hierarchy. The pair of numbers \((N, M)\) represents the band structure of the Lax matrix which has \(N\) upper and \(M\) lower diagonals, and the original one is referred to as the \((1, 1)\)-bigraded Toda hierarchy. Because of this band structure, one can introduce \(M + N - 1\) commuting flows which give a parametrization of a small phase space for a topological field theory. In this paper, first we show that there exists a natural symmetry between the \((N, M)\)- and \((M, N)\)-bigraded Toda hierarchies. We then derive the Hirota bilinear form for those commuting flows, which consist of two-dimensional Toda hierarchy, the discrete KP hierarchy and its Bäcklund transformations. We also discuss the solution structure of the \((N, M)\)-bigraded Toda equation in terms of the moment matrix defined via the wave operators associated with the Lax operator and construct some of the explicit solutions. In particular, we give the rational solutions which are expressed by the products of the Schur polynomials corresponding to the non-rectangular Young diagrams.

PACS numbers: 02.30.Ik, 02.30.Jr
Mathematics Subject Classification: 37K05, 37K10, 37K20

1. Introduction

The \((N, M)\)-bigraded Toda hierarchy, denoted by \((N, M)\)-BTH, is an integrable system (see e.g. [1, 2]), and its Lax operator is given by

\[
\mathcal{L} := \Lambda^N + u_{N-1} \Lambda^{N-1} + \cdots + u_0 + \cdots + u_{-M} \Lambda^{-M},
\]

where \(N, M \geq 1\), \(\Lambda\) is a shift operator which can be expressed as an infinite matrix in the form \(\Lambda = (E_{i,j})_{i,j\in\mathbb{Z}}\). In terms of an infinite-sized matrix, the Lax operator \(\mathcal{L}\) has the band structure with \(N\) upper and \(M\) lower nonzero diagonals. The \((N, M)\)-BTH is then defined
Here we call the \(N + M - 1\) numbers of the flows for \(n = 0\) the *primaries* of the BTH which describe the small phase space of a topological field theory (TFT), and the flows with \(n > 0\) correspond to the gravitational descendants in this TFT.

In the case of \(N = M = 1\), the \((1, 1)\)-BTH is the original Toda lattice hierarchy, and the primary flow is just the Toda lattice equation [3, 4]. The \((N, M)\)-BTH with \(N > 1\) or \(M > 1\) is then considered as an extension of the original Toda lattice hierarchy. One should also note that the case with infinite \(N\) and \(M\) corresponds to the two-dimensional Toda hierarchy where we have two independent Lax operators (one defined near infinity and the other defined near zero in the spectral space; more precisely, one considers two cases with \((1, \infty)\) and \((\infty, 1))\). Then the \((N, M)\)-BTH can naturally be considered as a reduction of the two-dimensional Toda hierarchy by imposing an algebraic relation to those two Lax operators (see [5–7]).

In [2], we showed the integrability of an extended version of the \((N, M)\)-BTH by writing the hierarchy as a bilinear identity and introduced the \(\tau\)-functions. Here the extension implies that the hierarchy has additional logarithmic flows, and this version is called the extended BTH (see [1, 8]). In this paper, we are interested in constructing several explicit solutions of the BTH, and as a first step, we consider only the non-extended version of the BTH based on our previous study [2].

The paper is organized as follows. In section 2, we give a brief summary of the original Toda lattice hierarchy whose Lax operator is given by a tridiagonal matrix. We also discuss briefly the \((2, 1)\)-BTH as an extension of the Toda hierarchy and describe the \(t_{2,0}\)-flow defined by the square root of the Lax operator. In particular, we mention that there exist nonlocal terms in the equation and remark that the flow also appears in the recent paper [9]. In section 3, we give the explicit form of the \((N, M)\)-BTH and the \(\tau\)-functions. This section is a brief summary of our previous paper [2] without the logarithmic flows. Here we express the coefficient functions in the Lax operator in terms of the \(\tau\)-functions in a similar manner to that discussed in [10, 13]. We also discuss some details of the \((2, 2)\)-BTH as an example. In section 4, we show the equivalence between the \((N, M)\)- and \((M, N)\)-BTHs using the Hirota bilinear equations (HBEs) found in [2] and the gauge transformation in [14]. To illustrate, we also give some simple concrete examples of equivalence between flows of \((2, 1)\)-BTH and \((1, 2)\)-BTH. This equivalence is explicitly shown in the examples given in section 5. In section 5, we derive the HBEs for the primaries of the \((N, M)\)-BTH. Then we construct the \(\tau\)-functions in terms of the moment matrix defined naturally via the wave operators introduced in section 3 (also see [2]). In section 6, we construct rational solutions based on the \(\tau\)-function formulas derived in the previous section. In particular, these rational solutions are given by the products of two Schur polynomials depending on two different sets of flow parameters \(t_{\alpha,n}\) and \(t_{\beta,n}\) in (1.1). Contrary to the case of the original Toda hierarchy where the rational solutions are given by the Schur polynomials of rectangular Young’s diagrams, the rational solutions of the BTH are parametrized by non-rectangular Young’s diagrams. Finally, in section 7, we summarize the results and give some discussions.

2. Tridiagonal Toda lattice hierarchy and generalization

Here we briefly explain the BTH as an extension of the original Toda equation and present some connection to the recent study in [9]. The main point is to explain the structure of an
additional symmetry generated by a fractional power of the Lax matrix. The Toda lattice equation is written in the form

\[
\begin{align*}
\frac{\partial a_{n+1}}{\partial t_1} &= a_{n+1}(b_{n+1} - b_n), \\
\frac{\partial b_n}{\partial t_1} &= a_{n+1} - a_n.
\end{align*}
\]

Equation (2.1) has the Lax representation with a tridiagonal semi-infinite matrix \( L \) given by

\[
\frac{\partial L}{\partial t_1} = [B_1, L], \quad B_1 = [L]_{\geq 0}.
\]

where \([L]_{\geq 0}\) is the upper triangular part of the matrix \( L \) given by

\[
L = \begin{pmatrix}
1 & 0 & 0 & \cdots \\
0 & 1 & 0 & \cdots \\
0 & 0 & 1 & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix}.
\]

If we consider the bounded Toda lattice equation, the Lax matrix will have finite size.

For a semi-infinite Toda equation, there exists a sequence of \( \tau \)-functions \( \{\tau_n : n \geq 0\} \) with \( \tau_0 = 1 \) defined by the \( a_n, b_n \) by the formulas

\[
a_n = \frac{\tau_n \tau_{n-2}}{\tau_{n-1}}, \quad b_n = \frac{\partial}{\partial t_1} \log \left( \frac{\tau_n}{\tau_{n-1}} \right).
\]

Then we can write the Toda lattice equation in the Hirota bilinear form:

\[
D_1^2 \tau_n \cdot \tau_n = 2\tau_n \tau_{n-1},
\]

where \( D_1 \) is the usual Hirota derivative. For the \( k \)th flow-parameter \( t_k \) of the Toda lattice hierarchy, \( D_k \) is defined by

\[
D_k f \cdot g := \left( \frac{\partial}{\partial t_k} - \frac{\partial}{\partial t_k^*} \right) f(t_k)g(t_k^*) \bigg|_{t_k = t_k^*}.
\]

The hierarchy of the Toda lattice is defined by

\[
\frac{\partial L}{\partial t_k} = [B_k, L], \quad B_k = [L^k]_{\geq 0}, \quad k = 1, 2, 3, \ldots.
\]

The \( \tau \)-functions of the Toda lattice hierarchy obey the following equations:

\[
[D_k - P_k(\hat{D})] \tau_{n+1} \cdot \tau_n = 0, \quad k = 2, 3, 4, \ldots.
\]

where the Schur polynomial \( P_k(\hat{D}) \) is defined by

\[
e^{\sum_{k=1}^N t_k \hat{D}_k z^k} = \sum_{k=0}^\infty P_k(\hat{D}) z^k, \quad \hat{D} = \left( D_1, \frac{1}{2} D_2, \frac{1}{3} D_3, \frac{1}{4} D_4, \ldots \right).
\]

When \( k = 2 \), the Hirota equation becomes

\[
(D_2 - D_1^2) \tau_{n+1} \cdot \tau_n = 0.
\]

Equation (2.10), together with equation (2.5), gives the nonlinear Schrodinger equation which can be seen as the second member of the Toda lattice hierarchy [11].

A natural question is: what is the generalized band structure of the Toda Lattice hierarchy which is just a \((1, 1)\) tridiagonal band matrix?
For example, for the \((2, 1)\) Heisenberg band structure of the Lax matrix
\[
\begin{pmatrix}
  b_1 & c_1 & 1 & 0 & \cdots \\
  a_2 & b_2 & c_2 & 1 & \cdots \\
  0 & a_3 & b_3 & c_3 & \cdots \\
  0 & 0 & a_4 & b_4 & \cdots \\
  \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix},
\] (2.11)
the equations for the Toda flow, i.e.
\[
\partial_t L = [L_{\geq 0}, L],
\] (2.12)
will lead to the following Blaszak–Marciniak lattice equation [14]:
\[
\begin{align*}
\partial_{t_1} c_n &= a_{n+2} - a_n \\
\partial_{t_1} b_n &= c_n a_{n+1} - a_n c_{n-1} \\
\partial_{t_1} a_n &= a_n (b_n - b_{n-1}).
\end{align*}
\] (2.13)
Equation (2.13) is also equivalent to the Bogoyavlensky–Narita lattice given in [15]. In [14], Blaszak and Marciniak considered the local flows which correspond to the integer powers of Lax operators. In this paper, we construct nonlocal flows using the fractional power of the Lax operator. Because of the \((2, 1)\)-band structure, one can define the square root of the Lax matrix which will be shown in detail in the next section. Using the operator \(L^{1/2}\), we give a new flow
\[
\partial_{t_2} L = [L_{\geq 0}, L^{1/2}],
\] (2.14)
which further leads to
\[
\begin{align*}
\partial_{t_2} c_n &= b_{n+1} - b_n + c_n (1 - \Lambda)(1 + \Lambda)^{-1} c_n \\
\partial_{t_2} b_n &= a_{n+1} - a_n \\
\partial_{t_2} a_n &= a_n (1 - \Lambda^{-1})(1 + \Lambda)^{-1} c_n.
\end{align*}
\] (2.15)
After denoting \(\mathcal{H}\) as \(\frac{1+\Lambda}{\Lambda-1}\), equation (2.15) can be rewritten as
\[
\begin{align*}
\partial_{t_2} c_n &= b_{n+1} - b_n + c_n \mathcal{H}^{-1} c_n \\
\partial_{t_2} b_n &= a_{n+1} - a_n \\
\partial_{t_2} a_n &= a_n \mathcal{H}^{-1} c_n.
\end{align*}
\] (2.16)
After the transformation
\[
c_n = \tilde{c}_{n+1} + \tilde{c}_n,
\] (2.17)
equation (2.15) becomes
\[
\begin{align*}
\partial_{t_2} \tilde{c}_{n+1} + \partial_{t_2} \tilde{c}_n &= b_{n+1} - b_n + \tilde{c}_n^2 - \tilde{c}_{n+1}^2 \\
\partial_{t_2} \tilde{b}_n &= a_{n+1} - a_n \\
\partial_{t_2} \tilde{a}_n &= a_n (\tilde{c}_n - \tilde{c}_{n-1}),
\end{align*}
\] (2.18)
which is just equation (40) in [15] and also related to the system (10)–(12) proposed in [17]. In [15], Svinin gives general constructions of such flows. Studying the relation between two ways of constructing nonlocal flows, i.e. the way given in [15] and that in this paper, might be an interesting problem.

Equations (2.13) and (2.16) are in fact the \(t_1\) and \(t_2\) flows of the case \(n = 3, \alpha = -1\) without center extension in [9], i.e. the solutions do not depend on the variable \(y\). In this paper, we directly introduce a new fractional Lax matrix instead of using the Casimir construction used in [9] for constructing the Lax equations. We also generalize these results to the \((N, M)\)-band matrix. For a finite-sized Lax matrix, its fractional power may not be well defined.
This leads to a difficulty to give symmetric flows generated by the fraction powers of the Lax matrix. But for a bi-infinite band matrix, one can define the fraction powers and further define other additional flows which commute with the original Toda flow. This generalization leads to the BTH which might also be seen as a general reduction of the two-dimensional Toda lattice hierarchy. In the next section, we give the continuous interpolated version of the BTH and later present the matrix version of the BTH in bi-infinite and semi-infinite band matrices.

3. The bigraded Toda hierarchy

The Lax operator of the BTH is given by the Laurent polynomial of $\Lambda$ \cite{1}:

$$L := \Lambda^N + u_{N-1} \Lambda^{N-1} + \cdots + u_0 + \cdots + u_{-M} \Lambda^{-M},$$  \hspace{1cm} (3.1)

where $N, M \geq 1$, $\Lambda$ represents the shift operator with $\Lambda := e^{\varepsilon b}$ and $\varepsilon$ is called the string coupling constant, i.e. for any function $f(x)$,

$$\Lambda f(x) = f(x + \varepsilon).$$

$L$ can be written in two different ways by dressing the shift operator

$$L = P_L \Lambda^N P_L^{-1} = P_R \Lambda^{-M} P_R^{-1},$$  \hspace{1cm} (3.2)

where the dressing operators have the form

$$P_L = 1 + w_1 \Lambda^{-1} + w_2 \Lambda^{-2} + \cdots,$$  \hspace{1cm} (3.3)

$$P_R = \tilde{w}_0 + \tilde{w}_1 \Lambda + \tilde{w}_2 \Lambda^2 + \cdots.$$  \hspace{1cm} (3.4)

Equation (3.2) is quite important because it gives the reduction condition from the two-dimensional Toda lattice hierarchy. The pair is unique up to multiplying $P_L$ and $P_R$ from the right by the operators in the form $1 + a_1 \Lambda^{-1} + a_2 \Lambda^{-2} + \cdots$ and $\tilde{a}_0 + \tilde{a}_1 \Lambda + \tilde{a}_2 \Lambda^2 + \cdots$ respectively with coefficients independent of $x$. Given any difference operator $A = \sum_k A_k \Lambda^k$, the positive and negative projections are defined by $A_+ = \sum_{k \geq 0} A_k \Lambda^k$ and $A_- = \sum_{k < 0} A_k \Lambda^k$.

To write out explicitly the Lax equations of the BTH, the fractional powers $L^{\frac{1}{N}}$ and $L^{\frac{1}{M}}$ are defined by

$$L^{\frac{1}{N}} = \Lambda + \sum_{k \geq 0} a_k \Lambda^k,$$

$$L^{\frac{1}{M}} = \sum_{k \geq -1} b_k \Lambda^k,$$

with the relations

$$(L^{\frac{1}{N}})^N = (L^{\frac{1}{M}})^M = L.$$  \hspace{1cm} (3.5)

Acting on a free function, these two fraction powers can be seen as two different local expansions around zero and infinity respectively. It was stressed that $L^{\frac{1}{N}}$ and $L^{\frac{1}{M}}$ are two different operators even if $N = M (N, M \geq 2)$ in \cite{1} due to two different dressing operators. They can also be expressed as follows:

$$L^{\frac{1}{N}} = P_L \Lambda P_L^{-1},$$

$$L^{\frac{1}{M}} = P_R \Lambda^{-1} P_R^{-1}.$$  \hspace{1cm} (3.6)

Let us now define the following operators for the generators of the BTH flows:
\[ B_{\gamma,n} := \begin{cases} L^{n+1 - \frac{\alpha}{N}} & \text{if } \gamma = \alpha = 1, 2, \ldots, N, \\ L^{n+1 + \frac{\beta}{M}} & \text{if } \gamma = \beta = -M + 1, \ldots, -1, 0. \end{cases} \] (3.5)

**Definition 3.1.** The BTH in the Lax representation is given by the set of infinite number of flows defined by

\[ \frac{\partial L}{\partial t_{\gamma,n}} = \begin{cases} \{ (B_{\alpha,n}), L \} & \text{if } \gamma = \alpha = 1, 2, \ldots, N, \\ \{- (B_{\beta,n}), L \} & \text{if } \gamma = \beta = -M + 1, \ldots, -1, 0. \end{cases} \] (3.6)

We need to remark that this kind of definition is equivalent to the definition in [1] which is just a scalar transformation about the time variables using gamma function. The original tridiagonal Toda (i.e. Kostant–Toda) hierarchy corresponds to the case with \( N = M = 1 \).

One can show [2] that the BTH in the Lax representation can be written in the equations of the dressing operators (i.e. the Sato equations).

**Theorem 3.2.** The operator \( L \) in (3.1) is a solution to the BTH (3.6) if and only if there is a pair of dressing operators \( P_L \) and \( P_R \) which satisfy the Sato equations:

\[ \partial_{\gamma,n} P_L = -(B_{\gamma,n})_{\gamma} P_L, \quad \partial_{\gamma,n} P_R = (B_{\gamma,n})_{\gamma} P_R \] (3.7)

for \( -M + 1 \leq \gamma \leq N \) and \( n \geq 0 \).

The dressing operators satisfying Sato equations will be called wave operators. By wave operators we will give the definition of the tau function for the BTH as follows.

According to [2], a function \( \tau \) depending only on the dynamical variables \( t \) and \( \epsilon \) is called the tau function of the BTH if it provides symbols related to wave operators as follows:

\[ P_L := 1 + \frac{w_1}{\lambda} + \frac{w_2}{\lambda^2} + \cdots := \frac{\tau(x, t - [\lambda^{-1}]^N; \epsilon)}{\tau(x, t; \epsilon)}, \] (3.8)

\[ P_L^{-1} := 1 + \frac{w_1'}{\lambda} + \frac{w_2'}{\lambda^2} + \cdots := \frac{\tau(x + \epsilon, t + [\lambda^{-1}]^N; \epsilon)}{\tau(x + \epsilon, t; \epsilon)}, \] (3.9)

\[ P_R := \tilde{w}_0 + \tilde{w}_1 \lambda + \tilde{w}_2 \lambda^2 + \cdots := \frac{\tau(x, t + [\lambda]^M; \epsilon)}{\tau(x, t; \epsilon)}, \] (3.10)

\[ P_R^{-1} := \tilde{w}_0' + \tilde{w}_1' \lambda + \tilde{w}_2' \lambda^2 + \cdots := \frac{\tau(x, t - [\lambda]^M; \epsilon)}{\tau(x, t; \epsilon)}, \] (3.11)

where \([\lambda^{-1}]^N\) and \([\lambda]^M\) are defined by

\[ [\lambda^{-1}]^N_{\gamma,n} := \begin{cases} \lambda^{-N(N+1)} & \gamma = N, N-1, \ldots, 1, \\ 0 & \gamma = 0, -1, \ldots, -(M-1). \end{cases} \]

\[ [\lambda]^M_{\gamma,n} := \begin{cases} 0 & \gamma = N, N-1, \ldots, 1, \\ \lambda^M(N+1) & \gamma = 0, -1, \ldots, -(M-1). \end{cases} \]

For a given pair of wave operators, the tau function is unique up to a non-vanishing function factor which is independent of \( x \) and \( t_{\gamma,n} \) with all \( n \geq 0 \) and \(-M + 1 \leq \gamma \leq N\).
Then we obtain
\[
\begin{align*}
\mathcal{P}_L &= \sum_{n=0}^{\infty} \frac{P_n(-\hat{\partial}_L)\tau(x, t; \epsilon)}{\tau(x, t; \epsilon)} \lambda^{-n}, \\
\mathcal{P}_L^{-1} &= \sum_{n=0}^{\infty} \frac{P_n(\hat{\partial}_L)\tau(x + \epsilon, t; \epsilon)}{\tau(x + \epsilon, t; \epsilon)} \lambda^{-n}, \\
\mathcal{P}_R &= \sum_{n=0}^{\infty} \frac{P_n(\hat{\partial}_R)\tau(x + \epsilon, t; \epsilon)}{\tau(x, t; \epsilon)} \lambda^n, \\
\mathcal{P}_R^{-1} &= \sum_{n=0}^{\infty} \frac{P_n(-\hat{\partial}_R)\tau(x, t; \epsilon)}{\tau(x, t; \epsilon)} \lambda^n,
\end{align*}
\]
where \(P_n\) are the elementary Schur polynomials as defined in (2.9). Here the operators \(\hat{\partial}_L\) and \(\hat{\partial}_R\) are defined by
\[
\begin{align*}
\hat{\partial}_L &= \left\{ \frac{1}{N(n + 1 - \frac{\alpha - 1}{N})} \partial_t \alpha, n : 1 \leq \alpha \leq N \right\}, \\
\hat{\partial}_R &= \left\{ \frac{1}{M(n + 1 + \frac{\beta}{M})} \partial_t \beta, n : -M + 1 \leq \beta \leq 0 \right\}.
\end{align*}
\]

The dressing operators \(\mathcal{P}_L\) and \(\mathcal{P}_R\) can be expressed by the function \(\tau(x, t; \epsilon)\):
\[
\begin{align*}
\mathcal{P}_L &= \sum_{n=0}^{\infty} \frac{P_n(-\hat{\partial}_L)\tau(x, t; \epsilon)}{\tau(x, t; \epsilon)} \Lambda^{-n}, \\
\mathcal{P}_L^{-1} &= \sum_{n=0}^{\infty} \Lambda^{-n} \frac{P_n(\hat{\partial}_L)\tau(x + \epsilon, t; \epsilon)}{\tau(x + \epsilon, t; \epsilon)}, \\
\mathcal{P}_R &= \sum_{n=0}^{\infty} \frac{P_n(\hat{\partial}_R)\tau(x + \epsilon, t; \epsilon)}{\tau(x, t; \epsilon)} \Lambda^n, \\
\mathcal{P}_R^{-1} &= \sum_{n=0}^{\infty} \Lambda^n \frac{P_n(-\hat{\partial}_R)\tau(x, t; \epsilon)}{\tau(x, t; \epsilon)}.
\end{align*}
\]

One can then find the explicit form of the coefficients \(u_i(x, t)\) of the operator \(L\) in terms of the \(\tau\)-function using equation (3.2) as \([10, 13]\)
\[
\begin{align*}
u_i(x, t) &= \frac{P_{N-i}(\hat{D}_L)\tau(x + (i + 1)\epsilon, t; \epsilon) \circ \tau(x, t; \epsilon)}{\tau(x, t; \epsilon) \tau(x + (i + 1)\epsilon, t; \epsilon)} \\
&= \frac{P_{M+i}(\hat{D}_R)\tau(x + \epsilon, t; \epsilon) \circ \tau(x + i\epsilon, t; \epsilon)}{\tau(x, t; \epsilon) \tau(x + (i + 1)\epsilon, t; \epsilon)},
\end{align*}
\]
where \(\hat{D}_L\) and \(\hat{D}_R\) are the Hirota derivatives corresponding to \(\hat{\partial}_L\) and \(\hat{\partial}_R\) respectively.

The BTH can also be written in the matrix form with the identification of the shift operator \(\Lambda\) as the infinite matrix having zero entries except 1’s in the upper diagonal elements and all other functions about \(x\) as the diagonal infinite matrix \([6]\). But now we consider only its reduction, i.e. the corresponding semi-infinite matrix form in the following. Then we rewrite the coefficient \(u_i(x, t)\) as \(u_{i,j}(t)\) and \(\tau(x + \epsilon, t)\) as \(\tau_j(t)\). We can find the semi-infinite matrix forms \(\mathcal{P}_L, \mathcal{P}_L^{-1}, \mathcal{P}_R, \mathcal{P}_R^{-1}\) corresponding to \(\mathcal{P}_L, \mathcal{P}_L^{-1}, \mathcal{P}_R, \mathcal{P}_R^{-1}\) as follows:
\[
\mathcal{P}_L = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & \cdots \\
P_1(-\hat{\partial}_L)\tau_1 & 1 & 0 & 0 & 0 & \cdots \\
P_2(-\hat{\partial}_L)\tau_2 & P_1(-\hat{\partial}_L)\tau_2 & 1 & 0 & 0 & \cdots \\
P_3(-\hat{\partial}_L)\tau_3 & P_2(-\hat{\partial}_L)\tau_3 & P_1(-\hat{\partial}_L)\tau_3 & 1 & \cdots \\
& \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix},
\]
\(\mathcal{P}_L^{-1}\), \(\mathcal{P}_R\), \(\mathcal{P}_R^{-1}\) follow as
After the following transformation \( u_{i,j} = a_{j,i+1} \), the matrix representation of \( L \) can be expressed by \((a_{i,j}),i,j \geq 1\) with

\[
a_{i,j}(t) = P_{i-j+M(\hat{D}R)}(\hat{\tau}_j \circ \hat{\tau}_{i-1}) = P_{j-i+M(\hat{D}R)}(\hat{\tau}_i \circ \hat{\tau}_{j-1}) \tag{3.21}
\]

Note here that these expressions immediately imply

\[ a_{i,j} = 0, \quad \text{if} \quad j < -M + i \quad \text{or} \quad j > N + i. \]

That is, the Lax matrix \( L \) has the \((N, M)\) band structure.

As an example, we will give some concrete results on the \((2,2)\)-BTH in the following subsection from which we can see some general pattern of the \((N, M)\)-BTH.

### 3.1. Example of the \((2,2)\)-BTH

Let us summarize this section by taking the \((2,2)\)-BTH. The Lax operator is

\[
L = \Lambda^2 + u_1 \Lambda + u_0 + u_{-1} \Lambda^{-1} + u_{-2} \Lambda^{-2}. \tag{3.22}
\]
Then there will be two different fraction powers of $L$, denoted as $L^\frac{1}{N}$ and $L^\frac{1}{M}$ respectively, of the following form:

$$ L^\frac{1}{N} = \Lambda + a_0 + a_{-1} \Lambda^{-1} + a_{-2} \Lambda^{-2} + \cdots , \quad \tag{3.23} $$

$$ L^\frac{1}{M} = a'_{-1} \Lambda^{-1} + a'_0 + a'_1 \Lambda + a'_2 \Lambda^2 + \cdots . \quad \tag{3.24} $$

We can obtain some relations of $\{a_i; i \leq 0\}, \{a'_j; j \geq -1\}$ with $\{u_i; -M \leq i \leq N - 1\}$ as follows:

$$ a_0(x) = (1 + \Lambda)^{-1} u_1(x), \quad a'_{-1} = e^{(1+\Lambda^{-1})^{-1} \log u_{-2}(x)}. \quad \tag{3.25} $$

Then by the Lax equation, we obtain the $t_{2,0}$ flow of the $(2,2)$-BTH:

$$ \partial_{t_{2,0}} L = [\Lambda + (1 + \Lambda)^{-1} u_1(x), L], \quad \tag{3.26} $$

which corresponds to

$$ \begin{align*}
\partial_{2,0} u_1(x) &= u_1(x + \epsilon) - u_0(x) + u_1(x)(1 - \Lambda)(1 + \Lambda)^{-1} u_1(x) \\
\partial_{2,0} u_0(x) &= u_{-1}(x + \epsilon) - u_{-1}(x) \\
\partial_{2,0} u_{-1}(x) &= u_{-2}(x + \epsilon) - u_{-2}(x) + u_{-1}(x)(a_0(x) - a_0(x) - 2\epsilon) \\
\partial_{2,0} u_{-2}(x) &= u_{-2}(x)(\epsilon) = u_0(x + \epsilon) - u_0(x) + u_1(x)(a_0(x) - a_0(x) + 2\epsilon).
\end{align*} \tag{3.27} $$

From equations (3.27), we find that the equations have infinite terms because of $(1 + \Lambda)^{-1}$ which comes from the fraction power of the Lax operator. Just like the method in [17], for avoiding infinite sums, we use the auxiliary function $a_0(x)$ with which we can rewrite equation (3.27) as

$$ \begin{align*}
\partial_{2,0} u_1(x) &= u_1(x + \epsilon) - u_0(x) + u_1(x)(a_0(x) - a_0(x + \epsilon)) \\
\partial_{2,0} u_0(x) &= u_{-1}(x + \epsilon) - u_{-1}(x) \\
\partial_{2,0} u_{-1}(x) &= u_{-2}(x + \epsilon) - u_{-2}(x) + u_{-1}(x)(a_0(x) - a_0(x) - 2\epsilon) \\
\partial_{2,0} u_{-2}(x) &= u_{-2}(x)(\epsilon) = u_0(x + \epsilon) - u_0(x) + u_1(x)(a_0(x) - a_0(x) + 2\epsilon).
\end{align*} \tag{3.28} $$

The $t_{1,0}$ flow will have finite terms as follows because it does not use the fraction power of the Lax operator $L$:

$$ \partial_{1,0} L = [\Lambda^2 + u_1 \Lambda + u_0, L], \quad \tag{3.29} $$

which correspond to

$$ \begin{align*}
\partial_{1,0} u_1(x) &= u_{-1}(x + 2\epsilon) - u_{-1}(x) \\
\partial_{1,0} u_0(x) &= u_{-2}(x + 2\epsilon) - u_{-2}(x) + u_{-1}(x)(1 - \epsilon) u_1(x - \epsilon) \\
\partial_{1,0} u_{-1}(x) &= u_{-1}(x)(\epsilon) = u_1(x + \epsilon) - u_1(x) + u_0(x)(a_0(x) - a_0(x) - 2\epsilon) \\
\partial_{1,0} u_{-2}(x) &= u_{-2}(x)(\epsilon) = u_0(x + \epsilon) - u_0(x) + u_1(x)(a_0(x) - a_0(x) + 2\epsilon).
\end{align*} \tag{3.30} $$

For the $t_{-1,0}$ flow, equations will also be complicated because of another fraction power of $L$. The equation is

$$ \partial_{-1,0} L = -[e^{(1+\Lambda^{-1})^{-1} \log u_{-2}(\Lambda^{-1})}, L], \quad \tag{3.31} $$

which corresponds to

$$ \begin{align*}
\partial_{-1,0} u_1(x) &= e^{(1+\Lambda^{-1})^{-1} \log u_{-2}(x + 2\epsilon) - e^{(1+\Lambda^{-1})^{-1} \log u_{-2}(x)} \\
\partial_{-1,0} u_0(x) &= e^{(1+\Lambda^{-1})^{-1} \log u_{-2}(x + \epsilon) - e^{(1+\Lambda^{-1})^{-1} \log u_{-2}(x)} u_1(x - \epsilon) \\
\partial_{-1,0} u_{-1}(x) &= e^{(1+\Lambda^{-1})^{-1} \log u_{-2}(x)}(u_0(x) - u_0(x - \epsilon)) \\
\partial_{-1,0} u_{-2}(x) &= u_{-1}(x)e^{(1+\Lambda^{-1})^{-1} \log u_{-2}(x - \epsilon) - e^{(1+\Lambda^{-1})^{-1} \log u_{-2}(x)}u_{-1}(x - \epsilon)}.
\end{align*} \tag{3.32} $$
From equations (3.32), we find that the equations have finite terms but every term has infinite multiplication because of the factor \( e^{i(x)} \). For each \( a \), the Lax operator. Similarly for avoiding infinite multiplication, we use the auxiliary function \( \lambda \) and the constraints (3.2).

\[
\begin{aligned}
\partial_{-1,0}u_1(x) &= a_{-1}(x + 2\epsilon) - a_{-1}(x) \\
\partial_{-1,0}u_0(x) &= u_1(x) a'_{-1}(x + \epsilon) - a'_{-1}(x) u_1(x - \epsilon) \\
\partial_{-1,0}u_{-1}(x) &= a'_{-1}(x) (u_0(x) - u_0(x - \epsilon)) \\
\partial_{-1,0}a_{-1}(x) &= u_{-1}(x) a'_{-1}(x - \epsilon) - a'_{-1}(x) u_{-1}(x - \epsilon) \\
\partial_{-1,0}a_{-1}(x) &= u_{-1}(x). 
\end{aligned}
\]

(33)

Infinite sums or infinite multiplications are important properties of the BTH because of nonlocal operators. For a finite Lax matrix, it is not easy to construct its fraction power. That is why we do not use the fraction power of a finite matrix to give Lax equations.

4. Equivalence between the \((N, M)\)-BTH and the \((M, N)\)-BTH

In this section, we prove that there is an equivalence between the \((N, M)\)-BTH and the \((M, N)\)-BTH in three ways. One is to prove the equivalence in the Hirota bilinear identities based on [2]. The second one is for equivalence in specific HBEs. Finally, we will prove the equivalence between their Lax equations using transformation in [14]. To see the equivalence clearly, one explicit example, i.e. the equivalence between the \((1, 2)\)-BTH and the \((2, 1)\)-BTH in Lax equations under transformation will be shown in detail.

4.1. Equivalence in the Hirota bilinear identities

Firstly after denoting \( \tau(x, t) \) as \( \tau(x - \frac{\lambda}{2}, t) \), we obtain the following Hirota bilinear identity [2], i.e. for each \( m \in \mathbb{Z}, r \in \mathbb{N} \):

\[
\text{Res}_a \{ \lambda^{N+r+1} \tau(x, t - [\lambda^{-1}]^N) \times \tau(x - (m - 1)\epsilon, t' + [\lambda^{-1}]^N) e^{-\xi_L(\lambda, t - t')} \} = \text{Res}_a \{ \lambda^{M+r+1} \tau(x + \epsilon, t + [\lambda]^M) \times \tau(x - m\epsilon, t' - [\lambda]^M) e^{-\xi_R(\lambda, t - t')} \},
\]

(4.1)

where

\[
\xi_L(\lambda, t) = \sum_{n \geq 0} \sum_{a=1}^N \lambda^{N(a-1)+\frac{a+1}{2}} t_{a,n}, \quad \xi_R(\lambda, t) = \sum_{n \geq 0} \sum_{\beta=-M+1}^0 \lambda^{M(a+1)+\frac{\beta}{2}} t_{\beta,n}.
\]

The most important property that the BTH has is that there are \( Nr \) and \( Mr \) on both sides of HBEs (4.1). These two terms show the principal difference of the BTH from the two-dimensional Toda hierarchy, i.e. the constraint (3.2).

The Hirota bilinear identity (4.1) can lead to the following identity under the transformation \( m \mapsto -m, x \mapsto x - m\epsilon, \lambda \mapsto \lambda^{-1} \):

\[
\text{Res}_a \{ \lambda^{-N+r+1} \tau(x - m\epsilon, t - [\lambda]^N) \times \tau(x + \epsilon, t' + [\lambda]^N) e^{\xi_L(\lambda, t - t')} \} = \text{Res}_a \{ \lambda^{-M+r+1} \tau(x - (m - 1)\epsilon, t + [\lambda^{-1}]^M) \tau(x, t' - [\lambda^{-1}]^M) e^{-\xi_R(\lambda^{-1}, t - t')} \},
\]

(4.2)

where

\[
\xi'_L(\lambda, t - t') = \sum_{n \geq 0} \sum_{a=1}^N \lambda^{-N(a-1)+\frac{a+1}{2}} (t_{a,n} - t'_{a,n}),
\]

\[
\xi'_R(\lambda^{-1}, t - t') = \sum_{n \geq 0} \sum_{\beta=-M+1}^0 \lambda^{M(a+1)+\frac{\beta}{2}} (t_{\beta,n} - t'_{\beta,n}).
\]
Equation (4.2) can be rewritten as the following identity after the interchange of $t$ and $t'$:

$$\text{Res}_s \{ \lambda^{-1} \lambda^{-N \epsilon_m} \tau(x + \epsilon, t + [\lambda]N) \times \tau(x - m \epsilon, t' - [\lambda]N) e^{i \xi (t, t - \tau)} \} = \text{Res}_s \{ \lambda^{-1} \lambda^{-N \epsilon_m} \tau(x, t - [\lambda]N) \times \tau(x - (m - 1) \epsilon, t' + [\lambda]N) e^{-i \xi (t, t - \tau)} \},$$

(4.3)

which can further be rewritten as follows:

$$\text{Res}_s \{ \lambda^{-1} \lambda^{M \epsilon_m} \tau(x, t - [\lambda]N) \times \tau(x - (m - 1) \epsilon, t' + [\lambda]N) e^{i \xi (t, t - \tau)} \} = \text{Res}_s \{ \lambda^{-1} \lambda^{-N \epsilon_m} \tau(x + \epsilon, t + [\lambda]N) \times \tau(x - m \epsilon, t' - [\lambda]N) e^{-i \xi (t, t - \tau)} \}.$$

(4.4)

Equation (4.4) is obviously the $(M, N)$-BTH compared to equation (4.1) if we change the time variable $t, \gamma, n$, i.e., the subscript $L \leftrightarrow R$. Therefore for the $(M, N)$-BTH, equation (4.4) is in fact changed into the following equation under the transformation $t, t' \rightarrow t - \gamma, n$:

$$\text{Res}_s \{ \lambda^{-1} \lambda^{M \epsilon_m} \tau(x, t - [\lambda]N) \times \tau(x - (m - 1) \epsilon, t' + [\lambda]N) e^{i \xi (t, t - \tau)} \} = \text{Res}_s \{ \lambda^{-1} \lambda^{-N \epsilon_m} \tau(x + \epsilon, t + [\lambda]N) \times \tau(x - m \epsilon, t' - [\lambda]N) e^{-i \xi (t, t - \tau)} \}.$$

(4.5)

Because the transforms $m \mapsto -m, \ x \mapsto x - m \epsilon, \lambda \mapsto \lambda^{-1}$ do not change the equation itself, we can say that the $(N, M)$-BTH is equivalent to the $(M, N)$-BTH under the transformation $t, t' \rightarrow t - \gamma, n$.

4.2. Equivalence in the HBEs

The bilinear identity (4.1) of the BTH can be equivalently expressed as [2]

$$\text{Res}_s \{ \lambda^{N \epsilon_m - 1} \tau_j - (m - 1) \} (t + y + [\lambda]N) \tau_j (t - y - [\lambda]N) e^{i \xi (t, y - \tau)} \} = \text{Res}_s \{ \lambda^{-M \epsilon_m - 1} \tau_{j+1} (t + y + [\lambda]N) \tau_{j+1} (t - y - [\lambda]N) e^{i \xi (t, y - \tau)} \}.$$

(4.6)

To be specific, we will give the equivalence from the concrete Hirota equations between the $(N, M)$-BTH and the $(M, N)$-BTH as follows, which will also be used to derive specific primary Hirota equations of the BTH in the next section.

For the $(N, M)$-BTH, in terms of \( \prod \gamma_{\alpha_{i_1} i_1} k^i_{\beta_{i_1} i_1} \cdots \gamma_{\alpha_{i_t} i_t} k^i_{\beta_{i_t} i_t} \prod \gamma_{\beta_{i_1} i_1} k^i_{\beta_{i_1} i_1} \cdots \gamma_{\beta_{i_t} i_t} k^i_{\beta_{i_t} i_t}\), $-M + 1 \leq \beta_i \leq 1$, $1 \leq \alpha_i \leq N$, the Hirota equation is written as

\[
\prod_{i=1}^s \left( -D_{\alpha_i i_1} k^i \right) \left( \sum_{i_0=0}^{k_i} \cdots \sum_{i_q=0}^{k_q} \prod_{p=1}^p \frac{2^{i_q-i_p}}{(k^i_i-i^p_p)!} P_{\sum q=1}^p (M_{q+1}+1) \frac{k^i_{\beta_{i_q} i_q}}{2^q} \right) \tau_n \tau_{n-m} \]

\[
= \prod_{i=1}^s \left( \sum_{i_0=0}^{k_i} \cdots \sum_{i_q=0}^{k_q} \prod_{p=1}^p \frac{2^{i_q-i_p}}{(k^i_i-i^p_p)!} P_{\sum q=1}^p (M_{q+1}+1) \frac{k^i_{\beta_{i_q} i_q}}{2^q} \right) \tau_n \tau_{n-m} \]

(4.7)

where $P_k$ are the Schur polynomials as defined in (2.9). After the transformation $m \mapsto -m, n \mapsto n - m$, the identity (4.7) becomes
\[ \prod_{i=1}^{p} \left( \sum_{i'_{q}\leq 0}^{p} \prod_{i=0}^{q-1} \frac{(-D_{(\alpha_{q}, l_{q})}^{+})^{y_{q}}}{\ell_{q}'} \right) \prod_{i=0}^{q-1} \frac{(-D_{(\beta_{q}, l_{q})}^{+})^{y_{q}}}{\ell_{q}'} \prod_{i=0}^{q-1} \frac{(-D_{(\gamma_{q}, l_{q})}^{+})^{y_{q}}}{\ell_{q}'} \frac{P_{\sum_{s=M(l_{q}+1)+1}^{k_{q}}}(\ell_{q}-\ell_{q}')^{M}N_{r}}{P_{\sum_{s=M(l_{q}+1)+1}^{k_{q}}}(\ell_{q}-\ell_{q}')^{M}N_{r}} \right) \]
One can choose the Lax operator for the anti-involution map \( \tilde{u_j}(x) \) of the \((M, N)\) which satisfies the condition
\[
L^\dagger = L^\dagger M u_{-M}(x) + \Lambda^{-1} u_{-M+1}(x) + \cdots + \Lambda^{-N+1} u_{N-1}(x) + \Lambda^{-N}.
\]

By calculation, one can prove the following two lemmas as in [14].

Lemma 4.2. For \(-M + 1 \leq \gamma \leq N, n \geq 0\), the following identity holds for any integer \(k\):
\[
((B_{\gamma, n})^k) = ((B_{\gamma, n})^k)_{\leq -k}.
\]

Using the above lemma, we can prove the following lemma directly.

Lemma 4.3. \( \partial_{\gamma,n} L = [(B_{\gamma,n})^0, L] \) can lead to \( \partial_{\gamma,n} L^\dagger = [(B_{\gamma,n})^1, L^\dagger] \), and \( \partial_{\gamma,n} L = [(B_{\gamma,n})^2, L^\dagger] \).

By the above two lemmas and proposition 4.1, we can prove the following important theorem.

Theorem 4.4. The Lax equation (3.6) of the \((N, M)\)-BTH with the Lax operator
\[
L_{N,M} = \Lambda^N + u_{N-1} \Lambda^{N-1} + \cdots + u_0 + u_{-M} \Lambda^{-M}
\]
is equivalent to the \((M, N)\)-BTH with the Lax operator
\[
\tilde{L}_{N,M} = \Lambda^M + \tilde{u}_{M-1} \Lambda^{M-1} + \cdots + \tilde{u}_0 + \cdots + \tilde{u}_{-N} \Lambda^{-N}
\]
under the Miura map: \( \tilde{u}_j(x, t) = u_{-j}(x + j \epsilon) e^{\frac{j \Lambda^M u_{M}(x,t)}{\Lambda^M}} \).

Proof. For the Lax operator \( \mathcal{L}_{N,M} = \Lambda^N + u_{N-1} \Lambda^{N-1} + \cdots + u_0 + \cdots + u_{-M} \Lambda^{-M} \) of the \((N, M)\)-BTH which satisfies Lax equations (3.6), one can choose \( \Psi = \Psi(u, t) = e^{(1 - \Lambda^{-M})^{-1} u_{-M}(x,t)} \) which satisfies the condition
\[
\partial_{\gamma,n} \Psi = (B_{\gamma,n})_0 \Psi, \quad -M + 1 \leq \gamma \leq N, \quad n \geq 0.
\]
Then \( \tilde{L}_{N,M} = \Psi^{-1} \mathcal{L}_{N,M} \Psi = \tilde{u}_{M} \Lambda^M + \tilde{u}_{M-1} \Lambda^{M-1} + \cdots + \tilde{u}_0 + \cdots + \Lambda^{-M} \) satisfies the hierarchy
\[
\partial_{\gamma,n} \tilde{L}_{N,M} = [(B_{\gamma,n})^1, \tilde{L}_{N,M}], \quad -M + 1 \leq \gamma \leq N, \quad n \geq 0,
\]
where
\[
\tilde{B}_{\gamma,n} = \Psi^{-1} B_{\gamma,n} \Psi, \quad \tilde{u}_j(x) = \Psi^{-1}(x) u_j(x) \Psi(x + i \epsilon).
\]

Using lemma 4.3, one can derive
\[
\partial_{\gamma,n} \mathcal{L}_{N,M} = [(B_{\gamma,n})_0, \mathcal{L}_{N,M}^\dagger], \quad -M + 1 \leq \gamma \leq N, \quad n \geq 0.
\]

One can choose the Lax operator \( \tilde{L}_{N,M} = \mathcal{L}_{N,M}^\dagger = \Lambda^M + \tilde{u}_{M-1} \Lambda^{M-1} + \cdots + \tilde{u}_0 + \cdots + \tilde{u}_{-N} \Lambda^{-N} \) of the \((M, N)\)-BTH with
\[
\tilde{u}_j(x) = \tilde{u}_{-j}(x + j \epsilon) = \Psi^{-1}(x + j \epsilon) u_{-j}(x + j \epsilon) \Psi(x) = u_{-j}(x + j \epsilon) e^{\frac{j \Lambda^M u_{M}(x,t)}{\Lambda^M}}.
\]
The \((M, N)\)-BTH with this Lax matrix will be equivalent to the original \((N, M)\)-BTH.

To illustrate, we give some simple concrete examples of equivalence in the next subsection which include nonlocal flows of the \((2, 1)\)-BTH and the \((1, 2)\)-BTH. We also consider the interpolated form of the BTH. At this time, we use a powerful tool called gauge transformation to prove that equivalence.
4.4. Equivalence between the (1, 2)-BTH and the (2, 1)-BTH

Using proposition 4.1, we see the equivalence between the (1, 2)-BTH and the (2, 1)-BTH in detail. Here we give only the primary flows of them to see the equivalence.

(1, 2)-BTH. The Lax operator of the (1, 2)-BTH is as follows:
\[ L_{1,2} = \Lambda + u_0 + u_1 \Lambda^{-1} + u_{-2} \Lambda^{-2}. \] (4.20)

The (1, 2)-BTH has the following primary equations:
\[ \partial_{1,0} L_{1,2} = [\Lambda + u_0, L_{1,2}], \] (4.21)
and
\[ \partial_{-1,0} L_{1,2} = -[e^{(1+\Lambda^{-1})^{-1}\log u_{-2} \Lambda^{-1}}, L_{1,2}], \] (4.22)
which further lead to
\[
\begin{align*}
\partial_{1,0} u_0(x) &= e^{(1+\Lambda^{-1})^{-1}\log u_{-2}(x+\epsilon)} - e^{(1+\Lambda^{-1})^{-1}\log u_{-2}(x)}, \\
\partial_{-1,0} u_{-1}(x) &= e^{(1+\Lambda^{-1})^{-1}\log u_{-2}(x)}(u_0(x) - u_0(x - \epsilon)), \\
\partial_{-1,0} u_{-2}(x) &= u_{-1}(x)e^{(1+\Lambda^{-1})^{-1}\log u_{-2}(x-\epsilon)} - e^{(1+\Lambda^{-1})^{-1}\log u_{-2}(x)}u_{-1}(x - \epsilon),
\end{align*}
\]
and
\[
\begin{align*}
\partial_{1,0} u_0(x) &= u_{-1}(x + \epsilon) - u_{-1}(x), \\
\partial_{1,0} u_{-1}(x) &= u_{-2}(x + \epsilon) - u_{-2}(x) + u_{-1}(x)(u_0(x) - u_0(x - \epsilon)), \\
\partial_{1,0} u_{-2}(x) &= u_{-2}(x)(u_0(x) - u_0(x - 2\epsilon)).
\end{align*}
\] (4.23)

(2, 1)-BTH. The Lax operator of the (2, 1)-BTH is as follows:
\[ L_{2,1} = \Lambda^2 + \tilde{u}_1 \Lambda + \tilde{u}_0 + \tilde{u}_{-1} \Lambda^{-1}. \] (4.25)

Equations (3.6) in this case are as follows:
\[ \partial_{2,0} L_{2,1} = [\Lambda + (1 + \Lambda)^{-1}\tilde{u}_1(x), L_{2,1}], \] (4.26)
\[ \partial_{1,0} L_{2,1} = [\Lambda^2 + \tilde{u}_1 \Lambda + \tilde{u}_0, L_{2,1}], \] (4.27)
which further lead to the following concrete equations:
\[
\begin{align*}
\partial_{2,0} \tilde{u}_1(x) &= \tilde{u}_1(x + \epsilon) - \tilde{u}_1(x) + \tilde{u}_1(x)(1 - \Lambda)(1 + \Lambda)^{-1}\tilde{u}_1(x), \\
\partial_{2,0} \tilde{u}_0(x) &= \tilde{u}_{-1}(x + \epsilon) - \tilde{u}_{-1}(x), \\
\partial_{2,0} \tilde{u}_{-1}(x) &= \tilde{u}_{-1}(x)(1 - \Lambda^{-1})(1 + \Lambda)^{-1}\tilde{u}_1(x),
\end{align*}
\] (4.28)
\[
\begin{align*}
\partial_{1,0} \tilde{u}_1(x) &= \tilde{u}_{-1}(x + 2\epsilon) - \tilde{u}_{-1}(x), \\
\partial_{1,0} \tilde{u}_0(x) &= \tilde{u}_{-2}(x + 2\epsilon) - \tilde{u}_{-2}(x) + \tilde{u}_1(x)\tilde{u}_{-1}(x + \epsilon) - \tilde{u}_{-1}(x)\tilde{u}_1(x - \epsilon), \\
\partial_{1,0} \tilde{u}_{-1}(x) &= \tilde{u}_{-1}(x)(\tilde{u}_0(x) - \tilde{u}_0(x - \epsilon)).
\end{align*}
\] (4.29)

It seems that equations (4.28) and (4.29) are quite different from equations (4.23) and (4.24) respectively. In fact, after performing gauge transformation on the (2, 1)-BTH as follows, we find the equivalent relation between these equations.

Now we consider that the function \( \phi \) has the form
\[ \phi = e^{(1-\Lambda^{-1})^{-1}\log \tilde{u}_{-1}(x)}, \] (4.30)
then \( \tilde{L}_{2,1} := \phi^{-1} L_{2,1} \phi \) will have the form
\[ \tilde{L}_{2,1} = v_2 \Lambda^2 + v_1 \Lambda + v_0 + \Lambda^{-1}. \] (4.31)

The relations between \( v_i \) \( (0 \leq i \leq 2) \) and \( \tilde{u}_i \) \( (-1 \leq i \leq 1) \) are as follows:
\[ v_2 = \phi^{-1} \phi(x + 2\epsilon), \quad v_1 = \phi^{-1} \phi(x + \epsilon), \quad v_0 = u_0. \] (4.32)
Therefore we obtain the following new flows on the new Lax operator $\hat{L}_{2,1}$ using proposition 4.1:

$$\partial_{2,0}\hat{L}_{2,1} = [e^{(1+\Lambda)^{-1}\log v_2(x)}\Lambda, \hat{L}_{2,1}],$$

$$\partial_{1,0}\hat{L}_{2,1} = [v_2\Lambda^2 + v_1\Lambda^1, \hat{L}_{2,1}],$$

which further leads to

$$\begin{cases}
\partial_{2,0}v_0(x) = e^{(1+\Lambda)^{-1}\log v_2(x)} - e^{(1+\Lambda)^{-1}\log v_2(x-\epsilon)}, \\
\partial_{2,0}v_1(x) = e^{(1+\Lambda)^{-1}\log v_2(x)}(v_0(x + \epsilon) - v_0(x)), \\
\partial_{2,0}v_2(x) = v_1(x + \epsilon) e^{(1+\Lambda)^{-1}\log v_2(x)} - e^{(1+\Lambda)^{-1}\log v_2(x+\epsilon)} v_1(x),
\end{cases}$$

and

$$\begin{cases}
\partial_{1,0}v_0(x) = v_1(x) - v_1(x - \epsilon), \\
\partial_{1,0}v_1(x) = v_2(x) - v_2(x - \epsilon) + v_1(x)(v_0(x + \epsilon) - v_0(x)), \\
\partial_{1,0}v_2(x) = v_2(x)(v_0(x + 2\epsilon) - v_0(x)).
\end{cases}$$

Comparing equations (4.35) and (4.36) with equations (4.23) and (4.24), we can find that these two pairs of flows are equivalent under the Miura map $u_j = v_{-j}(x + j\epsilon)$ ($0 \leq j \leq 2$). This proves the equivalence between the (1, 2)-BTH and the (2, 1)-BTH.

Using the bilinear identities obtained by comparing every term of the Hirota bilinear identities of the BTH in this section, in the next section we will derive all primary Hirota equations of the BTH to see its inner structure.

5. Hirota equations and solutions of the BTH

HBEs are the central object in Sato theory. From HBEs, we can derive the structure of the solution. This is a great motivation for us to consider the HBEs of the BTH. The HBEs of the BTH can be derived from (4.6) which comes from the HBEs in [2]. In particular, the following proposition will list all the Hirota equations for the primary variables, i.e. $t_{r,n}$ with $n = 0$.

**Proposition 5.1.** For the $(N, M)$-BTH, we have the following identities for primary derivatives which are equivalent to all the primary Hirota equations:

$$D_{\beta,0} - P_{M+\beta}(\hat{D}_K) \tau_{n+1} \circ \tau_n = 0,$$

$$D_{\beta,0}D_{-M+1,0} - 2P_{M+\beta+1}(\hat{D}_K) \tau_n \circ \tau_n = 0,$$

$$D_{\beta,0}D_{N,0} \tau_n \circ \tau_n = 2P_{M+\beta-1}(\hat{D}_K) \tau_{n+1} \circ \tau_{n-1},$$

$$D_{\alpha,0}D_{-M+1,0} \tau_n \circ \tau_n = 2P_{N-\alpha}(\hat{D}_L) \tau_{n+1} \circ \tau_{n-1},$$

$$D_{\alpha,0}D_{N,0} - 2P_{N-\alpha+2}(\hat{D}_L) \tau_n \circ \tau_n = 0,$$

$$D_{\alpha,0} - P_{N-\alpha+1}(\hat{D}_L) \tau_{n+1} \circ \tau_n = 0,$$

where $P_k$ are the Schur polynomials as defined in (2.9).

**Proof.** In equation (4.6), for the term $\lambda_{\beta,1}^k$, $-M + 1 \leq \beta \leq 0$, in the $(N, M)$-BTH, the following Hirota equation holds:

$$\left( \sum_{i=0}^{k} \frac{(-D_{\beta,1})^i}{i!} \frac{2^{k-i}}{(k-i)!} P_{M(i+1)}(\hat{D}_K) \right) \tau_{n+1} \tau_{n-m} = \frac{(D_{\beta,1})^k}{k!} P_{N-m}(\hat{D}_L) \tau_{n-m+1} \tau_n,$$

which can also be obtained from equation (4.7).
For $y^k_{\alpha,l}, 1 \leq \alpha \leq N$, in the $(N, M)$-BTH, the following Hirota equation holds:

$$
\left( \sum_{i=0}^{k} \frac{(-D_{\alpha,l})^i}{i!} \frac{2^{-k+i}}{(k-i)!} P_N(i+\frac{1-\alpha}{2})N_{\alpha\alpha \alpha}(\hat{D}_L) \right) \tau_{n-m+1} \tau_n = \frac{(D_{\alpha,l})^k}{k!} P_{M \alpha \alpha}(\hat{D}_R) \tau_{n+1} \tau_{n-m}.
$$

(5.8)

For $y_{\beta,0}, -M + 1 \leq \beta \leq 0$, in the $(N, M)$-BTH, the following Hirota equation holds:

$$
\left( \sum_{i=0}^{1} \frac{(-D_{\beta,0})^i}{i!} \frac{1}{(1-i)!} P_{M(i+\frac{1+\beta}{2})M \alpha \alpha \alpha}(\hat{D}_R) \right) \tau_{n+1} \tau_{n-m} = D_{\beta,0} P_{N \alpha \alpha \alpha}(\hat{D}_L) \tau_{n+1} \tau_{n-m+1} \tau_n.
$$

which is

$$(2P_{M(1+\frac{1+\beta}{2})M \alpha \alpha \alpha}(\hat{D}_R) - D_{\beta,0} P_{M \alpha \alpha \alpha}(\hat{D}_R)) \tau_{n+1} \tau_{n-m} = D_{\beta,0} P_{N \alpha \alpha \alpha}(\hat{D}_L) \tau_{n+1} \tau_{n-m+1} \tau_n.$$

Setting $r = 0$ and $m = -1$, for the term with $y_{\beta,0}, -M + 1 \leq \beta \leq 0$, we obtain

$$
(2P_{M(1+\frac{1+\beta}{2})M \alpha \alpha \alpha}(\hat{D}_R) - D_{\beta,0} P_{M \alpha \alpha \alpha}(\hat{D}_R)) \tau_{n+1} \tau_{n+1} \tau_n = 0.
$$

(5.9)

When $r = 0$ and $m = 1$, for the term with $y_{\beta,0}, -M + 1 \leq \beta \leq 0$, in the $(N, M)$-BTH, we obtain

$$
(P_{M \alpha \alpha \alpha}(\hat{D}_R) - D_{\beta,0} P_{N \alpha \alpha \alpha}(\hat{D}_L)) \tau_{n+1} \tau_n = 0.
$$

Setting $r = 0$ and $m = 1$, for the term $y_{\beta,0}, -M + 1 \leq \beta \leq 0$, in the $(N, M)$-BTH, we obtain

$$
2P_{M \alpha \alpha \alpha}(\hat{D}_R) \tau_{n+1} \tau_{n-1} = D_{\beta,0} D_{N \alpha \alpha \alpha}(\hat{D}_L) \tau_n.
$$

For the term $y_{\alpha,0}, 1 \leq \alpha \leq N$, in the $(N, M)$-BTH, the following Hirota equation holds:

$$
\left( \sum_{i=0}^{1} \frac{(-D_{\alpha,0})^i}{i!} \frac{1}{(1-i)!} P_{N(1+\frac{1-\alpha}{2})N \alpha \alpha \alpha}(\hat{D}_L) \right) \tau_{n-m+1} \tau_n = D_{\alpha,0} P_{M \alpha \alpha \alpha}(\hat{D}_R) \tau_{n+1} \tau_{n-m}.
$$

(5.10)

which is

$$(2P_{N(1+\frac{1-\alpha}{2})N \alpha \alpha \alpha}(\hat{D}_L) - D_{\alpha,0} P_{M \alpha \alpha \alpha}(\hat{D}_R)) \tau_{n-m+1} \tau_n = D_{\alpha,0} P_{M \alpha \alpha \alpha}(\hat{D}_L) \tau_{n+1} \tau_{n-m}.$$

For $r = 0$ and $m = -1$, it further leads to

$$
2P_{N(1-\frac{1-\alpha}{2})N \alpha \alpha \alpha}(\hat{D}_L) \tau_{n+1} \tau_n = D_{\alpha,0} D_{M \alpha \alpha \alpha}(\hat{D}_R) \tau_{n+1} \tau_n.
$$

Setting $r = 0$ and $m = 0$, for the term with $y_{\alpha,0}, 1 \leq \alpha \leq N$, in the $(N, M)$-BTH, the following equation succeeds:

$$
(P_{N(1-\frac{1-\alpha}{2})N \alpha \alpha \alpha}(\hat{D}_L) - D_{\alpha,0} \alpha \alpha \alpha \alpha)(\hat{D}_L) \tau_{n+1} \tau_n = 0.
$$

Setting $r = 0$ and $m = 1$, we obtain the equation

$$
\left( \frac{1}{2} D_{\alpha,0} D_{N \alpha \alpha \alpha}(\hat{D}_L) - P_{N(1-\alpha)N \alpha \alpha \alpha}(\hat{D}_R) \right) \tau_n \tau_n = 0.
$$

We can obtain the primary Hirota equations from the value of $k = 0, 1$. The other values of $k$ will give higher order derivatives of the hierarchy.

When $k = 0$, equations (5.7) and (5.8) give

$$
P_{N \alpha \alpha \alpha}(\hat{D}_L) \tau_{n-m+1} \tau_n = P_{M \alpha \alpha \alpha}(\hat{D}_R) \tau_{n+1} \tau_{n-m}.
$$

(5.11)

When $r = 1$ and $m = 0$, equation (5.11) becomes

$$
P_n(\hat{D}_L) \tau_{n+1} \tau_n = P_M(\hat{D}_R) \tau_{n+1} \tau_n.
$$

(5.12)
When \( r = 1 \) and \( m = -1 \), equation (5.11) becomes
\[
P_{N-1}(\hat{D}_L)\tau_n\tau_n = P_{M+1}(\hat{D}_R)\tau_{n+1}\tau_{n+1}.
\] (5.13)

When \( r = 1 \) and \( m = 1 \), we obtain
\[
P_{N+1}(\hat{D}_L)\tau_n\tau_n = P_{M-1}(\hat{D}_R)\tau_{n+1}\tau_{n+1}.
\] (5.14)

The three equations (5.12), (5.13) and (5.14) can be derived from equations (5.1)–(5.4). These equations discussed above are all the primary Hirota equations. The other values of \( r \) and \( m \) will include higher derivatives of the hierarchy.

From these primary Hirota equations, we can see that the BTH has ample structure information. We can say that the BTH contains a discrete KP equation (from equations (5.2) and (5.5)), NLS equations (from equations (5.1) and (5.6)) and a two-dimensional Toda lattice equation (from equations (5.3) and (5.4)). Comparing with the Hirota equations of the two-dimensional Toda hierarchy, we find that the BTH has more constraints on equations which come from the equivalence of \( \partial_{n_1} \) and \( \partial_{n_2} \). This information helps us to obtain the solution of the BTH which will be given in the next subsection. From all the primary Hirota equations mentioned above, we can obtain that the solution of the BTH should have a double-Wronskian structure [17]. If we impose the vanishing of the \( \partial_t \) derivatives on tau functions, the molecule equation in [17] in fact becomes our \((2,1)\)-BTH. Using the same method as in [17], we proved that the double-Wronskian structure satisfies all the primary Hirota equations of the \((N,M)\)-BTH. But later we found that there is another much simpler way to obtain the structure naturally which will be mentioned in the next subsection. So we prefer this simpler way instead of that in [17].

### 5.1. Solutions of the BTH in the semi-infinite matrix

Now we construct the tau function for the BTH in the semi-infinite matrix representation, that is, \((a_{i,j})_{i,j \geq 1}\). In order to do this, we first introduce the wave operators \( W_L \) and \( W_R \) associated with the dressing operators \( P_L \) and \( P_R \):
\[
W_L(x,t,\Lambda) = P_L(x,t,\Lambda) \circ \exp \left( \sum_{n \geq 0} \sum_{\alpha = 1}^N \Lambda^{N(n+1)-\alpha}_{-\alpha \pi} t_{\alpha,n} \right),
\] (5.15)
\[
W_R(x,t,\Lambda) = P_R(x,t,\Lambda) \circ \exp \left( -\sum_{n \geq 0} \sum_{\beta = -M+1}^0 \Lambda^{-M(n+1)+\beta}_{\pi} t_{\beta,n} \right).
\] (5.16)

By Sato equations (3.7), we have the following identities [2]:
\[
\partial_{n_1} W_L := \begin{cases} (B_{\alpha,n})_+ W_L, & \alpha = N, \ldots, 1, \\ -(B_{\alpha,n})_- W_L, & \alpha = 0, \ldots, -M + 1, \end{cases}
\]
\[
\partial_{n_2} W_R := \begin{cases} (B_{\alpha,n})_+ W_R, & \alpha = N, \ldots, 1, \\ -(B_{\alpha,n})_- W_R, & \alpha = 0, \ldots, -M + 1. \end{cases}
\]

One can then prove that the product \( W_L^{-1} W_R \) is invariant under all the flows, i.e.
\[
\partial_{y,n}(W_L^{-1} W_R) = 0.
\]
Therefore
\[
W_L^{-1} W_R(x,t,\Lambda) = W_L^{-1} W_R(x,0,\Lambda) = P_L^{-1} P_R(x,0,\Lambda).
\]
This implies that
\[
(\mathcal{P}_L^{-1}\mathcal{P}_R)(t) = \exp \left( \sum_{n \geq 0} \sum_{\alpha = 1}^N \Lambda^{N(n+1-\frac{\alpha}{N})} t_{\alpha,n} \right) \circ (\mathcal{P}_L^{-1}\mathcal{P}_R)(0) \circ \exp \left( \sum_{n \geq 0} \sum_{\beta = -M+1}^0 \Lambda^{-M(n+1+\frac{\beta}{M})} t_{\beta,n} \right). \tag{5.17}
\]

If we let \(\tau_0 = 1\) and \(\tau_i = 0\) \((i \in \mathbb{Z}_-\)) , then the bi-infinite matrix will become the semi-infinite matrix everywhere. Representing the product \(\mathcal{P}_L^{-1}\mathcal{P}_R\) in the semi-infinite matrix, i.e. \(\tilde{\mathcal{P}}_L^{-1}\tilde{\mathcal{P}}_R\), and considering identities (3.18) and (3.19), equation (5.17) can be written as follows:

\[
(\tilde{\mathcal{P}}_L^{-1}\tilde{\mathcal{P}}_R)(t) = \left( \begin{array}{cccc}
1 & P_1(t) & P_2(t) & \cdots \\
0 & 1 & P_1(t) & \cdots \\
0 & 0 & 1 & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{array} \right) (\tilde{\mathcal{P}}_L^{-1}\tilde{\mathcal{P}}_R(0))
\times \left( \begin{array}{cccc}
1 & 0 & 0 & \cdots \\
P_1(t) & 1 & 0 & \cdots \\
P_2(t) & P_1(t) & 1 & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{array} \right), \tag{5.18}
\]

where \(P_k\) are the Schur polynomials as defined in (2.9). To construct tau functions, we define the moment matrix \(M_\infty(t)\) as

\[
M_\infty(t) := \left( \begin{array}{cccc}
\tilde{\mathcal{P}}_L^{-1}\tilde{\mathcal{P}}_R(t)
\end{array} \right).
\]

**Proposition 5.2.** The constrained condition of the BTH is equivalent to the matrix \(M_\infty\) which satisfies the identity

\[
\Lambda^{Nn} M_\infty = M_\infty \Lambda^{-Mn}. \tag{5.19}
\]

**Proof.** Using the constraint on \(\mathcal{L}\), i.e. \(\partial_{t_{\alpha,n}} \mathcal{L} = \partial_{t_{\alpha,n}} \mathcal{L}\), we obtain \(\partial_{t_{\alpha,n}} M_\infty = \partial_{t_{\alpha,n}} M_\infty\) which leads to equation (5.19). \(\square\)

Direct calculation can lead to the following proposition.

**Proposition 5.3.** The matrix \(M_\infty\) satisfies the identity

\[
\partial_{t_{\alpha,n}} M_\infty = \Lambda^{N(n+1-\frac{\alpha}{N})} M_\infty \partial_{t_{\alpha,n}} M_\infty = M_\infty \Lambda^{-M(n+1+\frac{\beta}{M})}.
\]

Let \(M_i\) be the \(i \times i\) submatrix of \(M_\infty\) of the top-left corner. By equation (5.18), each \(\tau\)-function can be obtained by the determinant [12]

\[
\tau_i(t) = \det (M_i(t)).
\]

Now, we will consider the detailed structure of the tau functions of the BTH from the point of reduction of the two-dimensional Toda hierarchy.

As we all know, the tau functions of the two-dimensional Toda lattice hierarchy are given by

\[
\tau = \begin{vmatrix}
\bar{C}_{0,0} & \bar{C}_{0,1} & \cdots & \bar{C}_{0,j-1} \\
\bar{C}_{1,0} & \bar{C}_{1,1} & \cdots & \bar{C}_{1,j-1} \\
\vdots & \vdots & \ddots & \ddots \\
\bar{C}_{i-1,0} & \bar{C}_{i-1,1} & \cdots & \bar{C}_{i-1,j-1} \\
\end{vmatrix}, \tag{5.20}
\]

where \(\bar{C}_{i,j}\) are the Schur polynomials as defined in (2.9).
where

\[
\bar{C}_{i,j} = \int \int \rho(\lambda, \mu) \lambda^i \mu^j e^{\sum_{n=0}^{\infty} \lambda^{n+1} e^{x_n} + \sum_{n=0}^{\infty} \mu^{n+1} e^{y_n}} \ d\lambda \ d\mu
\]

\[
= \sum_{k,l=0}^{\infty} \bar{c}_{i,j,k,l} P_k(x) P_l(y).
\]

We should note here that the coefficients \(\bar{c}_{i,j,k,l}\) are totally independent.

As the original tridiagonal Toda lattice is \((1, 1)\) reduction of the two-dimensional Toda lattice hierarchy, therefore to obtain the solution of the tridiagonal Toda lattice we need to add a factor \(\delta(\lambda - \mu)\) under the integral in the definition of \(\bar{C}_{i,j}\), i.e.

\[
\int \int \rho(\lambda, \mu) \delta(\lambda - \mu) \lambda^i \mu^j e^{\sum_{n=0}^{\infty} \lambda^{n+1} e^{x_n} + \sum_{n=0}^{\infty} \mu^{n+1} e^{y_n}} \ d\lambda \ d\mu,
\]

which can further lead to

\[
\int \rho(\lambda, \lambda) \lambda^i \mu^j e^{\lambda (L(t_{\alpha}) + R(t_{\beta}))} \ d\lambda.
\]

After changing the time variables \(x, y\) to \(t_{\alpha}, t_{\beta}\), equation (5.22) becomes a new function

\[
\int \rho(\lambda, \lambda) \lambda^i \mu^j e^{\xi_L(\lambda, t_{\alpha}) + \xi_R(\lambda, t_{\beta})} \ d\lambda,
\]

which corresponds to the \((1, 1)\)-BTH.

Denote \(\omega_N\) and \(\omega_M\) as the \(N\)th root and \(M\)th root of the unit, respectively. For the \((N, M)\)-BTH, a new function \(\bar{C}_{i,j}\) (new form of \(\bar{C}_{i,j}\)) has the following form:

\[
\bar{C}_{i,j} = \int \int \rho(\omega_N^p \lambda, \omega_M^q \mu) \omega_N^p \lambda^i \omega_M^q \mu^j e^{\xi_L(\omega_N^p \lambda, t_{\alpha}) + \xi_R(\omega_M^q \mu, t_{\beta})} \ d\lambda \ d\mu
\]

\[
= \sum_{p=0}^{N-1} \sum_{q=0}^{M-1} \int \rho(\omega_N^p \lambda, \omega_M^q \mu) (\omega_N^p \lambda^i \omega_M^q \mu^j) e^{\xi_L(\omega_N^p \lambda, t_{\alpha}) + \xi_R(\omega_M^q \mu, t_{\beta})} \ d\lambda,
\]

where

\[
t_{\alpha,n}^p = (\omega_N^p)^{N(n+1)} t_{\alpha,n}
\]

\[
t_{\beta,n}^q = (\omega_M^q)^{M(n+1)} t_{\beta,n}.
\]

Transforms (5.23) and (5.24) can be seen as the twist of exponential solutions because of the bigraded structure.

Because in the next section we will consider the rational solutions of the BTH, we write \(\bar{C}_{i,j}\) further as

\[
\bar{C}_{i,j} = \sum_{p=0}^{N-1} \sum_{q=0}^{M-1} \sum_{k,l=0}^{\infty} c_{i,j,p,q} \ P_k(t_{\alpha}) P_l(t_{\beta}),
\]

where

\[
c_{i,j,p,q} = \int \rho(\omega_N^p \lambda, \omega_M^q \mu) (\omega_N^p \lambda^i \omega_M^q \mu^j) e^{\xi_L(\omega_N^p \lambda, t_{\alpha}) + \xi_R(\omega_M^q \mu, t_{\beta})} \ d\lambda.
\]

We find that the coefficients \(\{c_{i,j,p,q} ; i, j, k, l \geq 0 ; 0 \leq p \leq N - 1, 0 \leq q \leq M - 1\}\) satisfy

\[
c_{i,j,p,q}^{k+l} = c_{i,j,p,q}^{k+l+N}.
\]
which tells us that $P_m(t_a)P_n(t_b)$ and $P_m(t_a)P_n(t_b)$ always appear at the same time. So the element at a position of the $k$th row and $l$th column in the initial moment matrix can have the following form:

$$
\left(\tilde{P}_L^{-1}P_R(0)\right)_{k,l} = \sum_{p=0}^{N-1} \sum_{q=0}^{M-1} \sum_{k,l=0}^{\infty} c_{0,0,p,q}^{k-1,l-1}.
$$

Therefore the tau functions of the BTH can explicitly be written in the form

$$
\tau_i = \left| C_{0,0} \ C_{0,1} \ \ldots \ C_{0,i-1} \\
C_{1,0} \ C_{1,1} \ \ldots \ C_{1,i-1} \\
\ldots \ \ldots \ \ldots \ \ldots \\
C_{i-1,0} \ C_{i-1,1} \ \ldots \ C_{i-1,i-1} \right|
$$

(5.27)

With the definition of $C_{i,j}$, in the next section we will construct the Lax matrix solution using orthogonal polynomials which are nothing but the wavefunction in a semi-infinite vector form. Before we do that, we need some formula about the property of the Schur–Hirota derivatives described in the following lemma [18].

**Lemma 5.4.** Schur derivatives have the following formula:

$$
P_k(\hat{D}_L)\tau_i \cdot \tau_j = \sum_{k+l=n} P_k(\hat{D}_L)\tau_i \times P_l(\hat{D}_L)\tau_j,
$$

(5.28)

$$
P_l(\hat{D}_R)[0, 1, 2, \ldots, i - 1]_L = [0, 1, 2, \ldots, i - 2, i + l - 1],
$$

(5.29)

$$
P_l(-\hat{D}_L)[0, 1, 2, \ldots, i - 1]_L = (-1)^i[0, 1, \ldots, i - l - 1, i - l + 1, \ldots, i - 1, i],
$$

(5.30)

$$
P_k(\hat{D}_R)\tau_i \cdot \tau_j = \sum_{k+l=n} P_k(\hat{D}_R)\tau_i \times P_l(\hat{D}_R)\tau_j,
$$

(5.31)

$$
P_l(\hat{D}_R)[0, 1, 2, \ldots, i - 1]_R = [0, 1, 2, \ldots, i - 2, i + l - 1],
$$

(5.32)

$$
P_l(-\hat{D}_R)[0, 1, 2, \ldots, i - 1]_R = (-1)^i[0, 1, \ldots, i - l - 1, i - l + 1, \ldots, i - 1, i],
$$

(5.33)

where

$$
[k_0, k_1, k_2, \ldots, k_{i-1}]_L = \left| C_{k_0,0} \ C_{k_0,1} \ \ldots \ C_{k_0,i-1} \\
C_{k_1,0} \ C_{k_1,1} \ \ldots \ C_{k_1,i-1} \\
\ldots \ \ldots \ \ldots \ \ldots \\
C_{k_{i-1},0} \ C_{k_{i-1},1} \ \ldots \ C_{k_{i-1},i-1} \right|
$$

(5.34)

$$
[k_0, k_1, k_2, \ldots, k_{i-1}]_R = \left| C_{0,k_0} \ C_{0,k_1} \ \ldots \ C_{0,k_{i-1}} \\
C_{1,k_0} \ C_{1,k_1} \ \ldots \ C_{1,k_{i-1}} \\
\ldots \ \ldots \ \ldots \ \ldots \\
C_{i-1,k_0} \ C_{i-1,k_1} \ \ldots \ C_{i-1,k_{i-1}} \right|
$$

(5.35)

where $P_k$ are the Schur polynomials as defined in (2.9).

In fact, lemma 5.4 is a special case of the abstract general formula of the Schur function [18]:

$$
S_Y(\hat{D})\tau_\phi = \tau_Y(t).
$$

(5.36)
Here $\tau_\phi := [0, 1, 2, \ldots, i - 1]$ is the standard Wronskian determinant, $Y := (Y_0, Y_1, Y_2, \ldots, Y_{i-1})$, and

$$
S_Y = \begin{vmatrix}
P_{Y_{i+1}} & P_{Y_{i+2}} & \cdots & P_{Y_{2}} & P_{Y_{1}} & P_{Y_{0}} \\
P_{Y_{i+2}} & P_{Y_{i+3}} & \cdots & P_{Y_{2}} & P_{Y_{1}} & P_{Y_{0}} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
P_{Y_{i+1}} & P_{Y_{i+2}} & \cdots & P_{Y_{i+3}} & P_{Y_{i+4}} & \cdots \\
P_{Y_{i+1}} & P_{Y_{i+2}} & \cdots & P_{Y_{i+3}} & P_{Y_{i+4}} & \cdots \\
\end{vmatrix}_{1 \times j},
$$

$Y = (Y_0, Y_1, Y_2, \ldots, Y_{i-1})$ ($Y_0 > Y_1 > Y_2 > \ldots > Y_{i-1}$) corresponds to Young’s diagram with $Y_0$ boxes at the first row, $Y_1$ boxes at the second row and so on. Note the formula [18]

$$
S_Y(-t) = (-1)^{|Y|} S_Y'(t),
$$

where $Y'$ is the conjugate Young diagram of $Y$. Equation (5.36) together with (5.37) further leads to the lemma above easily.

Then we define the wavefunctions

$$W_L = (W_{L1}, W_{L2}, \ldots), \quad \hat{W}_R = (\hat{W}_{R1}, \hat{W}_{R2}, \ldots)^T$$

with

$$W_{Li}(\lambda^\frac{i}{2}, t) = \frac{e^{\xi_L(\lambda^\frac{i}{2}, t)}}{\tau_{i-1}} \begin{vmatrix}
C_{0,0} & C_{0,1} & \cdots & C_{0,i-2} & 1 \\
C_{1,0} & C_{1,1} & \cdots & C_{1,i-2} & \lambda^\frac{i}{2} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
C_{i-1,0} & C_{i-1,1} & \cdots & C_{i-1,i-2} & \lambda^\frac{i-2}{2} \\
\end{vmatrix},
$$

(5.38)

$$\hat{W}_{Rj}(\lambda^\frac{j}{2}, t) = \frac{e^{\xi_R(\lambda^\frac{j}{2}, t)}}{\tau_j} \begin{vmatrix}
C_{0,0} & C_{0,1} & \cdots & C_{0,j-1} \\
C_{1,0} & C_{1,1} & \cdots & C_{1,j-1} \\
\vdots & \vdots & \ddots & \vdots \\
C_{j-2,0} & C_{j-2,1} & \cdots & C_{j-2,j-1} \\
1 & \lambda^\frac{j}{2} & \cdots & \lambda^\frac{j-1}{2} \\
\end{vmatrix},
$$

(5.39)

and

$$\hat{W}_{Rj}(\lambda^\frac{j}{2}, t) = \frac{e^{\xi_R(\lambda^\frac{j}{2})}}{\tau_{j-1}} \begin{vmatrix}
C_{0,0} & C_{0,1} & \cdots & C_{0,j-1} \\
C_{1,0} & C_{1,1} & \cdots & C_{1,j-1} \\
\vdots & \vdots & \ddots & \vdots \\
C_{j-2,0} & C_{j-2,1} & \cdots & C_{j-2,j-1} \\
1 & \lambda^\frac{j}{2} & \cdots & \lambda^\frac{j-1}{2} \\
\end{vmatrix},
$$

(5.40)

which satisfy the following orthogonality relation:

$$\langle W_{Li}, \hat{W}_{Rj} \rangle = \delta_{i,j}, \quad \langle W_{Li}, \hat{W}_{Rj} \rangle = \delta_{i,j} h_j,
$$

(5.41)

where

$$h_j := \frac{\tau_j}{\tau_{j-1}}.
$$

and the inner product $\langle , \rangle$ of the functions $A$ and $B$ is defined as

$$\langle A, B \rangle := \int\int \rho(\lambda, \mu) \delta(\lambda^N - \mu^M) A(\lambda, t) B(\mu, t) \, d\lambda \, d\mu.
$$

Therefore tau functions have another form as

$$\tau_m := \text{det}(W_{Li}, \hat{W}_{Rj})_{1 \leq i, j \leq m}. $$

21
The entries of the matrix representation of the Lax operator $\mathcal{L}$ can then be calculated as

$$a_{ij} = \frac{P_{j-i+N} (\hat{D}_L) \tau_j \tau_{i-1}}{\tau_{i-1} \tau_j}$$

$$= \frac{1}{\tau_{j-1} \tau_j} \sum_{m=i=N-j}^{i-N+j} P_m (\hat{\delta}_L) \tau_j P_n (\hat{-\delta}_L) \tau_{i-1}$$

$$= \frac{1}{\tau_{j-1} \tau_j} \sum_{m=0}^{i-j+N} [0, 1, \ldots, j - 2, j + m - 1]_L$$

$$(-1)^{m+j-i-N} [0, 1, \ldots, j - N + m - 2, j - N + m, \ldots, i - 1]_L$$

$$= \begin{pmatrix}
\lambda e^{\delta_t (\lambda \frac{1}{\Lambda_1}, \tau)} \\
\tau_{j-1}
\end{pmatrix}
\begin{pmatrix}
C_{0,0} & C_{0,1} & \cdots & C_{0,i-2} & 1 \\
C_{1,0} & C_{1,1} & \cdots & C_{1,i-2} & \lambda^{-\frac{1}{2}} \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
C_{i-1,0} & C_{i-1,1} & \cdots & C_{i-1,i-2} & \lambda^{-\frac{i-1}{2}}
\end{pmatrix}
\begin{pmatrix}
\lambda e^{\delta_t (\lambda \frac{1}{\Lambda_1}, \tau)} \\
\tau_j
\end{pmatrix}
\begin{pmatrix}
C_{0,0} & C_{0,1} & \cdots & C_{0,j-1} \\
C_{1,0} & C_{1,1} & \cdots & C_{1,j-1} \\
\cdots & \cdots & \cdots & \cdots \\
C_{j-2,0} & C_{j-2,1} & \cdots & C_{j-2,j-2} & \lambda^{\frac{j-1}{2}} \\
1 & \lambda^{\frac{j-2}{2}} & \cdots & \lambda^{\frac{1}{2}}
\end{pmatrix}.$$ 

So

$$a_{i,j} = \langle \lambda \mathcal{W}_L \mathcal{W}_R \rangle_i, \quad (5.42)$$

which are given by the matrix representations of the eigenvalue problems $\mathcal{L} \mathcal{W}_L = \lambda \mathcal{W}_L$ and $\mathcal{W}_R \mathcal{L} = \lambda \mathcal{W}_R$ (see [12] for the details). Until now, we have solved the BTH using orthogonal polynomials.

If we denote $\tilde{\mathcal{W}}_L$ and $\tilde{\mathcal{W}}^{-1}_R$ as the corresponding matrix forms of $\mathcal{W}_L$ (5.15) and $\mathcal{W}_R^{-1}$ (5.16), respectively, then $\mathcal{W}_L$ and $\mathcal{W}_R$ can be represented by matrices $\tilde{\mathcal{W}}_L$ and $\tilde{\mathcal{W}}^{-1}_R$ respectively as follows.

Because

$$\Lambda \begin{pmatrix}
1 \\
\lambda^{\frac{1}{2}} \\
\lambda^{-\frac{1}{2}} \\
\lambda^{\frac{1}{2}} \\
\cdots
\end{pmatrix} = \lambda \begin{pmatrix}
1 \\
1 \\
1 \\
1 \\
\cdots
\end{pmatrix},$$

we can obtain

$$\mathcal{W}_L = \tilde{\mathcal{W}}_L \begin{pmatrix}
1 \\
\lambda^{\frac{1}{2}} \\
\lambda^{-\frac{1}{2}} \\
\lambda^{\frac{1}{2}} \\
\cdots
\end{pmatrix} = \tilde{\mathcal{P}}_L (x, t, \Lambda) \begin{pmatrix}
1 \\
\lambda^{\frac{1}{2}} \\
\lambda^{-\frac{1}{2}} \\
\lambda^{\frac{1}{2}} \\
\cdots
\end{pmatrix} e^{\delta_t (\lambda \frac{1}{\Lambda_1}, t)}$$

$$\mathcal{W}_R = \tilde{\mathcal{W}}^{-1}_R \begin{pmatrix}
1 \\
\lambda^{\frac{1}{2}} \\
\lambda^{-\frac{1}{2}} \\
\lambda^{\frac{1}{2}} \\
\cdots
\end{pmatrix} = \tilde{\mathcal{P}}^{-1}_L (x, t, \Lambda) \begin{pmatrix}
1 \\
\lambda^{\frac{1}{2}} \\
\lambda^{-\frac{1}{2}} \\
\lambda^{\frac{1}{2}} \\
\cdots
\end{pmatrix} e^{\delta_t (\lambda \frac{1}{\Lambda_1}, t)}.$$
\[ \begin{pmatrix} \frac{1}{\tau_1} + \lambda \frac{\partial L}{\partial \tau_1} & \frac{P_1(-\hat{\partial}_L)\tau_1}{\tau_1} + P_2(-\hat{\partial}_L)\tau_2 \lambda \frac{\partial L}{\partial \tau_2} + \lambda \frac{\partial L}{\partial \tau_2} \\ \frac{P_2(-\hat{\partial}_L)\tau_2}{\tau_2} & \frac{1}{\tau_2} + \frac{P_1(-\hat{\partial}_L)\tau_2}{\tau_2} + P_2(-\hat{\partial}_L)\tau_3 \lambda \frac{\partial L}{\partial \tau_3} + \lambda \frac{\partial L}{\partial \tau_3} \\ \vdots \end{pmatrix} e^{\xi L(\lambda, \tau, \Lambda)} \]

where \( \hat{P}_L(x, t, \Lambda) \) is as matrix (3.17). This agrees with the definition of \( W_L \) in equation (5.38). Also, similarly we can obtain

\[ \tilde{W}_R = \begin{pmatrix} 1 & \lambda^{\frac{1}{M}} & \lambda^{\frac{2}{M}} & \cdots \\ \lambda^{\frac{1}{M}} & 1 & \lambda^{\frac{1}{M}} & \cdots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix} \tilde{W}_R^{-1} = e^{\xi R(\lambda, \tau, \Lambda)} \begin{pmatrix} 1 & \lambda^{\frac{1}{M}} & \lambda^{\frac{2}{M}} & \cdots \\ \lambda^{\frac{1}{M}} & 1 & \lambda^{\frac{1}{M}} & \cdots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix} \]

where \( \tilde{P}_R(x, t, \Lambda) \) is as matrix (3.20). This also agrees with the definition of \( \bar{W}_R \) in equation (5.39).

The formal factorization mentioned above is about the infinite-sized Lax matrix. In the next section, we consider its finite-sized truncation. Then we find that the finite-sized Lax matrix is in fact nilpotent because of dressing structure (3.2). This finite-sized case corresponds to the rational solutions of the BTH which will be considered in the next section.

6. Rational solutions of the \((N, M)\)-BTH

It is well known that the \( \tau \)-function of the original tridiagonal Toda lattice, i.e. the \((1, 1)\)-BTH, has the Schur polynomial solutions associated with rectangular Young’s diagrams. So what kind of Young’s diagrams correspond to the BTH becomes an interesting question. In this section, we consider only the homogeneous rational solution of the \((N, M)\)-BTH which is one kind of most interesting solutions in nonlinear integrable systems.

In order to describe the homogeneous rational solution of the \((N, M)\)-BTH \((N \leq M)\), we firstly set \( \deg(t_{\alpha,0}) = \deg(t_{0,0}) = MN \). Then we obtain \( \deg(P_m(t_{\alpha})) = mM \) and \( \deg(P_n(t_{\beta})) = nN \).

Define

\[ c_{p,q}^{m,n} := c_{0,0,p,q}^{m,n}, \]

and choose the following special homogeneous polynomials \( \tilde{P}_k(t_{\alpha}, t_{\beta}) \) with degree \( k \) as \( \tau_1 \):

\[ \tilde{P}_k(t_{\alpha}, t_{\beta}) := \sum_{p=0}^{N-1} \sum_{q=0}^{M-1} \sum_{mM+nN=k} c_{p,q}^{m,n} P_m(t_{\alpha}) P_n(t_{\beta}), \]

where \( t_{\alpha} \) and \( t_{\beta} \) denote

\[ t_{\alpha} = \{ t_{\alpha,n} : 0 \leq \alpha \leq N, n = 0, 1, 2, \ldots \} \]

\[ t_{\beta} = \{ t_{\beta,n} : -M + 1 \leq \beta \leq 0, n = 0, 1, 2, \ldots \}. \]

and \( k \) can be any number in the set \( \{ mM+nN ; m, n \in \mathbb{Z}_+ \} \). The polynomials \( P_m(t_{\alpha}) \) and \( P_n(t_{\beta}) \) are the elementary Schur polynomials, and they satisfy the following relations:

\[ \frac{\partial P_m(t_{\alpha})}{\partial t_{\alpha}, \alpha'} = P_{m-N(\alpha'+1)+\alpha'}(t_{\alpha}) \]

\[ \frac{\partial P_n(t_{\beta})}{\partial t_{\beta}, \beta'} = P_{n-M(\beta'+1)+\beta'}(t_{\beta}). \]
Note here that \( \{c^m_{p,q}\}_{0 \leq m, p \leq N-1, 0 \leq n, q \leq M-1} \) can be arbitrary constants.

Define

\[
p^{l,f}_k = \sum_{p=0}^{N-1} \sum_{q=0}^{M-1} \sum_{mM+nN=k} c^{m+l,n+f}_{p,q} P_m(t_a) P_n(t_b),
\]

and the rational solutions for the \((N,M)\)-BTH have the following diagram representation:

\[
D_j = \{k - (j - 1)N, k - (j - 2)N - M, \ldots, k - (j - 1)M\},
\]

\[
k = 0, N, M, 2N, N + M, 2M, \ldots.
\]

The difference between two adjacent numbers in \(D_j\) is \(M-N\).

We denote the tau function corresponding to \(D_j(k)\) as \(\tau_{j,D_j(k)}\), which has the following form:

\[
\tau_{j,D_j(k)} = S_{D_j(k)} = \begin{vmatrix}
\bar{P}_0^{(0,0)} & \bar{P}_1^{(1,0)} & \cdots & \bar{P}_{k-M}^{(j-1,0)} \\
\bar{P}_0^{(0,1)} & \bar{P}_1^{(1,1)} & \cdots & \bar{P}_{k-M}^{(j-1,1)} \\
\vdots & \vdots & \ddots & \vdots \\
\bar{P}_0^{(0,j-1)} & \bar{P}_1^{(1,j-1)} & \cdots & \bar{P}_{k-M}^{(j-1,j-1)} \\
\end{vmatrix}
\]

where \(\tau_{j,D_j(k)}\) denotes the \(j\)th tau function \(j \times j\) determinant generated by \(\bar{P}_k^{(0,0)}\). The range of rank \(j\) depends on the choice of \(k\). This kind of diagram like \(D_j\) is not the classical Young’s diagram. It is a kind of generalized diagram which counts the homogeneous degree that comes from the multiplication of two classical Schur functions. We can call this kind of generalized diagram degree diagram. Similar to Young’s diagram, the tau functions represented by the degree diagram are also of Wronskian form, and the derivative is about \(\partial t^{N,M}\) or \(\partial t^{0,0}\). Because the scale of degree in the definition before (e.g. \(\deg(P_m(t_a)) = mM\)) is bigger than the common degree of a Schur polynomial \(\deg(P_m(t)) = m\), the difference of subscript between two adjacent rows (columns) is \(N(M)\) not \(1(1)\). From this point, it is also different from the Hankel determinant.

In the following, first we consider only the case when \(N\) and \(M\) are the co-prime integers. When they are not co-prime, we divide them by their GCD and use the theory of co-prime case in the following. In fact, we find that for a fixed value of \(p, q\), all the coefficients of a homogeneous polynomial will be the same because they all satisfy relation (5.25). Then we obtain

\[
p^{l,f}_k = \sum_{p=0}^{N-1} \sum_{q=0}^{M-1} \sum_{mM+nN=k} c^{m+l,n+f}_{p,q} P_m(t_a) P_n(t_b)
\]

\[
= \sum_{p=0}^{N-1} \sum_{q=0}^{M-1} c^{m+l,n+f}_{p,q} P_{m1}(t_a) P_{n1}(t_b) + c^{m+l,n+z+f}_{p,q} P_{m2}(t_a) P_{n2}(t_b) + \cdots
\]

\[
+ c^{m+i,n+i+f}_{p,q} P_{m(i)}(t_a) P_{n(i)}(t_b)
\]

\[
= \sum_{p=0}^{N-1} \sum_{q=0}^{M-1} c^{m+l,n+n+f}_{p,q} P_m(t_a) P_n(t_b)
\]

\[
= \sum_{mM+nN=k} P_m(t_a) P_n(t_b),
\]
where
\[
\sum_{p=0}^{N-1} \sum_{q=0}^{M-1} c_{p,q}^{m_1+1,l,n_1+1,l'} = \cdots = c_{p,q}^{m_0+1,l,n_0+1,l'},
\]
and the number \( l(k) \) depends on \( k \). For simplicity, here we consider only the case where all coefficients \( c_{l,l'}^k \) equal to 1 which is our central consideration in this section and define
\[
\bar{P}_k = \sum_{mM+nN=k} P_m(t_a)P_n(t_b) = P_{m_1}(t_a)P_{m_2}(t_a)P_{n_1}(t_b) + P_{m_2}(t_a)P_{n_2}(t_b) + \cdots + P_{m_{l(k)}}(t_a)P_{n_{l(k)}}(t_b).
\]

Then the \( \tau \)-function \( \tau_s \) generated by \( \tau_1 = \bar{P}_k \) can be expressed by a double-Wronskian determinant,
\[
\tau_s(k, t_a, t_b) = \begin{vmatrix}
\bar{P}_k & \bar{P}_k-M & \cdots & \bar{P}_k-(s-1)M \\
\bar{P}_{k-N} & \bar{P}_{k-N-M} & \cdots & \bar{P}_{k-N-(s-1)M} \\
\vdots & \vdots & \ddots & \vdots \\
\bar{P}_{k-(s-1)M} & \bar{P}_{k-(s-1)M-M} & \cdots & \bar{P}_{k-(s-1)M-(s-1)M} \\
\end{vmatrix}_{s \times s}
\]

\[
= \begin{vmatrix}
P_{m_1}(t_a) & P_{m_2}(t_a) & \cdots & P_{m_1}(t_a) \\
P_{m_1-1}(t_a) & P_{m_2-1}(t_a) & \cdots & P_{m_1-1}(t_a) \\
P_{m_1-2}(t_a) & P_{m_2-2}(t_a) & \cdots & P_{m_2-2}(t_a) \\
\vdots & \vdots & \ddots & \vdots \\
P_{m_{l-1}+1}(t_a) & P_{m_{l-1}+1}(t_a) & \cdots & P_{m_{l-1}+1}(t_a) \\
\end{vmatrix}_{s \times s}
\]

\[
\times \begin{vmatrix}
P_{n_1}(t_b) & P_{n_1-1}(t_b) & P_{n_1-2}(t_b) & \cdots & P_{n_1-1}(t_b) \\
P_{n_2}(t_b) & P_{n_2-2}(t_b) & \cdots & P_{n_2}(t_b) \\
\vdots & \vdots & \ddots & \vdots \\
P_{n_{l-1}+1}(t_b) & P_{n_{l-1}+1}(t_b) & \cdots & P_{n_{l-1}+1}(t_b) \\
\end{vmatrix}_{s \times s}.
\]

Conversely, for a fixed size \( j \) of the Lax matrix for the \((N, M)\)-BTH, the choices of \( k \) for \( \tau_1 = \bar{P}_k \) are in a set \( K_j \):
\[
K_j := \{ k | k = (j-1)NM + mM + nN, m, n \in \mathbb{Z}_+, 0 \leq m < N, 0 \leq n < M \}. \tag{6.1}
\]

We see that the number of elements in the set \( K_j \) is \( NM \) (when they are not co-prime, this number will be \( \frac{NM}{(N, M)} \), where \((N, M)\) is the GCD of \( N \) and \( M \)). When the values of \( N, M, j, m, n \) are chosen, a series of non-vanishing \( \tau \)-functions corresponding to them will be fixed.

In the following, we consider the rational solutions of the \((N, M)\)-BTH with a finite-sized Lax matrix. For the \((N, M)\)-BTH, the size of the minimal Lax matrix is \((M+1) \times (M+1)\). This minimal Lax matrix has \( M \) non-vanishing \( \tau \)-functions and the degree of any one has \( M-N \) jumps between adjacent rows. For the special case of \( N = M \), degree diagrams are always rectangular which can also be seen from the definition of \( D_j \).

Besides considering the degree diagrams, it is more interesting to consider the decomposition of degree diagrams into the representation of Young’s diagrams. In fact, the
Young diagram representation of the general BTH has a form of multiplication of two different groups of Young’s diagrams which will be shown in the following theorem.

**Theorem 6.1.** For \( j \times j (j \geq M + 1) \)-sized Lax matrix of the \((N, M)\)-BTH (denoted as \((N, M)_{j \times j}\)), after choosing the value of \( k = (j - 1)MN + mM + nN \), Young’s diagram representation of a series of corresponding tau functions is as follows:

\[
\begin{align*}
\tau_1 &= \sum_{0 \leq a \leq j - 1} S_{(j-1-a)N+m}(t_\alpha) S_{(a+aM)}(t_\beta), \\
\tau_2 &= \sum_{0 \leq a < b \leq j - 1} S_{((j-1-a)+(j-1-b)N+m)}(t_\alpha) S_{(a+bM-aM)}(t_\beta), \\
\tau_3 &= \sum_{0 \leq a < b < c \leq j - 1} S_{((j-1-a)N+m-2, (j-1-b)N+m-1, (j-1-c)N+m)}(t_\alpha) S_{(a+cM-2,a+bM-1,a+aM)}(t_\beta), \\
& \vdots \\
\tau_r &= \sum_{0 \leq a_1 < a_2 < \cdots < a_r \leq j - 1} S_{((j-1-a_1)N+m-r+1, (j-1-a_2)N+m-r+2, \ldots, (j-1-a_r)N+m)}(t_\alpha) S_{(a+rM-aM, a+rM-aM, \ldots, a+rM-aM)}(t_\beta), \\
& \vdots \\
\tau_j &= S_{((j-1)(N-1)+m, (j-2)(N-1)+m-1, \ldots, m)}(t_\alpha) S_{((j-1)(M-1), n+(j-2)(M-1), \ldots, n)}(t_\beta).
\end{align*}
\]

**Proof.** To prove this theorem, one needs to use the Cauchy–Binet formula. The process is quite complicated because of huge sizes of matrices. So we will omit the proof. One can understand the pattern by the following example, i.e. \((2, 3)\)-BTH. \(\square\)

To see it clearly, we give some specific examples as follows.

**Example 6.2.** The \((2, 3)\)-BTH has six sets of \(\tau\)-functions for each size of the Lax matrix and the degree diagram for every tau function has one jump between adjacent rows. See \((2, 3)_{4 \times 4}\) in detail as the following degree diagram:

\[
(2, 3)_{4 \times 4} = \begin{cases}
(0, 0) : & \tau_{1,18} \rightarrow \tau_{2,16,15} \rightarrow \tau_{3,14,13,12} \rightarrow \tau_{4,11,10,9}, \\
(0, 1) : & \tau_{1,19} \rightarrow \tau_{2,18,17} \rightarrow \tau_{3,16,15,14} \rightarrow \tau_{4,14,13,12,11}, \\
(1, 0) : & \tau_{1,21} \rightarrow \tau_{2,19,18} \rightarrow \tau_{3,17,16,15} \rightarrow \tau_{4,15,14,13,12}, \\
(0, 2) : & \tau_{1,22} \rightarrow \tau_{2,20,19} \rightarrow \tau_{3,18,17,16} \rightarrow \tau_{4,16,15,14,13}, \\
(1, 1) : & \tau_{1,23} \rightarrow \tau_{2,21,20} \rightarrow \tau_{3,19,18,17} \rightarrow \tau_{4,17,16,15,14}, \\
(1, 2) : & \tau_{1,25} \rightarrow \tau_{2,23,22} \rightarrow \tau_{3,21,20,19} \rightarrow \tau_{4,19,18,17,16}.
\end{cases}
\]  

(6.2)

where \([p, q], 0 \leq p < 2, 0 \leq q < 3\) denotes the value of \((m, n)\) in the value of \(k\), i.e. (6.1). Here \(k\) takes values in \(\{18, 20, 21, 22, 23, 25\}\), i.e. the values in the brackets of \(\tau_{i,\cdot}\). Every tau function \(\tau_{i,\cdot}\) generates a series of tau functions which are connected by a right arrow in (6.2). \(\tau_{i,\ldots}\) represents the \(l\)th tau function whose degree diagram is in the brackets \(\{\ldots\}\). In (6.2), we use the product of Young’s diagrams to represent the four tau functions of the first line, i.e. the \((0, 0)\) case, as follows by the Cauchy–Binet formula:
\[
\begin{align*}
\tau_1 &= \tau_1(18) = \left| \begin{array}{ccc}
P_0(t_a) & P_1(t_a) & P_2(t_a) \\
1 & 0 & 0 \\
S_{\alpha}(t_a)S_{\alpha}(t_\beta) + S_{\alpha}(t_a)S_{\alpha}(t_\beta) + S_{\alpha}(t_a)S_{\alpha}(t_\beta) \\
\end{array} \right|, \\
\tau_2 &= \tau_2(16,15) = \left| \begin{array}{ccc}
P_0(t_a) & P_1(t_a) & P_2(t_a) \\
1 & 0 & 0 \\
S_{\alpha}(t_a)S_{\alpha}(t_\beta) + S_{\alpha}(t_a)S_{\alpha}(t_\beta) + S_{\alpha}(t_a)S_{\alpha}(t_\beta) \\
\end{array} \right|, \\
\tau_3 &= \tau_3(14,13,12) = \left| \begin{array}{ccc}
P_0(t_a) & P_1(t_a) & P_2(t_a) \\
1 & 0 & 0 \\
S_{\alpha}(t_a)S_{\alpha}(t_\beta) + S_{\alpha}(t_a)S_{\alpha}(t_\beta) + S_{\alpha}(t_a)S_{\alpha}(t_\beta) \\
\end{array} \right|, \\
\tau_4 &= \tau_4(12,11,10,9) = \left| \begin{array}{ccc}
P_0(t_a) & P_1(t_a) & P_2(t_a) \\
1 & 0 & 0 \\
S_{\alpha}(t_a)S_{\alpha}(t_\beta) + S_{\alpha}(t_a)S_{\alpha}(t_\beta) + S_{\alpha}(t_a)S_{\alpha}(t_\beta) \\
\end{array} \right|.
\end{align*}
\]

After general theory on the homogeneous rational solutions of the \((N, M)\)-BTH, as a special but important case, the rational solutions of the \((1, M)\)-BTH will be considered in the next subsection.

**6.1. Rational solutions of the \((1, M)\)-BTH**

Since the \(t_{1,a}\) flows are same as the \(t_{0,a}\) flows, we use \(t_{1,a} + t_{0,a}\) as a new variable and identify \(t_{1,a}\) as \(t_{0,a}\). Then the rational solutions for the \((1, M)\)-BTH are obtained from the \(\tau\)-function:

\[
\tau_j(k, t) = \begin{vmatrix}
P_k & P_{k-M} & \cdots & P_{k-(j-1)M} \\
P_{k-1} & P_{k-M-1} & \cdots & P_{k-(j-1)M-1} \\
\vdots & \vdots & \ddots & \vdots \\
P_{k-(j-1)} & P_{k-M-(j-1)} & \cdots & P_{k-(j-1)(M+1)}
\end{vmatrix}_{j \times j},
\]

where \(P_n = P_n(t_\beta)\) with \(t_{0,a} = t_{0,a} + t_{1,a}\). This \(\tau\)-function can be given by the Schur polynomial associated with the Young diagram which is the same as the degree diagram mentioned above for the \((1, M)\)-BTH, i.e., \(\tau_j(k) = S_{\nu_j}(k)\) with

\[
Y_j(k) = (k - j + 1, k - j + 1 - (M - 1), \ldots, k - (j - 1)M)
\]

for \(j = 1, 2, \ldots, 1 + \left\lfloor \frac{k}{M} \right\rfloor\),

where \(\left\lfloor \frac{k}{M} \right\rfloor\) denotes the biggest integer which is less than or equal to \(\frac{k}{M}\). Note here that the number of boxes in the diagram increases by \(M - 1\) between adjacent rows. Let us now give some examples of the \(\tau\)-functions for the \((1, M)\)-BTH with specific size \(r\) of the Lax matrix, denoted by \((1, M)_{r \times r}\). For a given size \(r > M\), the choices of \(k\) are in the set...
\[(r - 1)M, \ldots, rM - 1\], that is, there are \(M\) choices of the first member of the \(\tau\)-functions, \(\tau_1 = P_k\). Then each value of \(k\) generates \(r\) tau functions ordered from \(\tau_1\) to \(\tau_r\).

**Example 6.3.** The \((1, 1)\)-BTH, i.e. the original Toda lattice, has only one set of \(\tau\)-functions for each size of Lax matrix:

\[
\begin{align*}
(1, 1)_{2 \times 2} & : \tau \rightarrow \tau_{\phi}, \\
(1, 1)_{3 \times 3} & : \tau \rightarrow \tau \rightarrow \tau_{\phi}, \\
(1, 1)_{4 \times 4} & : \tau \rightarrow \tau \rightarrow \tau \rightarrow \tau_{\phi}, \\
\cdots & : \cdots,
\end{align*}
\]

where \(\tau_{\phi} = 1\).

The \((1, 2)\)-BTH has two sets of \(\tau\)-functions, and the Young diagram for each \(\tau\)-function has one jump between adjacent rows:

\[
\begin{align*}
(1, 2)_{3 \times 3} & : \tau \rightarrow \tau \rightarrow \tau, \\
& \tau \rightarrow \tau \rightarrow \tau, \\
& \tau \rightarrow \tau \rightarrow \tau,
\end{align*}
\]

\[
\begin{align*}
(1, 2)_{4 \times 4} & : \tau \rightarrow \tau \rightarrow \tau \rightarrow \tau, \\
& \tau \rightarrow \tau \rightarrow \tau \rightarrow \tau, \\
& \tau \rightarrow \tau \rightarrow \tau \rightarrow \tau, \\
\cdots & : \cdots
\end{align*}
\]

Similarly the \((1, 3)\)-BTH has three sets of \(\tau\)-functions, and the Young diagram has two jumps between adjacent rows:

\[
\begin{align*}
(1, 3)_{4 \times 4} & : \tau \rightarrow \tau \rightarrow \tau \rightarrow \tau, \\
& \tau \rightarrow \tau \rightarrow \tau \rightarrow \tau, \\
\cdots & : \cdots
\end{align*}
\]

7. Conclusions and discussions

We proved the equivalence between the \((N, M)\)-BTH and the \((M, N)\)-BTH, derived the primary Hirota equations of the \((N, M)\)-BTH, and found several explicit formulas about the solutions for the BTH using orthogonal polynomials in the matrix form. We also constructed some rational solutions of the BTH which are parameterized by the products of Schur polynomials corresponding to the non-rectangular Young diagrams. It may be interesting to find their significance in terms of the representation theory, as in the case of the original Toda lattice where the rational solutions are given by the Schur polynomials of the rectangular Young diagrams, and they are the Virasoro singular vectors.
Acknowledgments

This work was carried out under the guidance of Professor Yuji Kodama during my visit to Ohio State University. I would like to thank Professor Kodama for his guidance and many useful discussions. I would also like to thank the Department of Mathematics at Ohio State University for providing me with generous support and making my visit so pleasant. I am grateful to Professor Jingsong He (NBU, China) for useful discussions and his general support.

References

[1] Carlet G 2006 The extended bigraded Toda hierarchy J. Phys. A: Math. Gen. 39 9411–35 (arXiv:math-ph/0604024)
[2] Li C Z, He J S, Wu K and Cheng Y 2010 Tau function and Hirota bilinear equations for the extended bigraded Toda hierarchy J. Math. Phys. 51 043514 (arXiv:0906.0624)
[3] Toda M 1967 Vibration of a chain with nonlinear interaction J. Phys. Soc. Japan 22 431–6
[4] Toda M 1989 Nonlinear Waves and Solitons (Dordrecht: Kluwer)
[5] Takasaki K 2010 Two extensions of 1D Toda hierarchy J. Phys. A: Math. Theor. 43 434032 (arXiv:1002.4688)
[6] Ueno K and Takasaki K 1984 Toda lattice hierarchy Group Representations and Systems of Differential Equations (Tokyo, 1982) (Advanced Studies in Pure Mathematics vol 4) (Amsterdam: North-Holland) pp 1–95
[7] Li C Z, He J S and Su Y C 2011 Block type symmetry of bigraded Toda hierarchy (in preparation)
[8] Milanov T and Tseng H H 2008 The spaces of Laurent polynomials, P1-orbiﬁolds, and integrable hierarchies J. Reine Angew. Math. 622 189–235 (arXiv:math.AG/0607012)
[9] Blaszak M and Szum A 2001 Lie algebraic approach to the construction of (2 + 1)-dimensional lattice-ﬁeld and ﬁeld integrable Hamiltonian equations J. Math. Phys. 42 225–59
[10] Adler M and Moerbeke P V 2001 Darboux transforms on band matrices, weights, and associated polynomials Int. Math. Res. Not. 18 935–84
[11] Kodama Y and Pierce V U 2009 Combinatorics of dispersionless integrable systems and universality in random matrix theory Commun. Math. Phys. 292 529–68 (arXiv:0811.0351)
[12] Kodama Y and Ye J 1996 Iso-spectral deformations of general matrix and their reductions on Lie algebras Commun. Math. Phys. 178 765–88
[13] Adler M and Moerbeke P van 1999 Generalized orthogonal polynomials, discrete KP and Riemann–Hilbert problems Commun. Math. Phys. 207 589–620
[14] Blaszak M and Marciniak K 1994 R-matrix approach to lattice integrable systems J. Math. Phys. 35 4661–82
[15] Svinin A K 2001 A class of integrable lattices and KP hierarchy J. Phys. A: Math. Gen. 34 10559–68
[16] Wu Y T and Hu X B 1999 A new integrable differential-difference system and its explicit solutions J. Phys. A: Math. Gen. 32 1515–21
[17] Yu G F, Li C X, Zhao J X and Hu X B 2005 On a special two-dimensional lattice by Blaszak and Szum: pfaffianization and molecule solutions J. Nonlinear Math. Phys. 12 316–32
[18] Ohta Y, Satsuma J, Takahashi D and Tokihiro T 1988 An elementary introduction to Sato theory Prog. Theor. Phys. Suppl. 94 210–41

29