Infinite dimensional continued fractions.

Leonel Robert*
Department of Mathematics
University of Toronto
Toronto, ON M5S 3G3

Luis Santiago†
Department of Mathematics
University of Toronto
Toronto, ON M5S 3G3

Abstract

In this paper we define a continued fraction for infinite dimensional vectors. This continued fraction is analogous to the Jacobi-Perron continued fraction of finite vectors. We develop the basic properties of the continued fraction when the entries of the coefficients are noncommutative formal variables and when they belong to a noncommutative ring. In this context we establish a connection with the quasideterminants of Hessenberg matrices. We also study the continued fraction applied to Laurent series and link the cotinued fraction to problems in rational approximation. Finally, we apply the continued fraction to the solution of certain integrable systems that generalize the nonabelian Toda lattice.

Keywords: Jacobi-Perron continued fraction, quasideterminants, Hessenberg matrices, Padé approximants, nonabelian Toda lattice.
MSC (2000): 30B70, 41A28, 35Q58.

0 Introduction.

This paper is devoted to the study of a notion of continued fraction for infinite vectors. In fact, we define a continued fraction for pairs of infinite vectors with entries in a unital ring R (we do not assume that R is commutative).

In order to be able to write a finite continued fraction of pairs of vectors we need to be able to add them, which we do componentwise, and to invert them. The inverse of a pair of vectors is defined as follows. Let $f = \binom{f_1}{f_2}$ be an element of $R^{\mathbb{N}} \times R^{\mathbb{N}}$. Let us write $f_1 = (f_{1,0}, f_{1,1}, \ldots)$ and $f_2 = (f_{2,0}, f_{2,1}, \ldots)$ and suppose that $f_{2,0}$ is invertible in R. We define the inverse of f in $R^{\mathbb{N}} \times R^{\mathbb{N}}$ by

$$\frac{1}{f} = \begin{pmatrix} f_{2,0}^{-1} & f_{1,0}f_{2,0}^{-1} & \cdots \\ f_{2,1}f_{2,0}^{-1} & f_{2,2}f_{2,0}^{-1} & \cdots \end{pmatrix}. \quad (1)$$

*lrobert@math.utoronto.ca, Telephone: (416) 978-3484, Fax: (416) 978-4107
†santiago@math.utoronto.ca, Telephone: (416) 978-3484, Fax: (416) 978-4107

Notice that, unlike the inversion of real numbers, this inverse operation is not an involution; that is, the inverse of the inverse of a vector is not the vector itself. This continued fraction can be seen as an infinite dimensional version of the Jacobi-Perron continued fraction.

We now review the contents of the paper. Throughout its sections, the dominant theme of the paper is to see to what extent the properties of the standard continued fraction have an analog in the infinite dimensional continued fraction.

Section 1: The continued fraction. Section 1 develops the basic properties of the continued fraction of pairs of infinite vectors with entries in a ring R. As in the case of the standard continued fraction, we obtain recurrence relations for the numerators and denominators of the convergents of the continued fraction. These recurrence formulas allow us to establish a relation between the continued fraction and the quasideterminants of a matrix defined by Gelfand and Retakh (see [13]). We express the entries of the convergents of the continued fraction in terms of certain quasideterminants of a Hessenberg matrix. We also prove the convergence of the infinite continued fraction in the case when the entries of the continued fraction are formal noncommutative variables. Subsection 1.4 discusses the continued fraction associated to the first column of D^{-1}, where D is a Hessenberg matrix. This continued fraction is analogous to the standard continued fraction associated to a Jacobi matrix (see [22]).

There are several continued fractions for finite vectors. Probably the best known of these is the Jacobi-Perron continued fraction. Subsection 1.5 discusses the relation of the infinite continued fraction to the Jacobi-Perron continued fraction.

Section 2: The ring of Laurent series. In the rings $\mathbb{R}$ and $\mathbb{C}((z^{-1}))$ (Laurent series in z^{-1}), not only one can write finite and infinite continued fractions, one also has an algorithm to expand elements of the ring in a continued fraction. Using the definition of inverse (1), a similar algorithm can be applied to pairs of infinite vectors with entries in $\mathbb{R}$ or $\mathbb{C}((z^{-1}))$. In this section we discuss several properties of the continued fraction algorithm for pairs of infinite vectors of Laurent series with coefficients in a unital ring. (The important case when the ring R is taken to be $\mathbb{R}$ will not be treated in this paper.)

The classical theory of continued fractions of a Laurent series and its relation to orthogonal polynomials and rational approximation has been generalized to the noncommutative setting by I. Gelfand, et. al. in [12]. We will see how to a great extent this noncommutative theory of the standard continued fraction generalizes to our new setting. Thus, we prove that if the continued fraction of an element of $R((z^{-1}))^{\mathbb{N}} \times R((z^{-1}))^{\mathbb{N}}$ never stops then it converges to this same element. In particular this happens for the continued fraction associated to the first column of $(zI - D)^{-1}$, where D is a Hessenberg matrix. This continued fraction is analogous to the standard continued fraction of the Weyl function of a Jacobi matrix.

Given a Laurent series $h \in \mathbb{C}((z^{-1}))$, the convergents of its continued fraction solve a problem of best approximation to h by rational functions. They are the diagonal Padé approximants corresponding to the normal indices of h (see [17]).

If a finite vector of Laurent series is given, the relation of the convergents of its Jacobi-Perron continued fraction to its simultaneous Padé approximants is less straightforward. Still, assuming normality of all indices the convergents of the Jacobi-Perron continued fraction are simultaneous Padé approximants (see [17]).

We discuss in Subsection 2.2 the relation between the continued fraction of infinite vectors and the problem of infinite simultaneous Padé approximation. The formal series $h(z)$ is replaced by an infinite vector of formal series. The Padé approximants are replaced by the infinite simultaneous Padé approximants defined by A. Bultheel and M. van Barel in Section 2.7 of [9]. In the case of normality of all indices we show that the convergents of the continued fraction are infinite simultaneous Padé approximants.

In Subsection 2.3 the continued fraction algorithm is used to find the LU decomposition of a matrix Υ with noncommutative entries. More explicitly, we decompose Υ in the product L_1BL_2 where L_1 is lower triangular with ones in the main diagonal, B is invertible and diagonal, and L_2 is upper triangular with ones in the main diagonal. Along with the entries of L_1, B and L_2, the continued fraction algorithm gives us extra information that will be used in Section 3 in the solution of certain nonlinear differential equations.

The analytic counterpart of an element of $\mathbb{C}((z^{-1}))$ is a meromorphic function on a neighborhood of infinity. Thus we can apply the continued fraction algorithm to a vector with entries in the ring of meromorphic functions on a neighborhood of infinity and ask for its point-wise or uniform convergence. In Subsection 2.4 we let the ring R be the complex numbers and rely on the results of [19] to prove the analytic convergence of the continued fraction of the first column of $LR(z, D)$. Here L is lower triangular and $R(z, D)$ is the resolvent function of the Hessenberg matrix D as an operator acting on l_2. We allow D to be an unbounded operator. This includes cases where the functions of the continued fraction are not necessarily analytic on a neighborhood of infinity (instead, they have determinate moments). Corollary 2 generalizes Stieljes's Theorem on the convergence of the continued fraction in the determinate case.

Section 3: Solving integrable systems. The relation of the classical continued fraction of Laurent series in $\mathbb{C}((z^{-1}))$ to the Toda lattice was first established by J. Moser (see [17, 16]). The continued fraction is used to solve the inverse problem of recovering the solutions of the Toda lattice from the spectral data of the initial condition. In Section 3 we apply the infinite dimensional continued fraction algorithm to the solution of a class of nonlinear systems of differential equations that includes the nonabelian Toda lattice. Following [1], these systems can be described as a generalization of the finite and semi-infinite nonabelian 2-Toda lattices. As in the case of the 2-Toda lattice, the solution of these systems can be reduced by a suitable change of variable to a linear system of differential equations with constant coefficients for the matrix of moments Υ. The continued fraction is used to go from the matrix Υ to the solution of the original system. Examples of systems to which these approach can be applied are: the relativistic Toda lattice, the full Kostant-Toda lattice, the Bogoyavlensky lattice, the KP discrete hierarchy and the 2-Toda lattice. These examples are examined in Subsection 3.4.

1 Continued fractions of infinite vectors.

Throughout this section R will denote a ring with a unit. Unless it is otherwise specified, we do not assume that R is commutative. Given $f = \binom{f_1}{f_2}$, an element of $R^{\mathbb{N}} \times R^{\mathbb{N}}$ such that $f_{2,0}$ is invertible in R, we define $\frac{1}{f}$ by the equation (1) of the introduction.

Notation convention 1. In many relevant cases the vectors that we will write in a continued fraction appear naturally as column vectors of matrices. Thus, we will always understand that the vectors f_1 and f_2 in $\binom{f_1}{f_2}$ are column vectors. Even if, for reasons of space, we will sometimes write them in row form.

1.1 The continued fraction with formal entries.

In this subsection we will let R be the free division ring generated by a set X. We refer to [10] and [13] for the explicit construction of this ring.

Let $a_i = \binom{b_i}{c_i}$, $i = 0, 1, \ldots$ be a sequence of pairs of infinite column vectors b_i and c_i, with $b_i = (b_{i,0}, b_{i,1}, \ldots)$ and $c_i = (c_{i,0}, c_{i,1}, \ldots)$. Let us define the set $X = \{b_{i,j}, c_{i,j} : i, j = 0, 1, \ldots\}$ formed by the entries of a_i. We consider the entries of the a_i as free variables and we denote by $F(X)$ the free division ring generated by X.

Definition 1. *The expression*

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_n}}}}, \tag{2}$$

is well defined in $F(X)^{\mathbb{N}} \times F(X)^{\mathbb{N}}$ (i.e, all the inversions are possible). We call this expression the continued fraction of the sequence $\{a_i\}_{i=0}^n$ and we denote it briefly by $\langle a_i \rangle_{i=0}^n$.

It is well known that the numerators and denominators of the standard continued fraction satisfy a three-terms recurrence relation (see [17]).

Theorem 1. *Let $\phi_n = \langle a_i \rangle_{i=0}^n$, $n \geq 0$. Then*

$$\phi_n = \begin{pmatrix} B_{n,0}P_n^{-1} & B_{n,1}P_n^{-1} & \cdots \\ C_{n,0}P_n^{-1} & C_{n,1}P_n^{-1} & \cdots \end{pmatrix}, \tag{3}$$

where $B_{n,i}$, $C_{n,i}$ and P_n satisfy $P_0 = 1$, $B_{0,k} = b_{0,k}$, $C_{0,k} = c_{0,k}$ and the *recurrence relations*

$$B_{n,k} = \sum_{i=0}^{n-1} B_{n-i-1,k} c_{n,i} + b_{n,k-n}, \tag{4}$$

$$C_{n,k} = \sum_{i=0}^{n-1} C_{n-i-1,k} c_{n,i} + c_{n,k+n}, \tag{5}$$

$$P_n = \sum_{i=0}^{n-1} P_{n-i-1} c_{n,i}, \tag{6}$$

for $n \geq 1$. In these formulas we make the convention that $b_{n,i} = 0$ for $i < -1$ and $b_{n,-1} = 1$.

Proof. The proof follows by induction similarly to the classical case. □

It is clear from the recurrence formulas for P_n, $B_{n,i}$ and $C_{n,i}$ that they are polynomials in the variables $b_{i,j}$ and $c_{i,j}$. For example we have

$$P_n = \sum_{1 \leq j_1 < j_2 < \ldots < j_k < n} c_{j_1,j_1-1} c_{j_2,j_2-j_1-1} c_{j_3,j_3-j_2-1} \cdots c_{n,n-j_k-1}. \tag{7}$$

This formula can be deduced from (9.1.1) and (9.1.2) of [13] and the relation between the continued fraction and quasideterminants that will be discussed in subsection 1.2. Similar formulas hold for $B_{n,i}$ and $C_{n,i}$.

1.1.1 Continued fraction in an arbitrary unital ring.

Let R be a unital ring and $a'_i = \binom{b'_i}{c'_i} \in R^{\mathbb{N}} \times R^{\mathbb{N}}$, $i = 0, 1, \ldots, n$ a sequence of pairs of infinite vectors with entries in R. Let us define $\alpha : X \to R$ such that $\alpha(b_{i,j}) = b'_{i,j}$ and $\alpha(c_{i,j}) = c'_{i,j}$. If α can be evaluated in the element $f = \langle a_i \rangle_{i=0}^n$, we say that the continued fraction $\langle a'_i \rangle_{i=0}^n$ *exists* and we put $\langle a'_i \rangle_{i=0}^n = \alpha(f)$. We say that $\langle a'_i \rangle_{i=0}^n$ *can be evaluated* if all the continued fractions $\langle a'_i \rangle_{i=k}^n$, $k = 0, 1 \ldots, n$ exist. It follows from (3) that $\langle a'_i \rangle_{i=0}^n$ exists if and only if $\alpha(P_n)$ is invertible in R.

1.2 Relation of the continued fraction (2) to quasideterminants.

An upper Hessenberg matrix $D = (d_{i,j})$ is a matrix such that $d_{i,j} = 0$ for all $i > j+1$. Unless it is otherwise specified, in this paper Hessenberg always means upper Hessenberg such that $d_{i,i-1}$ is invertible for all i.

Let D_n be the Hessenberg matrix

$$D_n = \begin{pmatrix} d_{0,0} & d_{0,1} & \ldots & d_{0,n-2} & d_{0,n-1} \\ -1 & d_{1,1} & \ldots & d_{1,n-2} & d_{1,n-1} \\ 0 & -1 & \ldots & d_{2,n-2} & d_{2,n-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & -1 & d_{n-1,n-1} \end{pmatrix}, \tag{8}$$

with entries in a unital ring R. For $k = 0, \ldots, n$ let us define the sequence $\{P^k_{i-k}\}_{i=0}^n$ by the recurrence relation (6) taking the coefficient $c_{i,j}$ of the recurrence formula equal to the entry $d_{i-j-1,i-1}$ of D_n. The initial values are set to $P^k_i = 0$ if $i < 0$ and $P^k_0 = -1$. It is known that the matrix D_n is invertible if and only if $P_n = P^0_n$ is invertible in R (see [13]). More explicitly, the matrix

$$(P^{i+1}_{j-i-1} - P^{i+1}_{n-i-1} P_n^{-1} P_j)_{i,j=0}^{n-1}, \tag{9}$$

is the inverse of D_n. (One can verify that (9) is a left inverse of D by a direct computation making use of the recurrence formulas that define the P^j_is. Since D is invertible, (9) must be the inverse of D. See [19] for a proof in the commutative case.)

In [13] the quasideterminant $|A|_{i,j}$ of a matrix A with entries in a unital ring is defined. If A is invertible and the entry (j,i) of A^{-1} is invertible, then the quasideterminant $|A|_{i,j}$ is the inverse of the entry (j,i) of A^{-1}. Thus by (9) we have that $|D_n|_{0,n} = -P_n$ and more generally $|D_n|_{0,k} = P_n(P^k_{n-k})^{-1}$. $(-1)^{n-1}|D|_{0,n}$ is a polynomial expression in terms of the entries of D_n which generalizes the determinant of a matrix with commutative entries (see [13]). We write $\det(D_n) = (-1)^{n-1}|D|_{0,n}$. By the recurrence relations defining P^k_{n-k} one can deduce that $P^k_{n-k} = (-1)^{n-k}\det(D_n^{(k)})$ for $0 \le k \le n$, where $D_n^{(k)}$ is the matrix obtained from D_n deleting the first k rows and columns (the determinant of the 0×0 matrix is taken equal to -1).

The entries of the vector $\phi_n = \langle \binom{b_i}{c_i} \rangle_{i=0}^n$ can be expressed as quasideterminants of suitably defined Hessenberg matrices. Define the Hessenberg matrices $H^B_{n,k}$ and $H^C_{n,k}$ by

$$H^B_{n,k} = \begin{pmatrix} b_{0,k} & b_{1,k-1} & \cdots & b_{n-1,k-n-1} & b_{n,k-n} \\ -1 & c_{1,0} & \cdots & c_{n-1,n-2} & c_{n,n-1} \\ 0 & -1 & \cdots & c_{n-1,n-3} & c_{n,n-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & -1 & c_{n,0} \end{pmatrix},$$

$$H^C_{n,k} = \begin{pmatrix} c_{0,k} & c_{1,k+1} & \cdots & c_{n-1,k+n-1} & c_{n,k+n} \\ -1 & c_{1,0} & \cdots & c_{n-1,n-2} & c_{n,n-1} \\ 0 & -1 & \cdots & c_{n-1,n-3} & c_{n,n-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & -1 & c_{n,0} \end{pmatrix},$$

and $H_n = (H^C_{n,k})^{(1)} = (H^B_{n,k})^{(1)}$.

Proposition 1. *Recall the notation of Theorem 1. We have*

$$P_n = -\det(H_n), \quad B_{n,k} = -\det(H^B_{n,k}), \quad C_{n,k} = -\det(H^C_{n,k}).$$

By our earlier discussion we know that for a Hessenberg matrix D_n we have $|D_n|_{0,0} = \det(D_n)\det(D_n^{(1)})^{-1}$. Thus, by the preceding proposition we have

$$\phi_n = \begin{pmatrix} |H^B_{n,0}|_{0,0} & |H^B_{n,1}|_{0,0} & \cdots \\ |H^C_{n,0}|_{0,0} & |H^C_{n,1}|_{0,0} & \cdots \end{pmatrix}.$$

Proof. By the previous discussion on determinants of Hessenberg matrices, we know that the sequence $-\det(H_n)$, $n = 0, 1, \ldots$ satisfies a recurrence relation in terms of the entries of H_n (the recurrence of the P_ns). This formula coincides with the recurrence (6) of the denominators of the continued fraction. The initial values of these two sequences are also equal. Thus we get $P_n = -\det(H_n)$. The same reasoning can be used to prove the other two formulas. □

1.3 Convergence of the formal continued fraction.

Let $D = (d_{i,j})_{i,j=0}^{\infty}$ be an infinite Hessenberg matrix such that $d_{i+1,i} = -1$. In [13] the quasideterminants $|D_n|_{0,0}$ of the finite sections of D were shown to converge to an expression of the form $P_\infty Q_\infty^{-1}$. Here P_∞ and Q_∞ are formal series ce (R, ies ce ([illegible] 0 [illegible] 3 - [illegible] ∘ m 6 m [illegible] 3 m

1.3.1 The continued fraction in a topological ring.

Suppose that R is a topological ring with a unit and $a_i = \binom{b_i}{c_i} \in R^{\mathbb{N}} \times R^{\mathbb{N}}$, $i = 0, 1, \ldots$ is an infinite sequence of pairs of vectors with coefficients in the ring. Suppose that all continued fractions $\langle a_i \rangle_{i=0}^{n}$, $n = 0, 1, \ldots$, exist and they converge entry-wise to some $f \in R^{\mathbb{N}} \times R^{\mathbb{N}}$. Then we say that the infinite continued fraction of the sequence $\{a_i\}_{i=0}^{\infty}$ converges and we write $f = \langle a_i \rangle_{i=0}^{\infty}$.

1.4 Continued fractions associated to a Hessenberg matrix.

Notation convention 2. The following convention will be used throughout the paper. Let us denote by Π_n the [redacted] $\times n$ matrix $\Pi_n = (e_0, e_1, \ldots, e_{n-1})$. To every $n \times n$ matrix M we can associate the infinite matrix $\Pi_n M \Pi_n^t$ that coincides with M in the upper left corner and has zeroes everywhere else. In the sequel, when talking about a finite matrix as infinite, for example D_n^{-1}, we always mean the infinite matrix obtained through Π_n. Also for an infinite matrix A we write $A_n = \Pi_n^t A \Pi_n$ for its finite sections. The same convention applies for finite and infinite vectors.

1.4.1 Finite Hessenberg matrices.

Let D_n be as in (8) with entries in a ring R. In [13] the following generalized continued fraction expansion of the quasideterminant $|D_n|_{0,0}$ is given (see also [4]):

$$|D_n|_{0,0} = d_{0,0} + \sum_{j_0 \neq 0} d_{0,j_0} \cfrac{1}{d_{1,j_0} + \sum\limits_{j_1 \neq 0, j_0} d_{1,j_1} \cfrac{1}{d_{2,j_1} + \cdots}}. \tag{11}$$

When D_n is a Jacobi matrix, the right side of this equality takes the form of a standard continued fraction.

In this subsection we prove a continued fraction expansion of the first column of D_n^{-1}. More specifically, let v_n be the infinite column vector which has in its first entries the entries of the first column of D_n^{-1} and zeroes everywhere else. Thus v_n has the form

$$\begin{aligned} v_n =& (|D|_{0,0}^{-1}, |D|_{0,1}^{-1}, \ldots, |D|_{0,n}^{-1}, 0, \ldots) \\ =& (P_{n-1}^1 P_n^{-1}, P_{n-2}^2 P_n^{-1}, \ldots, P_0^n P_n^{-1}, 0, \ldots). \end{aligned}$$

In Theorem 2 below we will prove that the vector $\binom{v_n}{0}$ can be written in a continued fraction in the following way:

$$\begin{pmatrix} v_n \\ 0 \end{pmatrix} = \cfrac{1}{\binom{0}{c_1} + \cfrac{1}{\binom{0}{c_2} + \cfrac{1}{\cdots + \cfrac{1}{\binom{0}{c_n}}}}}, \tag{12}$$

where

$$c_k = (d_{k-1,k-1}, d_{k-2,k-1}, \dots, d_{1,k-1}, d_{0,k-1}, 0, \dots).$$

Notice that the k-th coefficient of this continued fraction only depends on the k-th column of D_k. For $m \le n$ the continued fractions of $\binom{v_m}{0}$ given by these formulas are equal to the continued fractions of $\binom{v_n}{0}$ truncated in the m-th term.

The continued fraction (12) exists if and only if $\det(D_n)$ is invertible. For every k the continued fraction $\langle \binom{0}{c_i} \rangle_{i=k}^n$ is the continued fraction expansion of the first column of $(D_n^{(k)})^{-1}$. Thus, the continued fraction (12) can be evaluated if and only if $\det(D_n^{(k)})$ is invertible for all $k = 0, \dots, n$.

We now prove (12). More generally, we consider $L = (l_{i,j})_{i,j=0}^{\infty}$ an infinite lower triangular matrix (i.e., $l_{i,j} = 0$ for $i < j$) with invertible elements in the main diagonal and $D_n = (d_{i,j})_{i,j=0}^{n-1}$ an $n \times n$ invertible Hessenberg matrix. We do not assume that the entries $d_{i+1,i}$ of D_n are equal to -1. Instead we only assume that they are invertible elements of the ring R.

Let us denote by w_n the first column of $L\Pi_n D_n^{-1}$.

Theorem 2. *The vector* $\binom{w_n}{0}$ *can be written in a continued fraction in the following way:*

$$\begin{pmatrix} w_n \\ 0 \end{pmatrix} = \cfrac{1}{\binom{b_1}{c_1} + \cfrac{1}{\binom{b_2}{c_2} + \cfrac{1}{\dots + \cfrac{1}{\binom{b_n}{c_n}}}}}, \tag{13}$$

where

$$\begin{aligned} b_{k+1} =& (l_{k+1,k}, l_{k+2,k}, \dots, l_{n,k}, \dots) l_{k,k}^{-1}, \\ c_{k+1} =& - (l_{k-1,k-1} d_{k,k-1}^{-1} d_{k,k}, l_{k-2,k-2} d_{k-1,k-2}^{-1} d_{k-1,k}, \dots, l_{0,0} d_{1,0}^{-1} d_{1,k}, -d_{0,k}, 0, \dots) l_{k,k}^{-1}. \end{aligned}$$

Remark 1. Taking L equal to the identity matrix and $d_{k,k-1} = -1$, $k = 0, \dots, n$ we obtain the continued fraction (12).

Remark 2. Since

$$\frac{1}{\begin{pmatrix} f_{1,0} & f_{1,1} & \cdots \\ f_{2,0} & f_{2,1} & \cdots \end{pmatrix}} g = \frac{1}{\begin{pmatrix} f_{1,0} & f_{1,1} & \cdots \\ f_{2,0} g^{-1} & f_{2,1} & \cdots \end{pmatrix}},$$

we have that

$$\cfrac{1}{\binom{x_1}{y_1} + \cfrac{1}{\binom{x_2}{y_2} + \cfrac{1}{\dots + \cfrac{1}{\binom{x_n}{y_n}} \alpha_{n-1}} \alpha_2} \alpha_1} \alpha_0 = \cfrac{1}{\binom{x_1}{y_1'} + \cfrac{1}{\binom{x_2}{y_2'} + \cfrac{1}{\dots + \cfrac{1}{\binom{x_n}{y_n'}}}}},$$

with $y_i' = (y_{i,0}\alpha_{i-1}^{-1}, y_{i,1}\alpha_{i-2}^{-1}, \ldots, y_{i,i-1}\alpha_0^{-1}, y_{i,i}, y_{i,i+1}, \ldots)$. Using this we obtain an equivalent representation of (13) in the following way

$$\binom{w_n}{0} = \cfrac{1}{\binom{b_1'}{c_1'} l_{0,0}^{-1} - \cfrac{1}{\binom{b_2'}{c_2'} l_{1,1}^{-1} - \cfrac{1}{\cdots - \cfrac{1}{\binom{b_n'}{c_n'} l_{n-1,n-1}^{-1}} d_{n-1,n-2} l_{n-2,n-2}^{-1}} d_{2,1} l_{1,1}^{-1}} d_{1,0} l_{0,0}^{-1}}, \tag{14}$$

$$\begin{aligned} b_{k+1}' &= (l_{k+1,k}, l_{k+2,k}, \ldots, l_{n,k}, \ldots), \\ c_{k+1}' &= (d_{k,k}, d_{k-1,k}, \ldots, d_{1,k}, d_{0,k}, 0, \ldots), \end{aligned}$$

with $k = 0, 1, \ldots, n-1$.

Proof of Theorem 2. Let us write $q_i = (D_n^{-1})_{i,0}$, $i = 0, 1, \ldots, n-1$. For $k = 1, \ldots, n$ we let $r_k = \binom{r_{k,1}}{r_{k,2}}$ and $a_k = \binom{b_k}{c_k}$, where

$$\begin{aligned} b_k &= (l_{k,k-1}, l_{k+1,k-1}, \ldots, l_{n,k}, \ldots) l_{k-1,k-1}^{-1}, \\ c_k &= (-l_{k-2,k-2} d_{k-1,k-2}^{-1} d_{k-1,k-1}, -l_{k-3,k-3} d_{k-2,k-3}^{-1} d_{k-2,k-1}, \ldots \\ &\qquad \ldots, l_{0,0} d_{1,0}^{-1} d_{1,k-1}, d_{0,k-1}, 0, \ldots) l_{k-1,k-1}^{-1}, \\ r_{k,1} &= \left(\sum_{i=k}^{k} l_{k,i} q_i, \sum_{i=k}^{k+1} l_{k+1,i} q_i, \ldots, \sum_{i=k}^{n-1} l_{n-1,i} q_i, \sum_{i=k}^{n-1} l_{n,i} q_i, \ldots \right) q_{k-1}^{-1} l_{k-1,k-1}^{-1}, \\ r_{k,2} &= \left(-l_{k-2,k-2} d_{k-1,k-2}^{-1} \sum_{i=k}^{n-1} d_{k-1,i} q_i, -l_{k-3,k-3} d_{k-2,k-3}^{-1} \sum_{i=k}^{n-1} d_{k-2,i} q_i, \ldots \right. \\ &\qquad \left. \ldots, -l_{0,0} d_{1,0}^{-1} \sum_{i=k}^{n-1} d_{1,i} q_i, \sum_{k}^{n-1} d_{0,i} q_i, 0, \ldots \right) q_{k-1}^{-1} l_{k-1,k-1}^{-1}. \end{aligned}$$

Using that $d_{k-1,k-2} q_{k-1} + \sum_{i=k-1}^{n-1} d_{k-1,i} q_i = 0$ we can check that $r_{k-1} = \frac{1}{a_k + r_k}$ for $1 < k \leq n$. Also we can deduce from $d_{0,0} q_0 + \sum_{i=1}^{n-1} d_{0,i} q_i = 1$ that $\binom{w_n}{0} = \frac{1}{a_1 + r_1}$. Now the result follows from these two relations. □

1.4.2 Infinite Hessenberg matrices.

Let $D = (d_{i,j})_{i,j=0}^{\infty}$ be an infinite Hessenberg matrix. For simplicity we assume that $d_{i+1,i} = -1$. We also restrict our analysis to a Hessenberg matrix with formal entries above the diagonal $(i+1, i)$, since it is in this setting that we have proven the convergence of the continued fraction. The algebra $F[[X]]$ of formal series in $X = \{d_{i,j} : i, j = 0, 1, \ldots, i \geq j\}$ has a canonical embedding in a division algebra $F((X))$ such that the image of $F[[X]]$ generates $F((X))$ (see [10], Section 4).

Recall that $P^k_{i-k} = (-1)^{i-k}\det(D^{(k)}_i)$. Thus $\{P^k_{i-k}\}_{i=0}^{\infty}$ is the sequence of determinants of the finite sections of the infinite Hessenberg matrix $D^{(k)}$. By the results of [13], $P^k_{i-k}\gamma_i^{-1}$ converges to some $P^k_\infty \in F((X))$, where $\gamma_i = \Pi_{j=0}^{i-1} d_{j,j}$ (alternatively we can proceed as in the proof of Proposition 2). Thus the left side of (12) converges componentwise to the infinite column vector

$$v_\infty = (P^1_\infty P^{-1}_\infty, P^2_\infty P^{-1}_\infty, \ldots).$$

Hence we can write $\binom{v_\infty}{0} = \langle \binom{0}{c_i} \rangle_{i=1}^{\infty}$ with

$$c_k = (d_{k,k}, d_{k-1,k}, \ldots, d_{1,k}, d_{0,k}, 0, \ldots).$$

More generally, we let L be an infinite lower triangular matrix with formal entries. Then the continued fraction in Theorem 2 converges componentwise to $\binom{Lv_\infty}{0}$ when $n \to$ [redacted].

1.5 Relation of the continued fraction (2) to the Jacobi-Perron continued fraction.

Let us define the inverse of a finite vector by

$$\frac{1}{f} = (f_m^{-1}, f_0 f_m^{-1}, \ldots, f_{m-1} f_m^{-1}).$$

With this definition of inverse we can write a continued fraction of finite vectors. We shall call a continued fraction of this type a Jacobi-Perron continued fraction (see [17]).

Let R be a ring and $f \in R^{m+1}$ a vector with entries in R. We denote by $\widetilde{f}$ the infinite vector $\widetilde{f} = (f_m, f_{m-1}, \ldots, f_0, 1, 0, \ldots)$. Suppose that f can written as the Jacobi-Perron continued fraction $f = \langle a_i \rangle_{i=0}^{n}$. Then we have

$$\begin{pmatrix} g \\ \widetilde{f} \end{pmatrix} = \left\langle \begin{pmatrix} 0 \\ \widetilde{a_i} \end{pmatrix} \right\rangle_{i=0}^{n}$$

for some $g = (g_0, \ldots, g_m, 0, \ldots)$. This relation between the two continued fractions is easily verified evaluating both continued fractions starting at the bottom, and using that given $\alpha, \beta \in R^{m+1}$ and $x \in R^{\mathbb{N}}$ we have

$$\begin{pmatrix} 0 \\ \widetilde{\alpha} \end{pmatrix} + \frac{1}{\begin{pmatrix} x \\ \widetilde{\gamma} \end{pmatrix}} = \begin{pmatrix} y \\ \widetilde{\alpha + \frac{1}{\gamma}} \end{pmatrix}$$

for some $y \in R^{\mathbb{N}}$.

It is well known, at least in the case of commutative entries, that the Jacobi-Perron continued fraction is also related to Hessenberg banded matrices (see [14]). On the other hand, we have given in Theorem 2 a continued fraction for the first column of a Hessenberg matrix. Let us see what is the connection between them.

Let D_n be a Hessenberg like in (8). We will additionally assume that $d_{i,j} = 0$ for $i < j - m$ and $d_{i,i+m}$ is invertible for some $m \in \mathbb{N}$. Thus D_n has $m + 2$ nonzero diagonals with the lowermost diagonal formed by -1s and the uppermost nonzero diagonal formed by invertible elements. Assume that D_n is invertible. Let $v_n = D_n^{-1} e_0$ and $v_n^{(m)}$ the vector formed by the first m entries of v_n. In the notation of the previous sections this vector is

$$v_n^{(m)} = (P_{n-1}^1 P_n^{-1}, P_{n-2}^2 P_n^{-1}, \ldots, P_{n-m}^m P_n^{-1}).$$

Using the Jacobi-Perron algorithm the vector $v_n^{(m)}$ can be represented as the continued fraction $v_n^{(m)} = \langle a_i \rangle_{i=0}^n$ with

$$a_i = \begin{cases} 0, & i = 0 \\ (0, \ldots, 0, d_{0,i-1}, d_{1,i-1}, \ldots, d_{i-1,i-1}), & i = 1, \ldots, m \\ (d_{i-m,i-1}, d_{i-m+1,i-1}, \ldots, d_{i-1,i-1}) d_{i-m-1,i-1}^{-1}, & i = m+1, \ldots, n \end{cases}$$

See Theorem 2 of [14] for a proof in the commutative case.

This continued fraction can be compared to (12), which in this case can be written like $\binom{v_n}{0} = \langle \binom{0}{c_i} \rangle_{i=0}^n$, with

$$c_i = \begin{cases} 0, & i = 0, \\ (d_{i-1,i-1}, d_{i-2,i-1}, \ldots, d_{0,i-1}, 0, \ldots,), & i = 1, \ldots, m, \\ (d_{i-1,i-1}, d_{i-2,i-1}, \ldots, d_{i-m-1,i-1}, 0, \ldots), & i = m+1, \ldots, n. \end{cases}$$

2 Laurent series in $\frac{1}{z}$.

2.1 The continued fraction algorithm

Let R denote a unital ring. As before we do not assume that R is commutative. Let $R((z^{-1}))$ be the module over R of formal series of the form $\sum_{i=N}^{\infty} r_i z^{-i}$ with coefficients $r_i \in R$. $R((z^{-1}))$ becomes a ring using the standard multiplication of formal series and assuming that the variable z commutes with R. We shall denote by $R[z]$ the ring of polynomials in z with coefficients in R. Given $f \in R((z^{-1}))$ we shall denote by $[f]$ the polynomial part of f. The polynomial part of an element of $R((z^{-1}))^{\mathbb{N}} \times R((z^{-1}))^{\mathbb{N}}$ is defined taking the polynomial part of each of its components.

We now describe an algorithm to expand elements of $R((z^{-1}))^{\mathbb{N}} \times R((z^{-1}))^{\mathbb{N}}$ in a continued fraction with coefficients in $R[z]^{\mathbb{N}} \times R[z]^{\mathbb{N}}$.

In the algorithm to expand a vector of $R((z^{-1}))^{\mathbb{N}} \times R((z^{-1}))^{\mathbb{N}}$ in a continued fraction we need to write a given element r in the form $\frac{1}{\phi}$, where the inverse of ϕ is understood in the sense of (1). Since this inversion is not an involution, ϕ is not obtained from r by finding its inverse. Instead we have that $r = \frac{1}{\phi}$ for

$$\phi = \begin{pmatrix} r_{1,1} r_{1,0}^{-1} & r_{1,2} r_{1,0}^{-1} & \cdots \\ r_{1,0}^{-1} & r_{2,0} r_{1,0}^{-1} & \cdots \end{pmatrix}. \tag{15}$$

We call ϕ the anti-inverse of r. Clearly the anti-inverse exists if and only if $r_{1,0}$ is invertible. (An element of $R((z^{-1}))$ is invertible if and only if its first nonzero coefficient is invertible in R.)

2.1.1 The algorithm.

Starting with a given $f \in R((z^{-1}))^{\mathbb{N}} \times R((z^{-1}))^{\mathbb{N}}$ we put $f_0 = f$ and apply repeatedly the following two steps:

1. Find a_i and r_i such that $f_i = a_i + r_i$ and $a_i = [f_i]$ is the polynomial part of f_i.
2. Find f_{i+1} the anti-inverse of r_i. If r_i does not have an anti-inverse the algorithm stops.

Continuing this procedure we get a sequence of coefficients $a_i \in R[z]^{\mathbb{N}} \times R[z]^{\mathbb{N}}$ and residuals $r_i \in R((z^{-1}))^{\mathbb{N}} \times R((z^{-1}))^{\mathbb{N}}$ such that

$$f = a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_n + r_n}}}. \tag{16}$$

If r_{n-1} has an anti-inverse we call $\langle a_i \rangle_{i=0}^n$ the n-th convergent of f. The convergents of f can always be evaluated (recall definition in Subsection 1.1). This follows from the fact that the coefficients obtained by the algorithm satisfy the restrictions

$$\begin{aligned} \deg c_{i,0} &> \deg b_{i,k} \quad \text{for } k \geq 0, \\ \deg c_{i,0} &> \deg c_{i,k} \quad \text{for } k > 0, \end{aligned} \tag{17}$$

for all $i \geq 1$. So we do not get an arbitrary continued fraction by the continued fraction algorithm. On the other hand, if a sequence $\{a_i\}_{i=0}^n$ with $a_i \in R[z]^{\mathbb{N}} \times R[z]^{\mathbb{N}}$, satisfies (17), then expanding $\langle a_i \rangle_{i=0}^n$ in a continued fraction we obtain the sequence of coefficients $\{a_i\}_{i=0}^n$ and the n-th remainder term r_n equal to zero.

2.1.2 Convergence of the continued fraction algorithm.

Let $\phi_n = \langle a_i \rangle_{i=0}^n$ be the n-th convergent of f. Let us write $a_i = \binom{b_i}{c_i}$, $i \geq 0$, $m_n = \sum_{i=1}^n \deg(c_{i,0})$, and $\phi_n = \binom{\phi_{n,1}}{\phi_{n,2}}$.

Theorem 3. *We have*

$$\begin{aligned} f_{1,k} - \phi_{n,1,k} &= \begin{cases} O\left(z^{-(m_n+2)}\right) & \text{if } k < n, \\ O\left(z^{-(m_n+1)}\right) & \text{otherwise,} \end{cases} \\ f_{2,k} - \phi_{n,2,k} &= O\left(z^{-(m_n+1)}\right). \end{aligned} \tag{18}$$

Proof. We assume without lost of generality that $f = O(z^{-1})$. We will prove the theorem by induction on n. The theorem is easily checked for $n = 1$. Suppose that it is true for $n-1$. Let us write $f = \frac{1}{a_1+r}$ and $\phi_n = \frac{1}{a_1+\rho}$. Since ρ is the $(n-1)$-th convergent of r, we have by induction that

$$r_{1,k} - \rho_{1,k} = \begin{cases} O\left(z^{-(m_n - \deg(c_{1,0})+2)}\right) & \text{if } k < n, \\ O\left(z^{-(m_n - \deg(c_{1,0})+1)}\right) & \text{otherwise,} \end{cases}$$

$$r_{2,k} - \rho_{2,k} = O\left(z^{-(m_n - \deg(c_{1,0})+1)}\right).$$

Using (1) we can compute $f - \phi_n$ in terms of r, ρ and a_1. After some computations we get

$$\begin{aligned} f_{1,k} - \phi_{n,1,k} = (c_{1,0} + r_{2,0})^{-1} \big(r_{1,k-1} - \rho_{1,k-1} + \\ + (\rho_{2,0} - r_{2,0})(c_{1,0} + \rho_{2,0})^{-1}(b_{1,k-1} + \rho_{1,k-1})\big), \\ f_{2,k} - \phi_{n,2,k} = (c_{1,0} + r_{2,1})^{-1} \big(r_{2,k+1} - \rho_{2,k+1} + \\ + (\rho_{2,0} - r_{2,0})(c_{1,0} + \rho_{2,1})^{-1}(c_{1,k+1} + \rho_{2,k+1})\big), \end{aligned}$$

for $k \geq 0$, with the conventions $h_{1,-1} = q_{1,-1} = 0$ and $b_{1,-1} = 1$.

The degree of the right side can be estimated to get (18). For this we use (17) and the induction hypothesis. □

The continued fraction algorithm stops or continues indefinitely. We observe that the continued fraction of $\binom{f_1}{f_2}$ stops if and only if the continued fraction of $\binom{f_1}{0}$ stops.

Corollary 1. *If the continued fraction algorithm never stops it converges.*

Proposition 3. *Suppose that the base ring R is a division ring and that the continued fraction algorithm of $f = \binom{f_1}{f_2}$ stops in the n-th step. Then the n-th component of f_1 is equal to a linear combination with left polynomial coefficients of the first $n-1$ components of f_1.*

Proof. The proof proceeds by induction on n. If the algorithm stops in step 0 we must have $r_{0,1,0} = 0$. Thus the induction hypothesis is true for $n = 0$. Assume that the theorem is true for $n-1$ and that the continued fraction algorithm of f stops in the n-th step. Let us write $f = \frac{1}{a+g}$ with $a = \binom{a_1}{a_2} \in R[z]^{\mathbb{N}} \times R[z]^{\mathbb{N}}$, $a_{2,0} \neq 0$, and $g = O(z^{-1})$. It follows from the formula of the anti-inverse of f that $f_{1,k} = (a_{1,k-1} + g_{1,k-1})f_{1,0}$ for $k \geq 0$. Applying the induction hypothesis to g we get the desired result. □

2.1.3 The Hessenberg matrix $zI - D$.

Let D be an infinite Hessenberg matrix and D_n its n-th truncate. Then $zI - D$ and $zI_n - D_n$ are always invertible over $R((z^{-1}))$ with inverses

$$\rho(z, D) = \sum_{i=0}^{\infty} D^i z^{-i-1} \quad \text{and} \quad \rho(z, D_n) = \sum_{i=0}^{\infty} D_n^i z^{-i-1}.$$

Let $v(z) = \rho(z, D)e_0$ and $v_n(z) = \Pi_n \rho(z, D_n)e_0$. Since $(D^k)_{i,0} = (D_n^k)_{i,0}$ for $k = 0, 1, \dots, n$ and $i = 0, 1, \dots, n-1$, the vector v_n converges componentwise to $v(z)$. On the other hand, by Theorem 12 we have that $\binom{v_n}{0} = \langle \binom{b_i}{c_i} \rangle_{i=1}^n$ are the convergents of an infinite continued fraction. Thus we get that $\binom{v_n}{0}$ are the convergents of the continued fraction expansion of $\binom{v}{0}$.

More generally, let us take $w(z) = L\rho(z, D)e_0$ and $w_n(z) = L\Pi_n \rho(z, D_n)e_0$ for some upper triangular matrix L with invertible elements along the main diagonal. Then we have that $\binom{w_n}{0}$ are the convergents of the infinite continued fraction $\binom{w(z)}{0} = \langle \binom{b_i}{c_i} \rangle_{i=1}^{\infty}$. By Theorem 12 we have

$$\begin{aligned} b_{k+1} &= (l_{k+1,k}, l_{k+2,k}, \dots, l_{n,k}, \dots) l_{k,k}^{-1}, \\ c_{k+1} &= (l_{k-1,k-1} d_{k,k-1}^{-1}(z - d_{k,k}), -l_{k-2,k-2} d_{k-1,k-2}^{-1} d_{k-1,k}, \dots \\ &\qquad \dots - l_{0,0} d_{1,0}^{-1} d_{1,k}, -d_{0,k}, 0, \dots) l_{k,k}^{-1}. \end{aligned} \tag{19}$$

The following important observation will be used in Subsection 2.3 and Section 3. Using the equations (19) we can put the entries of L and D in terms of the coefficients of the continued fraction $b_{i,k}, c_{i,k}$ and the entries $d_{i+1,i}$. It thus follows that once we have fixed the diagonal $(d_{i+1,i})$ of D, we can recover the entries of D and L from the first column of $L\rho(z, D)$ by expanding it in a continued fraction.

2.2 Rational approximation

2.2.1 Overview of the classical case.

Given a Laurent series $h \in \mathbb{C}((z^{-1}))$, the convergents of its continued fraction solve a problem of best approximation to h by rational functions. They are the diagonal Padé approximants corresponding to the normal indices of h (see [17]). An important case of Laurent series is the one obtained from a positive linear functional $\Lambda : \mathbb{C}[\omega] \to \mathbb{C}$ (i.e., a measure on $\mathbb{R}$) by the formula

$$h(z) = \Lambda\left(\frac{1}{z - \omega}\right) = \sum_{i=0}^{\infty} \frac{\Lambda(\omega^i)}{z^{i+1}}.$$

In this case all the indices of $h(z)$ are normal; thus all the diagonal Padé approximants appear as convergents of the continued fraction of $h(z)$. The denominators $p_n(z)$ of these convergents satisfy $\Lambda(p_n(z)p_m(z)) = \delta_{n,m}$. That is, they are orthonormal polynomials with respect to the scalar product defined by Λ on $\mathbb{C}[\omega]$. These polynomials satisfy a three-term recurrence relation which can be put in matrix form like $\mathbf{p}J = z\mathbf{p}$, where J is a tridiagonal Jacobi matrix and $\mathbf{p} = (p_0, p_1, \dots)$ is an infinite row vector of polynomials. The matrix J is connected to Λ and $h(z)$ by the formulas $\Lambda(p(z)) = (p(J))_{0,0}$ and $h(z) = \sum_{i=0}^{\infty} (J^i)_{0,0} z^{-(i+1)}$. Expanding $h(z)$ in a continued fraction we can recover the entries of J from the coefficients of the continued fraction.

In Subsection 2.2.2 we will discuss a generalization of the setting described before. In this generalized setting, the formal series $h(z)$ is replaced by an infinite vector of formal series. The Padé approximants are replaced by the infinite simultaneous Padé approximants defined by A. Bultheel and M. van Barel in Section 2.7 of [9]. The functional Λ is replaced by a sequence of functionals (or a bilinear form) on the polynomials. The Jacobi matrix J is replaced by a Hessenberg matrix. We will focus on the relation of these objects to the continued fraction algorithm described before. Also, we will work mostly on an arbitrary unital ring R. We refer to [9] and [20] for a more thorough discussion of these questions when the base ring is taken to be $\mathbb{C}$.

2.2.2 Infinite simultaneous Padé approximants and biorthogonal polynomials.

Let $f \in R((z^{-1}))^{\mathbb{N}}$ be an infinite vector of Laurent series. The suitable definition of infinite simultaneous Padé approximant of f is also the most natural one.
Problem A. *Find a pair (p_n, q_n), with $p_n(z) \in R[z]$ a nonzero polynomial of degree at most n and $q_n \in R[z]^{\mathbb{N}}$, such that*

$$f_i p_n - q_{n,i} = \begin{cases} O\left(z^{-2}\right) & \text{if } i < n, \\ O\left(z^{-1}\right) & \text{otherwise.} \end{cases} \tag{20}$$

If (p_n, q_n) is a solution of Problem A, we call $q_n(p_n)^{-1}$ an *n-th infinite simultaneous Padé approximant* of f. It follows from (20) that $q_n(z)$ must be the polynomial part of $f(z)p_n(z)$ and the coefficients of $p_n(z)$ are found as a nonzero solution of the homogeneous system of equations $\sum_{i=0}^{n} a_{ik}x_i = 0$, $k = 0, \dots, n-1$. If R is taken to be a division ring then there is always at least one n-th infinite simultaneous Padé approximant for all $n \geq 0$.

Let $\binom{f}{0} = \langle a_i \rangle_{i=1}^{\infty}$ with $a_i = \binom{b_i}{c_i}$ and let $\phi_j = \langle a_i \rangle_{i=1}^{j}$ be the j-th convergent of $\binom{f}{0}$. Suppose that $\deg c_{i,0} = 1$ for $1 \leq i \leq k$. Then it is clear from Theorem 3 that ϕ_j^1 are infinite Padé approximants of f for all $1 \leq j \leq k$. This observation applies in particular to the convergents of the continued fraction of $\binom{v(z)}{0}$, where $v(z) = \rho(z, D)$ for some Hessenberg matrix D.

Let us see how the denominators of the infinite simultaneous Padé approximants satisfy certain orthogonality relations. Let $f = (f_1, f_2, \dots) \in R^{\mathbb{N}}((z^{-1}))$ be such that $f_i = \sum_{j=0}^{\infty} a_{i,j} z^{-j-1}$. Let us define the sequence of linear functionals $\Lambda_i : R[\omega] \to R$ such that $\Lambda_i(\omega^j) = a_{i,j}$. We extend Λ_i to the Laurent series in z^{-1} with coefficients in $R[\omega]$ by $\Lambda_i(h(\omega)z^{-j}) = \Lambda_i(h(\omega))z^{-j}$. Hence we have $f_i(z) = \Lambda_i(\frac{1}{z-\omega})$. One can check that if $\alpha, \beta \in R[z]$ satisfy the equation $f_i\alpha - \beta = O(z^{-2})$ then

$$\beta(z) = \Lambda_i\left(\frac{\alpha(z) - \alpha(\omega)}{z - \omega}\right), \quad \Lambda_i(\alpha(\omega)) = 0. \tag{21}$$

Thus if $q_n p_n^{-1}$ is the n-th simultaneous Padé approximant of f then p_n satisfies the biorthogonality relations $\Lambda_i(p_n) = 0$ for all $i < n$. In particular if $\Lambda_i(\cdot) = \langle \cdot, \omega^i \rangle$ for some inner product over the polynomials, then p_n is the n-th orthogonal polynomial with respect to this inner product.

The denominators of the convergents of an infinite continued fraction also satisfy orthogonality relations. Let P_n denote the denominator of the n-th convergent of the continued fraction of $\binom{f}{0}$. Then we have that $\deg P_n = m_n \geq n$ (recall notation of Theorem 3) and $\Lambda_i(P_n) = 0$ for $i < n$.

We say that n is a *normal* index of f if Problem A has a solution such that p_n is monic of degree n, and every other solution is a scalar multiple of this solution. We say that n is *strongly normal* if n is normal and $f_n p_n - q_{n,n} = c_n z^{-1} + \dots$ with c_n an invertible element of R and p_n the monic solution of Problem A. We remark that in the case when R is a division ring the index n is strongly normal if and only if n and $n+1$ are both normal.

Lemma 1. *Let D be an upper Hessenberg matrix with ones below the main diagonal. Let $v(z) = \rho(z, D)e_0$ and $p_n(z)$ be the monic solution of Problem A for $v(z)$. The following statements hold:*

(i) $v_i(z) = \alpha_i z^{-(i+1)} + \dots$ with α_i invertible element of R.

(ii)

$$p_n v_i - [p_n v_i] = \begin{cases} O\left(z^{-2}\right) & \text{if } i \neq n, \\ \alpha_i z^{-1} + \dots & \text{if } i = n. \end{cases} \tag{22}$$

(iii) The polynomials p_n satisfy that $p_0 = 1$ and the recurrence relation $z p_n = \sum_{i=0}^{n+1} p_i d_{n,i}$ for all $n \geq 0$.

Proof. Part (i) follows immediately from the definition of $\rho(z, D)$ and the fact that D is Hessenberg. Part (ii) follows from (i) and the definition of p_n.

(iii) As noted above, the polynomials p_n are the denominators of the continued fraction of $\rho(z, D)$. By Theorem 1, these polynomials satisfy the given recurrence relation. □

Let $f \in R((z^{-1}))^{\mathbb{N}}$ be such that $f_i = \sum_{j=0}^{\infty} a_{i,j} z^{-j-1}$. We define the matrix $(\Upsilon)_{i,j} = a_{i,j}$. The following theorem gives a characterization of the formal series f such that all their indices are strongly normal.

Theorem 4. *Let f and Υ be as above. The following statements are equivalent:*

(i) All indices of f are strongly normal.

(ii) The quasideterminants $|\Upsilon_n|_{n,n}$ are invertible for all n.

(iii) There are an infinite upper Hessenberg matrix D and an infinite, lower triangular, matrix L with invertible entries along the main diagonal, such that $f = L\rho(z, D)e_0$.

(iv) The continued fraction of $\binom{f}{0}$ never stops and $\deg c_{i,0} = 1$ for all $i \geq 1$.

Suppose that we are in the above situation. Let $\binom{\phi_{n,1}}{0}$ be the n-th convergent of $\binom{f}{0}$. Then $\phi_{n,1}$ is n-th infinite simultaneous Padé approximant of f for all $n \geq 0$ and $\phi_{n,1} = L\Pi_n \rho(z, D_n)e_0$.

Proof. (i) $\Rightarrow$ (iii). Let p_n be the monic solution of Problem A for f. Since $\{p_n\}_{n=0}^{\infty}$ is a basis of $R[z]$ we can write $z p_n$ as linear combination of $\{p_k\}_{k=0}^{n+1}$. That is, we have $p_0 = 1$ and

$$z p_n = \sum_{i=0}^{n+1} p_i d_{n,i}. \tag{23}$$

We define the Hessenberg matrix $D = (d_{i,j})_{i,j=0}^{\infty}$. By Lemma 1, the components of the vector $\rho(z, D)e_0$ form a basis of $R(z^{-1})$. Thus, there is a unique infinite matrix L such that $f = L\rho(z, D)e_0$.

We now prove that L is lower triangular with invertible elements in the main diagonal. In virtue of the discussion in 2.1.3 we know that the polynomials p_n are also the monic solutions of Problem A for $\rho(z, D)e_0$. We have

$$p_n f - [p_n f] = L(p_n \rho(z, D)e_0 - [p_n \rho(z, D)e_0]).$$

The left side of this equality satisfies (20). On the other hand, by Lemma 1 (ii) we know that the right side has the form $L(z^{-1}e_n + O(z^{-2}))$. Using this we get that L must be lower triangular and that $(L)_{n,n} = c_n$ (recall notation in the definition of strongly normal index). Since n is strongly normal c_n is invertible.

(iii) $\Rightarrow$ (i). First notice that if (p_n, q_n) is a solution of Problem A for a vector g and M is a lower triangular matrix, then (p_n, Mq_n) is a solution of Problem A for Mg. It follows that if M has invertible entries along the main diagonal, then g and Mg have the same strongly normal indices. Since all indices of $\rho(z, D)e_0$ are normal, we are done.

(iii) $\Rightarrow$ (iv) The coefficients of the expansion of $L\rho(z, D)e_0$ in continued fraction are given by formula (19). From this the desired implication follows.

(iv) $\Rightarrow$ (iii). Let $\binom{f}{0} = \langle a_k \rangle_{k=1}^{\infty}$ with $a_k = \binom{b_k}{c_k}$, $c_{k,0} = \alpha_k(z - \beta_k)$. We want to find D and L such that $f = L\rho(z, D)e_0$. Let us take $d_{i+1,i} = 1$. Using the equations (19) we can find $l_{i,k}$ and $d_{i,k}$ in terms of α_k, β_k, $b_{k,i}$ and $c_{k,i}$ such that the continued fraction of $L\rho(z, D)e_0$ coincides with the continued fraction of $\binom{f}{0}$. Since the continued fraction expansion is unique, we get (iii).

(ii) $\Rightarrow$ (iii) By Theorem 4.9.7 of [13], (ii) is equivalent to having the decomposition $\Upsilon = L_1 L_2^{-1}$ with L_1 lower triangular matrix with invertible entries in the main diagonal, and L_2 upper triangular matrix with ones in the main diagonal. We have $f = \Upsilon\rho(z, S)e_0$, where S is the shift to the right. Hence, $f = L_1\rho(z, L_2^{-1}SL_2)e_0 = L_1\rho(z, D)e_0$.

(iii) $\Rightarrow$ (ii) We have $\Upsilon\rho(z, S)e_0 = L_1\rho(z, D)e_0$. Any infinite upper Hessenberg matrix is conjugate to S by an upper triangular matrix with invertible elements in the main diagonal. That is, there is L_2 upper triangular such that $D = L_2^{-1}SL_2$. From this we get $\Upsilon\rho(z, S)e_0 = L_1L_2^{-1}\rho(z, S)e_0$. Hence $\Upsilon = L_1L_2^{-1}$. Therefore by Theorem 4.9.7 of [13] we get (ii).

By Theorem 3, (iii) implies that $\phi_{n,1}$ is the n-th infinite Padé approximants of f. Now, it follows from (ii) and the discussion in Subsection 2.1.3 that $\phi_{n,1} = L\Pi_n\rho(z, D_n)e_0$. □

Remark. If L is an invertible lower triangular matrix with entries in a non-commutative ring R, it need not be the case that L^{-1} is lower triangular, nor that the entries in the main diagonal of L are invertible. Nevertheless, we do have that L is invertible with lower triangular inverse if and only if the entries along the main diagonal of L are invertible. We made use of this fact, and of its analog for upper triangular matrices, in the proof of the preceding theorem.

2.2.3 Connection with multidimensional simultaneous Padé approximation.

Let R be a division ring. Let $f \in R((z^{-1}))^{m+1}$ be of the form $f = \sum_{i=N}^{\infty} a_i z^{-i-1}$, with $a_i \in R^{m+1}$. Thus $f = (f_0, f_1, \ldots, f_m)$ is a finite vector of Laurent series in $R((z^{-1}))$. For a multi-index $\mathbf{n} = (n_0, n_1, \ldots, n_m)$ we write $|\mathbf{n}| = n_0 + \ldots + n_m$.

Problem B. Find a pair $(p_{\mathbf{n}}, q_{\mathbf{n}})$, where $p_{\mathbf{n}}(z)$ is a nonzero polynomial of degree at most $|\mathbf{n}|$ and $q_{\mathbf{n}} = (q_{\mathbf{n},0}, q_{\mathbf{n},1}, \ldots, q_{\mathbf{n},m}) \in R[z]^{m+1}$, such that

$$f_j p_{\mathbf{n}} - q_{\mathbf{n},j} = O(z^{-(n_j+1)}), \; j = 0, \ldots, m. \tag{24}$$

Since R is a division ring Problem B always has a solution. The fraction $q_{\mathbf{n}} p_{\mathbf{n}}^{-1}$ is called an $\mathbf{n}$-th Hermite-Padé approximant for f. If $\deg p_{\mathbf{n}} = |\mathbf{n}|$ for all the solutions of Problem B then the multi-index $\mathbf{n}$ is called normal.

We can use the solutions of Problem A to solve Problem B. More specifically, let $f = (f_0, \ldots, f_m)$ be a finite vector and $\mathbf{n} = (n_0, \ldots, n_m)$ a multi-index. Let $\mathbf{g}$ be an infinite vector of formal series which has in its first components a permutation of the formal series

$$f_0, zf_0, \ldots, z^{n_0-1} f_0, f_1, \ldots, f_m, zf_m, \ldots, z^{n_m-1} f_m.$$

Then $(p, [pf])$ is a solution of Problem B with multi-index $\mathbf{n}$ if and only if $(p, [pg])$ is an $|\mathbf{n}|$-th solution of Problem A (recall that $[\cdot]$ denotes polynomial part). It follows that a multi-index $\mathbf{n}$ is normal for f with respect to Problem A if and only if $|\mathbf{n}|$ is a normal index for $\mathbf{g}$ with respect to Problem B.

We can use the infinite continued fraction to find Hermite-Padé approximants associated to a sequence of increasing normal multi-indices. Specifically, let f be a finite vector of formal series and $\{\mathbf{n}_k\}_{k=0}^{\infty}$ an increasing sequence of multi-indices; that is, $\mathbf{n}_0 = 0$ and $\mathbf{n}_{k+1} - \mathbf{n}_k = e_{j_k}$ for some j_k. Let g be the infinite vector such that $g_k = z^{n_{k,j_k}} f_{j_k}$. This vector satisfies that if $[p_k g] p_k^{-1}$ is a k-th infinite simultaneous Padé approximant for g then $[p_k f] p_k^{-1}$ is an Hermite-Padé approximant for f of multi-index $\mathbf{n}_k$. Now suppose that all the multi-indices $\mathbf{n}_k$ are normal. We expand g in a continued fraction and call $\binom{\pi_j}{0}$ its j-th convergent. Since all the indices of g are normal, Theorem 4 implies that π_j is the j-th infinite Padé approximant of g for all j. Let $i_0, i_1, \ldots, i_m$ be the indices such that $g_{i_k} = f_k$ for $k = 0, 1, \ldots, m$. Then the vector $(\pi_{j,i_0}, \pi_{j,i_1}, \ldots, \pi_{j,i_m})$ is the Hermite-Padé approximant of f associated to the multi-index $\mathbf{n}_j$ for all j.

In [18] Parusnikov finds a characterization of the finite vectors of formal series such that all their proper multi-indices are normal. This characterization is given in terms of the coefficients of their Jacobi-Perron continued fraction (see [18]). Theorem 4 and the discussion above allow us to give a similar characterization for the finite vectors of formal series such that the multi-indices in a given increasing sequence are normal. The increasing sequence of multi-indices $\{\mathbf{n}_i\}_{i=0}^{\infty}$ is normal for the vector $f = (f_0, \ldots, f_m)$ if and only if the continued fraction of the associated infinite vector g is of the form described in Theorem 4 (iii).

2.3 Decomposition of matrices.

Let R be a ring. We denote by $R^{op} = (R, *, +)$ the opposite ring of R. This is R with the same operation of addition and multiplication such that $a * b = ba$.

Let Υ be an $N \times N$ matrix with entries in a ring R, with $N \in \mathbb{N}$ or $N =$ [redacted]. Consider the problem of decomposing Υ in a product $L_1 B L_2$ where L_1 is a lower triangular matrix with ones in the main diagonal, L_2 is upper triangular matrix with ones in the main diagonal and B is an invertible diagonal matrix. In this subsection we will show how the continued fraction algorithm can be used to solve this problem.

First let us suppose that $\Upsilon = L_1 B L_2$ with L_1, L_2 and B as before. Consider the formal series $g_1(z) = \Upsilon \rho(z, H_2) e_0$, where H_2 is any $N \times N$ upper Hessenberg matrix with ones below the main diagonal (e.g. $(H_2)_{i,j} = \delta_{i,j+1}$ is the shift to the right). Then we have that $g_1(z) = L_1 B \rho(z, D_2) e_0$, where $D_2 = L_2 H_2 L_2^{-1}$. Since D_2 is an upper Hessenberg matrix, we know by the discussion in 2.1.3 that the continued fraction algorithm of $\binom{g_1}{0}$ never stops before the N-th step. Moreover, in the k-th step of the continued fraction algorithm for $\binom{g_1}{0}$ we can obtain the k-row of L_1, the entry $B_{k,k}$ of B, and the k-th column of D_2 from the coefficients of the continued fraction. Specifically, if $\binom{g_1}{0} = \langle \binom{b_i}{c_i} \rangle_{i=1}^{N}$ then

$$\begin{aligned} b_{k+1} &= (L_{1,k,k+1}, L_{1,k,k+2}, \ldots, L_{1,k,n}, \ldots), \\ c_{k+1} &= -(B_{k-1,k-1}(d_{k,k} - z), B_{k-2,k-2} d_{k-1,k}, \ldots B_{0,0} d_{1,k}, d_{0,k}, 0, \ldots) B_{k,k}^{-1} \end{aligned}$$

where $(D_2)_{i,j} = d_{i,j}$.

Taking transpose we adapt the previous argument to find the lower triangular matrix L_1. Taking into account that $(AB)^t = A^t * B^t$, we need to perform all the operations in R^{op}. More specifically, we apply the continued fraction to $\binom{g_2}{0}$, where $g_2(z) = \Upsilon^t * \rho(z, H_1^t) e_0$ for some lower Hessenberg matrix H_1 with ones over the main diagonal. All the operations of the continued fraction and $\rho(z, H_1^t)$ are computed in R^{op} (i.e., $\rho(z, H_1^t) = I + H_1^t z^{-1} + H_1^t * H_1^t z^{-2} + \ldots$). As a byproduct of the continued fraction, we also get the entries of the lower Hessenberg matrix $D_1 = L_1^{-1} H_1 L_1$.

Conversely, let us suppose that the continued fraction of $g_1(z) = \Upsilon \rho(z, H_2) e_0$ never stops before the N-th step and that its coefficients satisfy $\deg c_{k,0} = 1$ for all $k \leq N$. Using the equations (19) we can find an $N \times N$ upper Hessenberg matrix D_2 and an $N \times N$ lower triangular matrix P with ones in the main diagonal such that $P\rho(z, D_2)e_0$ has the same expansion in continued fraction of $g_1(z)$. This implies that $g_1(z) = P\rho(z, D_2) e_0$ (In the case $N =$ [redacted] by Corollary 1). Thus, we have that $P^{-1}\Upsilon \rho(z, H_2) e_0 = \rho(z, D_2) e_0$, but this is only possible when $P^{-1}\Upsilon$ is equal to an invertible upper triangular matrix Q, hence $\Upsilon = PQ$. Taking B the diagonal matrix formed by the diagonal elements of Q, $L_2 = QB^{-1}$ and $L_1 = P$, we obtain the desired decomposition for Υ.

Summarizing, given $N \leq$ [redacted] and an $N \times N$ matrix Υ, we choose a lower Hessenberg matrix H_1 with ones over the main diagonal and an upper Hessenberg matrix H_2 with ones below the main diagonal. We apply the continued fraction algorithm to the vectors $\binom{f_1}{0}$ and $\binom{f_2}{0}$, with $g_1 = \Upsilon \rho(z, H_2) e_0$ and

$g_2 = \Upsilon^t * \rho(z, H_1^t)e_0$. If one of these two continued fractions stops before the N-th step or one of their coefficients does not satisfy $\deg c_{k,0} = 1$, then the desired decomposition for Υ does not exist. Otherwise, we obtain the entries of the matrices L_1, L_2 and B such that $\Upsilon = L_1 B L_2$ from the coefficients of these continued fractions. Additionally, we also get the coefficients of the Hessenberg matrices $D_1 = L_1^{-1} H_1 L_1$ and $D_2 = L_2 H_2 L_2^{-1}$. This last observation will be used in Section 3 in the solution of nonlinear differential equations.

2.4 Analytic convergence

In this subsection we let the field of scalars R be the complex numbers.

Let D be an infinite upper Hessenberg matrix with complex entries such that $d_{i+1,i} \neq 0$. We already know that the convergents of the continued fraction (19) are the vectors $\binom{w_n(z)}{0}$ where $w_n(z) = L\Pi_n R(z, D_n)e_0$. Thus the convergence of this continued fraction reduces to the convergence of $R(z, D_n)e_0$ to an analytic counterpart of $\rho(z, D)e_0$.

We consider D as an operator in l_2 with domain the infinite vectors with finitely many nonzero entries. Since D is densely defined, it has an adjoint D^* which is a closed operator with domain $\{x \in l_2 : \overline{D}^t x \in l_2\}$.

Let us denote by $p_{n-k}^k(z)$ the polynomials defined by the recurrence relation

$$z p_{n-k}^k(z) = \sum_{i=0}^{n+1} p_{i-k}^k(z) d_{i,n},$$

with initial conditions $p_0^k(z) = d_{k,k-1}^{-1}$ and $p_i^k(z) = 0$ for $i < 0$ (recall notation of Subsection 1.2). We say that D is *determinate* if $\sum_{n,k=0}^{\infty} |p_{n-k}^k(z)|^2 =$ [redacted].

This definition agrees with the definition of determinate given by H. S. Wall in [22] for complex Jacobi matrices. There is also a theorem of invariance of determinacy by bounded lower triangular perturbations (Theorem 3 of [19]).

Definition 2. *For an infinite matrix A and an infinite sequence of indices $S \subset \{0, 1, \dots\}$ define the set:*

$$\Theta_S(A) = \{z : \limsup_{n \in S} \|R(z, (A)_n)\| < \text{[redacted]}\}.$$

When $S = \{0, 1, \dots\}$ we omit the subscript and write simply $\Theta(D)$.

The following theorem is proved in [19].

Theorem 5. *Suppose we have one of the two cases:*

(i) D is determinate and $S = \{0, 1, \dots\}$,

(ii) $\sum_{n \in S} |d_{n,n-1}|^{-2} =$ [redacted].

Then

$$\Theta_S(D) = \{z : R(z, D_n) \to R(\overline{z}, D^*)^*, n \in S\} \subset \Omega(D),$$

and for $x \in l_2$ the convergence $\lim_{n \in S} R(z, D_n)x = R(\overline{z}, D^*)^* x$ is uniform in compact subsets of $\Theta_S(D)$.

Thus the analytic counterpart of $\rho(z, D)$ is $R(\overline{z}, D^*)^*$ and we have strong convergence of $R(z, D_n)$ to $R(\overline{z}, D^*)^*$ on $\Theta(D)$.

Corollary 2. *Suppose L is an infinite lower triangular matrix and D satisfies (i) or (ii) of the previous theorem. Then the continued fraction of the first column of $LR(\overline{z}, D^*)^*$ converges component-wise and uniformly on compact subsets of $\Theta(D)$. If $L \in B(l_2)$ then the convergence is in l_2 and uniform on compact subsets of $\Theta(D)$.*

The preceding theorems show that it is important to know the properties of the set $\Theta(D)$. Let us define the sets

$$\widetilde{\Gamma}(A) = \overline{\{\langle Ax, x\rangle : \|x\| = 1, x \in C_0\}},$$
$$\widetilde{\Gamma}_{ess}(A) = \bigcap_C \widetilde{\Gamma}(A + C),$$

where C ranges through all the compact operators with lower triangular matrix. The following properties of the sets $\Theta_S(D)$, $\widetilde{\Gamma}(D)$ and $\widetilde{\Gamma}_{ess}(D)$ are proven in [19].

(i) $\Theta_S(D)$ is an open set.

(ii) $(\widetilde{\Gamma}(D))^c \subset \Theta(D)$.

(iii) If $z \in \Theta_S(D) \cap \Omega(D + C)$ and C is a compact operator, then $z \in \Theta_S(D + C)$.

(iv) $\Omega(D)\backslash\widetilde{\Gamma}_{ess}(D) \subset \Theta_S(D)$.

(v) If $\lim_{n\in S} d_{n-1,n} = 0$ then $\Theta_S(D) = \Omega(D)$.

Let us denote by $\Omega^\infty(\mathcal{D})$ be the union of the connected components of $\Omega(\mathcal{D})$ which have nonempty intersection with $(\widetilde{\Gamma}(D))^c$. Let us also define the set $Z_S^\infty(D)$ by

$$Z_S^\infty(D) = \{z : \exists\{z_{n_k}\}, z_{n_k} \in \sigma(D_{n_k}), z_{n_k} \to z, n_k \to$$ [redacted] $$, n_k \in S\},$$

where $\sigma(D_n)$ is the set of eigenvalues of D_n. Finally for $K \subset \mathbb{C}$ let us denote by $v(K, D_n)$ the number of eigenvalues of D_n in K.

Remark. It is known that the eigenvalues of D_n are the zeroes of $p_n(z)$, which in turn are the poles of the n-th convergent of the continued fraction of $R(\overline{z}, D^*)^* e_0$. Thus $Z_S^\infty(D)$ and $v(K, D_n)$ can be described in terms of the poles of the convergents of the continued fraction of $R(\overline{z}, D^*)^* e_0$.

The following theorem is proved in [19].

Theorem 6. *Suppose that $\widetilde{\Gamma}(D) \neq \mathbb{C}$ and that $\{|d_{n-1,n}|\}_{n\in S}$ is a bounded sequence. Then we have*

(i) For every compact set $K \subset \Omega^\infty(D)$, there is an infinite subsequence $S_1 \subset S$ such that $K\backslash\Theta_{S_1}(D)$ is at most finite.

(ii) $\Omega^\infty(D)\backslash Z_S^\infty(D) \subset \Theta_S(D)$.

(iii) For every compact set $K \subset \Omega^\infty(D)$, the sequence $\{v(K, D_n)\}_{n\in S}$ is uniformly bounded.

3 Solving nonlinear integrable systems.

3.1 Overview of the Toda lattice.

The finite nonperiodic Toda lattice is the system of ordinary differential equations

$$\frac{d}{dt}a_n = b_{n+1}a_n - a_n b_n,$$
$$\frac{d}{dt}b_n = a_n - a_{n-1},$$

$n = 0, 1, \ldots, N$, where $a_{-1} = b_{N+1} = 0$. These equations can be put in the Lax form

$$\frac{dL}{dt} = [L, (L)_-], \tag{25}$$

where L is the $N \times N$ tridiagonal matrix

$$L = \begin{pmatrix} b_0 & 1 & & & \\ a_0 & b_1 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{N-2} & b_{N-1} & 1 \\ & & & a_{N-1} & b_N \end{pmatrix}, \tag{26}$$

and $(L)_-$ denotes the strictly lower triangular part of L.

One has the following procedure to solve the Cauchy problem for the Toda equations. Let us define the matrix $\Upsilon(t) = e^{tL(0)}$ and the rational function $h_t(z) = \sum_{i=0}^{\infty} \Upsilon_{i,0} z^{-i-1}$. Then the entries of L can be obtained from the coefficients of the continued fraction expansion of $h_t(z)$ (see [17, 16]). A similar procedure applies when $N =$ [redacted]; that is, for semi-infinite systems. In this section we will see how this procedure can be applied to a broad class of nonlinear systems of differential equations that generalizes the Toda equations. The standard continued fraction is replaced by the continued fraction of infinite vectors described in the previous sections and the matrix L is replaced by a pair of Hessenberg matrices. Instead of the exponential evolution of Υ, we will have that Υ satisfies a linear system of differential equations with constant coefficients.

3.2 Preliminaries

Let R be a unital ring. Throughout this section N is a natural number or $N =$ [redacted] and all matrices have size $N \times N$ and entries in R.

For a matrix A we denote by $(A)_-$ and $(A)_+$ the strictly lower triangular and the upper triangular part of A, respectively. If A and B are upper Hessenberg matrices, we write $A \sim_u B$ if there is an upper triangular matrix Δ with invertible elements along the main diagonal such that $A = \Delta^{-1}B\Delta$. This defines an equivalence relation. If A and B are lower Hessenberg matrices we write $A \sim_l B$ if they are conjugate by a lower triangular matrix with ones in the main diagonal.

Let ∂_i, $i \in I$ be a family of derivations on R. We call $c \in R$ a constant if $\partial_i c = 0$ for all $i \in I$.

Lemma 2. *Let D be an upper Hessenberg matrix, A_i, $i \in I$, upper triangular matrices and $\xi \in R$ an invertible element of R. The following are equivalent:*

(i) $\partial_i D = [D, A_i]$, for all $i \in I$ and $\partial_i \xi = \xi (A_i)_{0,0}$.

(ii) There are a constant upper Hessenberg matrix H and an upper triangular matrix L with invertible entries in the main diagonal such that $D = L^{-1}HL$, $\xi = (L)_{0,0}$ and $A_i = L^{-1}\partial_i L$ for all $i \in I$.

Moreover, if either (i) or (ii) hold, then we also have that:

(iii) Given a constant upper Hessenberg matrix H such that $D \sim_u H$, there is unique upper triangular matrix L with invertible elements in the main diagonal that satisfies (ii) for H.

(iv) When $N <$ [redacted] we have that $\partial_i p_N = 0$ for all $i \in I$, where $p_N = \det(zI - D)$ is the characteristic polynomial of D. ($\det(\cdot)$ is taken as defined in Subsection 1.2.)

Proof. First we assume that $N <$ [redacted].

(i) $\Rightarrow$ (ii). Let us define the row vector of polynomials $\mathbf{p} = (p_0, p_1, \ldots, p_{N-1})$ by $p_0 = \xi$ and the recurrence relation

$$\mathbf{p}D = z\mathbf{p} + p_N c_N e_{N-1}^t, \tag{27}$$

where c_N is chosen so that p_N be monic and $e_{N-1}^t = (0, \ldots, 1)$. We remark that p_N in the above equation is the characteristic polynomial of D. Taking derivatives in (27) and using (i) we get

$$(\partial_i \mathbf{p} - \mathbf{p}A_i)(D - zI) = (\partial_i(p_N c_N) - p_N c_N A_{N-1,N-1})\, e_{N-1}^t. \tag{28}$$

The first component of $\partial_i \mathbf{p} - \mathbf{p}A_i$ is zero. Since $D - zI$ is a Hessenberg matrix, $v(D - zI) = 0$ and $v_0 = 0$ implies that $v = 0$ for any row vector v. We apply this in the last equation and get that $\partial_i \mathbf{p} - \mathbf{p}A_i = 0$.

Let $\chi(z) = (\lambda_0(z), \lambda_1(z), \ldots, \lambda_{N-1}(z))$ be a vector of polynomials such that $\lambda_0(z) = 1$, $\lambda_k(z) = \alpha_k z^k + \ldots$ with α_k invertible, and $\partial_i \chi(z) = 0$ for all $i \in I$ (e.g. $\chi(z) = (1, z, \ldots, z^{N-1})$). Let L be the upper triangular matrix defined by $\mathbf{p}(z) = \chi(z)L$. Then L has invertible elements in the main diagonal and $(L)_{0,0} = \xi$. Since $\partial_i \mathbf{p} - \mathbf{p}A_i = 0$, we obtain that $\partial_i L = LA_i$. Using this and (i) we get that $\partial_i(LDL^{-1}) = 0$ for all $i \in I$. Hence $D = L^{-1}HL$ where H is a constant upper Hessenberg matrix.

(ii) $\Rightarrow$ (i) This follows computing $\partial_i(L^{-1}HL)$ and using that $\partial_i L = LA_i$ and $\partial_i H = 0$ for all $i \in I$.

The proof of the case $N =$ [redacted] can be made in the same way as in the case $N <$ [redacted]. In this case, instead of equation (27), we define the infinite vector $\mathbf{p}(z)$ by $p_0 = \xi$ and $\mathbf{p}D = z\mathbf{p}$.

(iii) Suppose that $D \sim_u H$. One can prove that there is a unique upper triangular matrix L with invertible elements in the main diagonal such that $D = L^{-1}HL$ and $(L)_{0,0} = \xi$. By (ii) there is a constant upper Hessenberg matrix

H_1 and an upper triangular matrix L_1 such that $D = L_1^{-1} H_1 L_1$, $\xi = (L_1)_{0,0}$ and $A_i = L_1^{-1} \partial_i L_1$ for all $i \in I$. Hence $H_1 = (LL_1^{-1})^{-1} H LL_1^{-1}$ and $(LL_1^{-1})_{0,0} = 1$. Since H and H_1 are constant, we obtain that LL_1^{-1} is constant too. Hence

$$L^{-1}\partial_i L = ((LL_1^{-1})L_1)^{-1}\partial_i((LL_1^{-1})L_1) = L_1^{-1}\partial_i L_1 = A_i,$$

which proves (iii).

(iv) It is enough to prove that (i) implies that $\partial_i p_N = 0$ for all $i \in I$. Since (i) implies that $\partial_i \mathbf{p} - \mathbf{p} A_i = 0$, we have by (28) that

$$\partial_i(p_N \gamma_N) = p_N c_N A_{N-1,N-1}. \tag{29}$$

Comparing the principal coefficient on both sides and using that p_N is monic we get that $\partial_i c_N = c_N A_{N-1,N-1}$. Plugging this back in (29) we get that $\partial_i p_N = 0$ for all $i \in I$. □

Taking transpose in the statement of the previous lemma and changing the ring R to R^{op}, we see that it is still true if we change "upper Hessenberg" by "lower Hessenberg" and "upper triangular" by "lower triangular".

3.3 A generalization of the Toda lattice.

Let $f(u,v) = \sum_{0 \le n,m \le k} \alpha_{n,m} u^n v^m$ be a polynomial with coefficients in R. Let D_1, D_2 and M be $N \times N$ matrices, with $N \le$ [redacted]. We denote by $\widehat{f}(D_1, M, D_2)$ the matrix valued function

$$\widehat{f}(D_1, M, D_2) = \sum_{0 \le n,m \le k} \alpha_{n,m} D_1^n M D_2^m.$$

When $M = I_N$ we simply write $f(D_1, D_2)$. Notice that $\widehat{f}(D_1, M, D_2)$ is well defined when D_1 is a lower Hessenberg matrix, D_2 is an upper Hessenberg matrix and M is any matrix.

Let $f_i(u,v)$, $i \in I$ be a family of polynomials with coefficients in the center of R. We denote by $\mathcal{F}$ the set of triples (D_1, D_2, γ) such that D_1 is a lower Hessenberg matrix with ones below the main diagonal, D_2 is a upper Hessenberg matrix, $\gamma \in R$ is an invertible element of R, and for all $i \in I$ they satisfy the equations

$$\partial_i D_1 = [D_1, (f_i(D_1, D_2))_-], \tag{30}$$

$$\partial_i D_2 = [D_2, -(f_i(D_1, D_2))_+], \tag{31}$$

$$\partial_i \gamma = -\gamma (f_i(D_1, D_2))_{0,0}. \tag{32}$$

Let H_1 and H_2 be constant matrices, H_1 lower Hessenberg matrix with ones below the main diagonal and H_2 upper Hessenberg matrix. We denote by $\mathcal{F}_{H_1,H_2}$ the subset of $\mathcal{F}$ of triples (D_1, D_2, γ) such that $D_1 \sim_l H_1$ and $D_2 \sim_u H_2$

Let H_1 and H_2 be as before. Let us denote by $\mathcal{S}_{H_1,H_2}$ the set of $N \times N$ matrices Υ that satisfy

$$\partial_i \Upsilon = \widehat{f}_i(H_1, \Upsilon, H_2), \quad \forall i \in I, \tag{33}$$

and that can be decomposed in the form $\Upsilon = L_1 L_2^{-1}$, with L_1 lower triangular matrix with ones in the main diagonal and L_2 upper triangular matrix with invertible elements in the main diagonal. We remark that such decomposition, if it exists, is unique.

Theorem 7. *The following statements hold true:*

(i) $\mathcal{F} = \bigcup_{H_1,H_2} \mathcal{F}_{H_1,H_2}$.

(ii) The mapping $\Phi_{H_1,H_2} : \mathcal{S}_{H_1,H_2} \to \mathcal{F}_{H_1,H_2}$ *defined by*

$$\Phi_{H_1,H_2}(\Upsilon) = (L_1^{-1} H_1 L_1, L_2^{-1} H_2 L_2, (\Upsilon)_{0,0}^{-1})$$

is a bijection from $\mathcal{S}_{H_1,H_2}$ *to* $\mathcal{F}_{H_1,H_2}$.

(iii) Suppose that $\Phi_{H_1,H_2}(\Upsilon) = (D_1, D_2, \gamma)$. *The entries of* D_1 *and* D_2 *can be computed from the continued fraction of* $g_1(z) = \Upsilon\rho(z, H_2)e_0$ *and* $g_2(z) = \Upsilon^t * \rho(z, H_1^t)e_0$, *where the entries of the vector* $\rho(z, H_1^t)e_0$ *are computed in* R^{op}. *More specifically, we have*

$$\begin{pmatrix} g_1 \\ 0 \end{pmatrix} = \cfrac{1}{a_1 - \cfrac{1}{a_2 - \cfrac{1}{a_3 - \cfrac{1}{\cdots} d_{2,3,2}} d_{2,2,1}} d_{2,1,0}}, \quad \begin{pmatrix} g_2 \\ 0 \end{pmatrix} = \cfrac{1}{b_1 - \cfrac{1}{b_2 - \cfrac{1}{b_3 - \cfrac{1}{\cdots} * \eta_2} * \eta_1} * \eta_0},$$

where $a_k = \binom{a_{k,1}}{a_{k,2}}$, $b_k = \binom{b_{k,1}}{b_{k,2}}$ *with*

$$\begin{aligned} a_{k+1,2} &= -(d_{2,k,k} - z, d_{2,k-1,k}, \ldots, d_{2,0,k}, 0, \ldots), \\ b_{k+1,2} &= -(d_{1,k,k} - z, d_{1,k-1,k}, \ldots, d_{1,0,k}, 0, \ldots) * \eta_k, \end{aligned}$$

and $\eta_k = \left(\prod_{i=k}^{0} h_{2,i,i-1}\right) \gamma \prod_{i=0}^{k} d_{2,i,i-1}^{-1}$ *for all* $k \geq 0$. *The continued fraction of* $\binom{g_2}{0}$ *is taken over the ring* R^{op}.

Note: Recall that, in order to apply the continued fraction algorithm when $N <$ [redacted], the finite vectors $g_1(z)$ and $g_2(z)$ are transformed into infinite vectors by adding zeroes after their last component.

Proof. (i) Suppose that (D_1, D_2, γ) satisfy (30)-(32). One can check that the conditions in Lemma 2 (i) hold true taking $D = D_2$, $A_i = -(f_i(D_1, D_2))_+$ and $\xi = \gamma$. Thus, there is a constant upper Hessenberg matrix H_2 and an upper triangular matrix L_2 such that $D_2 = L_2^{-1} H_2 L_2$ and $-(f_i(D_1, D_2))_+ = L_2^{-1} \partial_i L_2$, $i \in I$.

By the remark made at the end of Lemma 2, this lemma can also be applied to lower Hessenberg matrices. In this case one takes $D = D_2$, $A_i = (f_i(D_1, D_2))_-$ and $\xi = 1$. Thus, there are H_1 and L_1 such that $D_1 = L_1^{-1} H_1 L_1$ and $(f_i(D_1, D_2))_- = L_1^{-1} \partial_i L_1$, $i \in I$.

(ii) It is follows from (iii) of Lemma 2 that the mapping Φ_{H_1,H_2} is injective.

Suppose that Υ satisfies (33) and $\Upsilon = L_1 L_2^{-1}$. Taking derivatives on both sides of this equality we get

$$L_1^{-1} \partial_i L_1 - L_2^{-1} \partial_i L_2 = f_i(L_1^{-1} H_1 L_1, L_2^{-1} H_2 L_2) = f_i(D_1, D_2)$$

From this we get that $L_1^{-1}\partial_i L_1 = (f_i(D_1, D_2))_-$ and $L_2^{-1}\partial_i L_2 = -(f_i(D_1, D_2))_+$. Thus, it follows from Lemma 2 that $\Phi_{H_1,H_2}(\Upsilon) \in \mathcal{F}_{H_1,H_2}$.

Finally, we prove that Φ_{H_1,H_2} is surjective. Suppose that $(D_1, D_2, \gamma) \in \mathcal{F}_{H_1,H_2}$. From (iii) of the Lemma 2 we have that there are L_k, $k = 1, 2$ such that $D_k = L_k^{-1} H_k L_k$ and

$$L_1^{-1}\partial_i L_1 = (f_i(D_1, D_2))_-, \qquad L_2^{-1}\partial_i L_2 = -(f_i(D_1, D_2))_+, \quad i \in I. \tag{34}$$

Let $\Upsilon = L_1 L_2^{-1}$, then using (34) we have that

$$\partial_i \Upsilon = L_1(L_1^{-1}\partial_i L_1 - L_2^{-1}\partial_i L_2)L_2^{-1} = L_1 f_i(D_1, D_2)L_2^{-1} = \widehat{f}_i(H_1, \Upsilon, H_2)$$

Hence $\Upsilon \in \mathcal{S}_{H_1,H_2}$ and $\Phi_{H_1,H_2}(\Upsilon) = (D_1, D_2, \gamma)$. Therefore Φ_{H_1,H_2} is surjective.

(iii) This was discussed in Subsection 2.3. We have also made use of the continued fraction described in the Remark 2 after Theorem 2. □

Remarks.

1. Let H_1' and H_2' be constant Hessenberg matrices such that $H_1' \sim_l H_1$ and $H_2' \sim_u H_2$. Since $\mathcal{F}_{H_1,H_2} = \mathcal{F}_{H_1',H_2'}$, it follows that $\Phi_{H_1',H_2'}$ is a bijection from $\mathcal{S}_{H_1',H_2'}$ to $\mathcal{F}_{H_1,H_2}$.

2. In the case when $N =$ [redacted] all Hessenberg matrices are conjugate to each other by a triangular matrix; thus $\mathcal{F} = \mathcal{F}_{H_1,H_2}$ for any H_1, H_2. When $N <$ [redacted], two Hessenberg matrices are conjugate by a triangular matrix if and only if they have the same characteristic polynomial.

3. This theorem reduces the problem of finding solutions of the nonlinear differential equations (30)-(32) to finding solutions of the linear equations with constant coefficients (33), and among those, the ones that decompose in the form $L_1 L_2^{-1}$.

Let us see how this works. We choose H_1 and H_2 constant Hessenberg matrices and find Υ that satisfies (33). Expanding $\binom{g_1(z)}{0}$ and $\binom{g_2(z)}{0}$ as defined in the theorem, in a continued fraction we obtain the entries of the matrices D_1 and D_2. In the case $N <$ [redacted] both continued fractions are finite and stop in the N-th step. If one of these continued fraction stops before the N-th step, or its coefficients do not satisfy that $\deg c_{k,0} = 1$ for all $k \leq N$, then by the discussion in Subsection 2.3 we know that $\mathcal{S}_{H_1,H_2} = \emptyset$. Thus, in this case we have $\mathcal{F}_{H_1,H_2} = \emptyset$.

3.3.1 The Cauchy Problem.

Let A be a Banach algebra. Let $R = C^\infty(\mathbb{R}^I, A)$ be the ring of A-valued functions in the variable $\mathbf{t} = (t_i)_{i \in I}$ with continuous partial derivatives of all orders. Let us take $\partial_i = \frac{\partial}{\partial t_i}$, $i \in I$. Then we can pose the Cauchy problem of finding (D_1, D_2, γ) that satisfy equations (30)-(32) with given initial conditions in $\mathbf{t} = 0$. In order to find (D_1, D_2, γ) we take $H_1 = D_1(0)$, $H_2 = \gamma(0)D_2(0)\gamma(0)^{-1}$ and find Υ that satisfies (33) with initial condition $\Upsilon = \gamma(0)^{-1} I_N$. This solution is taken by the map Φ_{H_1,H_2} to the solution of the given Cauchy problem. If such Υ

problem for (D_1, D_2, γ) has no solution. Otherwise, we can use the continued fraction algorithm to find (D_1, D_2, γ) in terms of the entries of Υ and the initial conditions.

Next we discuss the solubility of the equation for Υ. For this we will need a notion of "bounded initial condition" of the Cauchy problem. Let $l_2(\mathbb{N}, A)$ be the Banach space of sequences of elements of A with square summable norms. Let $B(l_2(\mathbb{N}, A))$ be the Banach algebra of bounded A-module endomorphisms of $l_2(\mathbb{N}, A)$. To every element $M \in B(l_2(\mathbb{N}, A))$ we associate the matrix $(M)_{i,j} = (Me_i)_j$. In the sequel we will say that an infinite matrix with entries in A is bounded if it comes from an element of $B(l_2(\mathbb{N}, A))$ in this way. We need bounded matrices in order to apply the holomorphic calculus of Banach algebras to the initial conditions of the Cauchy problem.

Suppose that $I = \{1, 2, \ldots, m\}$, $m \in \mathbb{N}$ and that $D_1(0)$, $D_2(0)$ are bounded operators (if $N <$ [redacted] then $D_1(0)$ and $D_2(0)$ are always bounded). Let U_1 and U_2 be open sets that contain the spectrum of $D_1(0)$ and $D_2(0)$ and such that ∂U_2 and ∂U_2 consists of a finite number of rectifiable Jordan curves oriented in the positive sense. Let $f(u, v, \mathbf{t}) = \sum_{0 \le n,m \le k} \alpha_{n,m}(\mathbf{t}) u^n v^m$ with $f(u, v, 0) = 0$. If we take $f_i = \frac{\partial f}{\partial t_i}$, then $\gamma(\mathbf{t})$ is invertible and the solution of the equation (33) with initial condition $\Upsilon(0) = \gamma(0)^{-1} I_N$ is given by

$$\Upsilon(\mathbf{t}) = -\frac{1}{4\pi^2} \int_{\partial U_1} \int_{\partial U_2} e^{f(u,v,\mathbf{t})} R(u, D_1(0)) R(v, D_2(0)) \gamma(0)^{-1} du dv.$$

When $f(u, v, \mathbf{t}) = g(u, \mathbf{t}) + h(v, \mathbf{t})$ the last equation becomes

$$\Upsilon(\mathbf{t}) = e^{g(D_1(0),\mathbf{t})} e^{h(D_2(0),\mathbf{t})} \gamma(0)^{-1}.$$

If $\mathbf{t}$ consists of just one variable t, or if the Banach algebra A is commutative, the system of equations

$$\frac{\partial x}{\partial t_i} = x \frac{\partial f}{\partial t_i}, \quad i \in I$$

with initial condition $\gamma(0) = 1$ always has one invertible solution. In the case of one variable the solution is given by Araki's expansional $\mathrm{Exp}(\cdot)$ (see [3]). In the case when A is commutative the solution is e^f. In these two cases we can drop the function γ and its equation from the system (30)-(32). That is, the existence of D_1 and D_2 that satisfy (30) and (31) automatically implies the existence of an invertible γ that satisfies (32).

Another case when we can simplify the system (30)-(32) is when $N <$ [redacted], $\mathbf{t}$ consists of just one variable t and $f = f(u)$ only depends of the variable u. In this case, given D_1 that satisfies equation (30) and initial conditions $(D_1(0), D_2(0), \gamma(0))$ there is a unique D_2 and a unique γ such that (D_1, D_2, γ) is a solution of the system (30)-(32). In order to see this we notice that D_2 and γ can be expressed in terms of expansionals of $(f(D_1(t)))_+$ and $f(D_1(t))_{0,0}$. We have

$$D_2(t) = \mathrm{Exp}_l \left(\int_0^t ; -(f(D_1([redacted])))_+ d[redacted] \right) D_2(0) \mathrm{Exp}_r \left(\int_0^t ; -(f(D_1([redacted])))_+ d[redacted] \right).$$

where Exp_l and Exp_r are Araki's left and right expansionals (see [3] for details). Conversely, it is clear that from a solution of (30)-(32) we get a solution of (30).

3.4 Examples

In this subsection we discuss some important examples of the system of equations (30)-(32).

3.4.1 The nonabelian Toda lattice.

Let $R = C^1(\mathbb{R}, M_k(\mathbb{R}))$ be the ring of $k \times k$ matrix valued functions with continuous derivative and we let the family of derivations consist of $\partial = \frac{d}{dt}$. Let L be an $N \times N$ tridiagonal matrix of the form (26) with entries in R, and such that a_k is invertible for $k = 0, \ldots, N-1$. Let us take $D_1 = D_2 = L$ and $f(u,v) = \frac{u+v}{2}$ in equations (30)-(32). In this way we get the nonabelian Toda lattice $\frac{d}{dt}L = [L, (L)_-]$. It follows from the discussion in 3.3.1 that the entries of L can be obtained from the continued fraction of the vector $e^{L(0)t}\rho(z, L(0))e_0$. Notice that when $N <$ [redacted] we have that $\rho(z, L(0)) = (zI - L(0))^{-1}$, so we expand in continued fraction a vector of rational functions in the variable z.

3.4.2 The relativistic nonabelian Toda lattice.

Let $N <$ [redacted]. Let us consider the system

$$\frac{da_n}{dt} = a_n b_{n-1} - b_n a_n, \quad 1 \le n \le N \tag{35}$$

$$\frac{db_n}{dt} = b_n(b_{n-1} + a_n) - (a_{n+1} + b_{n+1})b_n, \quad 1 \le n \le N-1 \tag{36}$$

with $a_i, b_i \in C^1(\mathbb{R}, M_k(\mathbb{R}))$ and $b_0 = b_N = 0$. Let $\lambda_{k,m} = b_k b_{k-1} \ldots b_m$ and L_1 and L_2 be the $N \times N$ matrices

$$L_1 = \begin{pmatrix} a_1 + b_1 & 1 & & & & \\ (a_2+b_2)\lambda_{1,1} & a_2 + b_2 & 1 & & & \\ (a_3+b_3)\lambda_{2,1} & (a_3+b_3)\lambda_{2,2} & a_3+b_3 & 1 & & \\ \vdots & \vdots & & & \ddots & \\ & & & & & 1 \\ a_N\lambda_{N-1,1} & a_N\lambda_{N-1,2} & & & \ldots & a_N \end{pmatrix}, \tag{37}$$

$$L_2 = \begin{pmatrix} a_1 & 1 & & & & \\ \lambda_{1,1}a_1 & b_1 + a_2 & 1 & & & \\ \lambda_{2,1}a_1 & \lambda_{2,2}(b_1+a_2) & b_2+b_3 & 1 & & \\ \vdots & \vdots & & \ddots & & \\ & & & & & 1 \\ \lambda_{N-1,1}a_1 & \lambda_{N-1,2}(b_1+a_2) & & \ldots & & b_{N-1}+a_N \end{pmatrix}. \tag{38}$$

Then the system (35) can written in the following Lax pairs

$$\frac{dL_k}{dt} = [L_k, -(L_k)_-], \quad k = 1, 2.$$

This system is the nonperiodic Relativistic Toda lattice and was studied in [11] (see also [8] for the periodic case).

As before, we take $R = C^1(\mathbb{R}, M_k(\mathbb{R}))$ and $\partial = \frac{d}{dt}$ in order to apply Theorem 7. Let us take $D_1 = L_1$ and $f(u, v) = -u$ in the system (30)-(32). Since f only depends on u, by the discussion at the end of 3.3.1 this system is equivalent to the equation $\frac{dL_1}{dt} = [L_1, -(L_1)_-]$. Thus by Theorem 7 the entries of L_1 are obtained from the continued fraction of the vector $e^{-tL_1(0)^t} * (zI - L_1(0)^t)^{-1} e_0$ (computations in $M_k(\mathbb{R})^{op}$). In a similar way we find L_2. From L_1 and L_2 we can easily compute the solutions of the original system.

3.4.3 The full Kostant-Toda lattice.

Let $R = C^1(\mathbb{R})$ and $\partial = \frac{d}{dt}$. We take $f(u) = u$ in (30)-(32) and $D = D_1$. Since f only depends on u this system reduces to the equation $\frac{dD}{dt} = [D, (D)_-]$. This equation is called the Full Kostant-Toda lattice and it was studied in [15] in connection with the representation of semisimple Lie groups. In particular, when D is a tridiagonal matrix the Full Kostant-Toda is just the Toda lattice.

3.4.4 Finite and Semi-infinite 2-Toda lattice.

Case 1: Ring of formal series. Let $I = \{(i, j) : 0 \le i, j \le$ [redacted]$, (i, j) \neq (0, 0)\}$ and $R = \mathbb{R}[[\mathbf{t}]]$, with $\mathbf{t} = \{t_{i,j}\}_{(i,j)\in I}$ the ring of formal series with real coefficients in the variables $t_{i,j}$. Taking $f_{i,j}(u, v) = (-1)^{\delta_{i,0}} u^i v^j$ and $\partial_{i,j} = \frac{\partial}{\partial t_{i,j}}$ in equations (30)-(32), we obtain the equations

$$\frac{\partial D_1}{\partial t_{i,j}} = [D_1, (D_1^i D_2^j)_-], \qquad \frac{\partial D_1}{\partial t_{0,j}} = [D_1, -(D_2^j)_-], \tag{39}$$

$$\frac{\partial D_2}{\partial t_{i,j}} = [D_2, -(D_1^i D_2^j)_+], \qquad \frac{\partial D_2}{\partial t_{0,j}} = [D_2, (D_2^j)_+] \tag{40}$$

with $(i, j) \in I$ and $i \neq 0$. This system is called the 2-Toda lattice and was studied in [2]. Let us consider the Cauchy Problem for these equations with initial value $D_1(0)$ and $D_2(0)$. The solution of the equations

$$\frac{\partial \Upsilon}{\partial t_{i,j}} = -D_1(0)^i \Upsilon D_2(0)^j, \qquad \frac{\partial \Upsilon}{\partial t_{0,j}} = \Upsilon D_2(0)^j, \tag{41}$$

with initial value $\Upsilon(0) = I$ is given by $\Upsilon(\mathbf{t}) = \sum_{i=0}^{\infty} G_{i,j}(\mathbf{t}) D_1(0)^i D_2(0)^j$. Here $G_i(\mathbf{t})$ are the polynomials defined by

$$\exp\left(\sum_{(i,j)\in I} (-1)^{\delta_{i,0}} t_{i,j} y^i z^j \right) = \sum_{(i,j)\in I} G_{i,j}(\mathbf{t}) y^i z^j.$$

By Theorem 7 we have that the solution of the 2-Toda lattice with initial values $D_1(0)$ and $D_2(0)$ is obtained from the continued fraction expansion of the vectors $\Upsilon(\mathbf{t})\rho(z, D_1(0))e_0$ and $\Upsilon(\mathbf{t})^t\rho(z, D_2(0)^t)e_0$.

Case 2: Ring of C^1 functions. Let $I = \{(i,j) : 0 \leq i,j \leq n-1, (i,j) \neq (0,0)\}$ and $R = C^1(\mathbb{R}^{n^2-1})$ be the ring of real functions in the variable $t = \{t_{i,j}\}_{(i,j)\in I}$ with continuous partial derivatives. Let us consider the Cauchy Problem for the equations (39) and (40) with bounded initial conditions $D_1(0)$ and $D_2(0)$. By the discussion in 3.3.1 the solution of the equations (41) with initial value $\Upsilon(0) = I$ is given by

$$\Upsilon(\mathbf{t}) = -\frac{1}{4\pi^2}\int_{\partial U_1}\int_{\partial U_2} e^{\sum_{(i,j)\in I}(-1)^{\delta_{i,0}} t_{i,j}u^i v^j} R(u, D_1(0))R(v, D_2(0))dudv.$$

Hence the solution of (39)-(40), with bounded initial conditions $D_1(0)$ and $D_2(0)$ is obtained by the continued fraction expansion of the vectors $\Upsilon(\mathbf{t})\rho(z, D_2(0))e_0$ and $\Upsilon(\mathbf{t})^t\rho(z, D_1(0)^t)e_0$.

3.4.5 The Bogoyavlensky lattice.

For this example we can not apply Theorem 7 directly the way it was stated before. Instead we use a variant of it assuming D_1 to be a lower m-Hessenberg matrix and D_2 an upper k-Hessenberg matrix. This analogous theorem again holds, and it can be proved adapting the proof of Theorem 7 to this case. Nevertheless, we will not state here this analog of Theorem 7 in its full generality. Instead, in Theorem 8 below we state the variant that will be needed later for the solution of the Bogoyavlensky lattice.

We say that a matrix D is a lower m-Hessenberg matrix if $(D)_{i,j} = 0$ for $j + m < i$, and $(D_1)_{j+m,j}$ is nonzero $\forall j \geq 0$. A matrix D is an upper k-Hessenberg matrix if it is the transpose of a lower k-Hessenberg matrix. Every m-Hessenberg matrix is a 1-Hessenberg matrix of $m \times m$ blocks.

Let $N \leq$ [redacted] such that N is multiple of two natural numbers k and m. Let D_1 and D_2 be $N \times N$ matrices with entries in $C^1(\mathbb{R})$ such that D_1 is a lower m-Hessenberg matrix with ones in the upper diagonal and D_2 is an upper k-Hessenberg matrix. Let $f(u,v,t)$ a polynomial in the variables u and v with coefficients in $C^1(\mathbb{R})$. We now state the version of Theorem 7 that we need in the Bogoyavlensky lattice.

Theorem 8. *Let us consider the system of equations*

$$\frac{dD_1}{dt} = [D_1, (f(D_1, D_2, t))_-], \tag{42}$$

$$\frac{dD_1}{dt} = [D_2, -(f(D_1, D_2, t))_+], \tag{43}$$

with bounded initial conditions $D_1(0)$ and $D_2(0)$ (i.e., $D_1(0), D_2(0) \in B(l_2(\mathbb{N}))$). Let Υ be the matrix defined by

$$\Upsilon(t) = -\frac{1}{4\pi^2}\int_{\partial U_1}\int_{\partial U_2} e^{\int_0^t f(u,v,s)ds} R(u, D_1(0))R(v, D_2(0))dudv.$$

where ∂U_1 and ∂U_2 are simple Jordan curves enclosing the spectra of $D_1(0)$ and $D_2(0)$ respectively. Then the solution of the equations (42) and (43) are obtained in the following way. The entries of D_2 as a 1-Hessenberg matrix over $M_k(\mathbb{R})$ are obtained by the expansion in continued fraction of $g_2(z) = \Upsilon \rho(z, D_2(0))\Pi_k$ over the ring $M_k(\mathbb{R})$, where $\Pi_k = (e_0, e_1, \ldots, e_{k-1})$. The entries of D_2 are computed according to formula (19). In the same way we obtain the entries of D_1 by the expansion in continued fraction of $g_1(z) = \Upsilon^t * \rho(z, D_1(0)^t)\Pi_m$ over the ring $M_m(\mathbb{R})^{op}$.

Let $N \leq$ [redacted] such that N is multiple of k and m. Let D be an $N \times N$ band matrix with entries in $C^1(\mathbb{R})$ such that D is an upper k-Hessenberg matrix and a lower m-Hessenberg matrix with ones in its upper diagonal. We consider for D the equation

$$\frac{dD}{dt} = [D, (D^{m+k})_-]. \tag{44}$$

If we suppose that D has only two diagonals different from zero, then this equation in the case $N =$ [redacted] is a discretization of the KdV equation and was studied by Bogoyavlensky in [5] and [6] (see also [21]).

D is a solution of (44) if and only if $D_1 = D$, $D_2 = D$ is a solution of (42)-(43) with $f(u, v) = u^m v$

[7] M. Bruschi, S. V. Manakov, O. Ragnisco, D. Levi (1980): *The nonabelian Toda lattice-discrete analogue of the matrix Schrdinger spectral problem.* J. Math. Phys. 21 , no. 12, 2749–2753.

[8] M. Bruschi, O. Ragnisco (1989): *The periodic relativistic Toda lattice: direct and inverse problem.* Inverse Problems 5, no. 3, 389–405.

[9] A. Bultheel, M. Van Barel (1995): *General block-biorthogonal polynomials. In A. Draux and V. Kaliaguine, editors, Orthogonal polynomials: the non-definite case.* Proceedings workshop, Rouen, France, April 24-26, 1995, Actes de l'atelier de Rouen, pages 3–24. INSA de Rouen, Lab. de Mathématique, 1997.

[10] P. M. Cohn (1995): Skew fields. Theory of general division rings. Encyclopedia of Mathematics and its applications, 57. Cambridge University Press, Cambridge.

[11] J. Coussement, A. B. J. Kuijlaars, W. Van Assche (2002): *Direct and inverse spectral transform for the relativistic Toda lattice and the connection with Laurent orthogonal polynomials.* Inverse Problems 18, no. 3, 923–942.

[12] I. Gelfand, D. Krob, A. Lascoux, B. Leclerc, V. S. Retakh, J.-Y. Thibon (1995): *Noncommutative symmetric functions* Adv. Math. 112, no. 2, 218–348.

[13] I. Gelfand, S. Gelfand, V. Retakh, R. Wilson (2004): *Quasideterminants*, Advances in Math. (in press)

[14] V. Kaliaguine (1995): *The operator moment problem, vector continued fractions and an explicit form of the Favard theorem for vector orthogonal polynomials.* Proceedings of the International Conference on Orthogonality, Moment Problems and Continued Fractions (Delft, 1994). J. Comput. Appl. Math. 65, no. 1-3, 181–193.

[15] B. Kostant (1979): *The solution to a generalized Toda lattice and representation theory.* Adv. in Math. 34 , no. 3, 195–338.

[16] J. Moser (1975): *Finitely many mass points on the line under the influence of an exponential potential–an integrable system.* Dynamical systems, theory and applications, pp. 467–497. Lecture Notes in Phys., Vol. 38, Springer, Berlin.

[17] E. M. Nikishin, V. N. Sorokin (1991): Rational approximations and orthogonality. Translated from the Russian by Ralph P. Boas. Translations of Mathematical Monographs, 92. American Mathematical Society, Providence, RI.

[18] V. I. Parusnikov (1984): *Weakly complete systems of functions and multidimensional continued fractions.*(Russian) Vestnik Moskov. Univ. Ser. I Mat. Mekh., no. 2, 13–17.

[19] L. Robert, L. Santiago (2003): *Finite sections method for Hessenberg matrices.* J. Approx. Theory 123, no. 1, 68–88.

[20] L. Robert, L. Santiago (2000): *General orthogonal polynomials.* University of Havana, Preprint.

[21] V. Sorokin, J. Van Iseghem (2000): *Matrix Hermite-Padé problem and dynamical systems.* Numerical analysis 2000, Vol. II: Interpolation and extrapolation. J. Comput. Appl. Math. 122, no. 1-2, 275–295.

[22] H. S. Wall (1948): Analytic Theory of Continued Fractions. Van Nostrand, New York.