SEMI-TENSOR PRODUCT OF MATRICES AND ITS SOME APPLICATIONS TO PHYSICS [*]

DAIZHAN CHENG[†] AND YALI DONG[†]

Abstract. In this paper we first give a general definition of a new kind of matrix products, called the semi-tensor product, which was firstly proposed in [4]. Certain new properties related to the later applications are proved. Using them, some problems in physics are investigated. First of all, the Carleman linearization of some dynamic physical systems is considered. It is used to investigate the invariants. A rigorous proof for the solvability is presented. Secondly, the problems of invariants of planar polynomial systems is converted to the solvability of a set of algebraic equations. Thirdly, we consider the contraction of a tensor field. A simple proof for general contraction is obtained.

1. Introduction. A new matrix product, called the semi-tensor product was introduced in [4]. Some of its applications have also been revealed there. In this paper, we first generalize the definition of semi-tensor product to matrices with arbitrary dimensions. Then certain new properties are proved, which are necessary for further investigation.

As an auxiliary tool, some properties of the commutation matrix are also investigated.

Some physical problems are discussed in the paper by using semi-tensor products. First, we consider the Carleman linearization of a nonlinear system, particularly, a polynomial system. We give a rigorous proof for the necessity of the equations for the polynomial type of first integrals. Moreover, for planar case, a more general type of first integrals is considered. A set of algebraic equations are presented for the existence of more general type of integrals.

Another problem considered is the contraction of tensor fields. A general proof for the formula of contraction of tensor fields is presented. It is used a lot in general relativity, but we did not see a proof for general case.

It is obvious that the new matrix product is useful in many other physical problems. The main purpose of this paper is to introduce this new matrix product and to show some of its applications.

The rest of the paper is organized as the follows: In section 2 the general definition for left and right semi-tensor products are given. Several examples are presented in section 3. The commutation matrix and its properties are discussed in section 4. Section 5 gives a tensor form expression for polynomials. Some properties, which are required in later discussion are presented in section 6. Section 7 gives certain formulas and properties of Carleman Linearization. The invariants of a planar polynomial systems is discussed in section 8. Section 9 devotes to the general property of the contraction of tensor fields. Section 10 is the conclusion.

2. Semi-tensor Product of matrices. Given two matrices $A \in M_{m \times n}$ and $B \in M_{p \times q}$, where $M_{s \times t}$ is the set of $s \times t$ matrices. In this section we consider the left and the right semi-tensor products of A and B. We need the following notations: Let $a, b \in Z^+$, where Z^+ is the set of positive integers. We denote by $a \wedge b$ the largest common divisor of a and b, and $a \vee b$ the least common multiplier. For instance, $6 \wedge 8 = 2$ and $6 \vee 8 = 24$.

DEFINITION 2.1.   *Let $A \in M_{m \times n}$, $B \in M_{p \times q}$ and $\alpha = n \vee p$. Then the left semi-tensor product of A and B, denoted by $A \ltimes B$, is defined as*

$$A \ltimes B = (A \otimes I_{\frac{\alpha}{n}})(B \otimes I_{\frac{\alpha}{p}}). \tag{2.1}$$

The right semi-tensor product of A and B, denoted by $A \rtimes B$, is defined as

$$A \rtimes B = (I_{\frac{\alpha}{n}} \otimes A)(I_{\frac{\alpha}{p}} \otimes B). \tag{2.2}$$

Where $\otimes$ is the Kronecker product of matrices.

Note that if $n = p$ these two products are degenerated to the conventional matrix product. So both of them are generalizations of the conventional matrix product.

The following properties are fundamental. Using definition, they can be proved by straightforward but tedious computations. So the proofs are omitted.

PROPOSITION 2.2. *The semi-tensor products satisfy*
1. Distributivity

$$\begin{aligned}
(A + B) \ltimes C &= (A \ltimes C) + (B \ltimes C), \\
(A + B) \rtimes C &= (A \rtimes C) + (B \rtimes C), \\
C \ltimes (A + B) &= (C \ltimes A) + (C \ltimes B), \\
C \rtimes (A + B) &= (C \rtimes A) + (C \rtimes B).
\end{aligned} \tag{2.3}$$

2. Associativity

$$\begin{aligned}
(A \ltimes B) \ltimes C &= A \ltimes (B \ltimes C), \\
(A \rtimes B) \rtimes C &= A \rtimes (B \rtimes C).
\end{aligned} \tag{2.4}$$

Many properties of the conventional matrix product remain true for this (left or right) semi-tensor product. For instance, we give the following:

PROPOSITION 2.3. *1. $(A \ltimes B)^T = B^T \ltimes A^T$; $(A \rtimes B)^T = B^T \rtimes A^T$.*
2. If $M \in M_{m \times pn}$, then $M \ltimes I_n = M$ and $M \rtimes I_n = M$; If $M \in M_{pm \times n}$, then $I_m \ltimes M = M$ and $I_m \rtimes M = M$.
  In the following let A, B be two square matrices.
3. $A \ltimes B$ and $B \ltimes A$ ($A \rtimes B$ and $B \rtimes A$) have the same characteristic function.
4. $\operatorname{tr}(A \ltimes B) = \operatorname{tr}(B \ltimes A)$; $\operatorname{tr}(A \rtimes B) = \operatorname{tr}(B \rtimes A)$.
5. If both A and B are orthogonal, (or upper triangular, or diagonal), then so is $A \ltimes B$ or $A \rtimes B$.
6. If either A or B is invertible, then $A \ltimes B \sim B \ltimes A$ and $A \rtimes B \sim B \rtimes A$, where $\sim$ means two matrices are similar.
7. If both A and B are invertible, then

$$(A \ltimes B)^{-1} = B^{-1} \ltimes A^{-1}; \quad (A \rtimes B)^{-1} = B^{-1} \rtimes A^{-1}.$$

8. The determinants of the products are

$$\det(A \ltimes B) = (\det(A))^{\frac{\alpha}{n}}(\det(B))^{\frac{\alpha}{p}}; \quad \det(A \rtimes B) = (\det(A))^{\frac{\alpha}{n}}(\det(B))^{\frac{\alpha}{p}},$$

where α, n, and p are as in Definition 2.1.

PROPOSITION 2.4.   *1. Let $A \in M_{m \times n}$ and $B \in M_{p \times q}$. Then $C = A \ltimes B = (C^{ij})$,* $i = 1, \cdots, m$, $j = 1, \cdots, q$, *with*

$$C^{ij} = A^i \ltimes B_j. \tag{2.5}$$

where A^i is i-th row of A and B_j is the j-th column of B.
    2. Let A and B be split into blocks as $A = (A^{ij})$ and $B = (B_{ij})$. Assume

$$\frac{\dim row(A^{ik})}{\dim column(B_{kj})} = \frac{n}{p}, \quad \forall i, j, k.$$

Then $C = (C^{ij})$, with

$$C^{ij} = \sum_k A^{ik} \ltimes B_{kj}. \tag{2.6}$$

It is worth to emphasize that the right semi-tensor product does not have similar properties as in Proposition 2.4. So the left semi-tensor product is more convenient in use. In the sequel, we consider the left semi-tensor product as a conventional extension of conventional matrix product.
    Let $A \in M_{m \times n}$ and $m \wedge n = t$. Set $m = m_0 t$ and $n = n_0 t$, then m_0 and n_0 are co-prime. We define a power of A as

$$\begin{cases} A^1 = A, \\ A^{k+1} = A \ltimes A^k, \quad k \geq 1. \end{cases}$$

Similarly,

$$\begin{cases} A^{\rtimes 1} = A, \\ A^{\rtimes(k+1)} = A \rtimes A^{\rtimes k}, \quad k \geq 1. \end{cases}$$

As for the dimension of A^k or $A^{\rtimes k}$, it is easy to prove by mathematical induction that $A^k \in M_{m_0^k t \times n_0^k t}$ and $A^{\rtimes k} \in M_{m_0^k t \times n_0^k t}$.
    Let $A \in M_{m \times n}$ and $B \in M_{p \times q}$. In practical problems the most useful case for the semi-tensor product is either n is a factor of p, i.e., $nt = p$ for some integer t, or p is a factor of n, i.e., $n = pt$ for some integer t. For the first case we denote

$$A \prec_t B. \tag{2.7}$$

For the second case we denote

$$A \succ_t B. \tag{2.8}$$

In the rest of this paper we assume, without any exception, that
A1. Either (2.7) or (2.8) holds.

As a convention, sometimes we simply use $AB = A \ltimes B$. There is no confusion because when the dimensions of A and B are suitable for conventional matrix product the left semi-tensor product coincides with the conventional matrix product.

3. Some Examples. In this section we give some simple examples to show the new matrix products.

EXAMPLE 3.1. Let $X \in \mathbb{R}^m$ and $Y \in \mathbb{R}^n$. Then

$$X \ltimes Y = (X \otimes I_n)Y = (x_1 y_1 \cdots x_1 y_n \ \cdots \ x_m y_1 \cdots x_m y_n)^T = X \otimes Y;$$

$$X \rtimes Y = (I_n \otimes X)Y = (x_1 y_1 \cdots x_m y_1 \ \cdots \ x_1 y_m \cdots x_m y_n)^T = Y \otimes X;$$

$\square$

EXAMPLE 3.2. Let

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ \vdots \\ b_{61} & b_{62} \end{pmatrix}.$$

Then $\alpha = 4 \vee 6 = 12$, $\beta = 4 \wedge 6 = 2$, and

$A \ltimes B = (A \otimes I_3)(B \otimes I_2)$
$$= \begin{pmatrix} a_{11}b_{11} + a_{13}b_{41} & a_{12}b_{21} + a_{14}b_{51} & a_{11}b_{12} + a_{13}b_{42} & a_{12}b_{22} + a_{14}b_{52} \\ a_{12}b_{31} + a_{14}b_{61} & a_{11}b_{11} + a_{13}b_{41} & a_{12}b_{32} + a_{14}b_{62} & a_{11}b_{12} + a_{13}b_{42} \\ a_{11}b_{21} + a_{13}b_{51} & a_{12}b_{31} + a_{14}b_{61} & a_{11}b_{22} + a_{13}b_{52} & a_{12}b_{32} + a_{14}b_{62} \\ a_{21}b_{11} + a_{23}b_{41} & a_{22}b_{21} + a_{24}b_{51} & a_{21}b_{12} + a_{23}b_{42} & a_{22}b_{22} + a_{24}b_{52} \\ a_{22}b_{31} + a_{24}b_{61} & a_{21}b_{11} + a_{23}b_{41} & a_{22}b_{32} + a_{24}b_{62} & a_{21}b_{12} + a_{23}b_{42} \\ a_{21}b_{21} + a_{23}b_{51} & a_{22}b_{31} + a_{24}b_{61} & a_{21}b_{22} + a_{23}b_{52} & a_{22}b_{32} + a_{24}b_{62} \end{pmatrix}$$

$A \rtimes B = (I_3 \otimes A)(I_2 \otimes B)$
$$= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} + a_{14}b_{41} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} + a_{14}b_{42} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} + a_{24}b_{41} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} + a_{24}b_{42} \\ a_{11}b_{51} + a_{12}b_{61} & a_{11}b_{52} + a_{12}b_{62} \\ a_{21}b_{51} + a_{22}b_{61} & a_{21}b_{52} + a_{22}b_{62} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ a_{13}b_{11} + a_{14}b_{21} & a_{13}b_{12} + a_{14}b_{22} \\ a_{23}b_{11} + a_{24}b_{21} & a_{23}b_{12} + a_{24}b_{22} \\ a_{11}b_{31} + a_{12}b_{41} + a_{13}b_{51} + a_{14}b_{61} & a_{11}b_{32} + a_{12}b_{42} + a_{13}b_{52} + a_{14}b_{62} \\ a_{21}b_{31} + a_{22}b_{41} + a_{23}b_{51} + a_{24}b_{61} & a_{21}b_{32} + a_{22}b_{42} + a_{23}b_{52} + a_{24}b_{62} \end{pmatrix}$$

$\square$

The following example shows how the multi-fold of cross-product in $\mathbb{R}^3$ be performed semi-tensor product.

EXAMPLE 3.3. Consider the cross-product in R^3. Let $e_1 = i$, $e_2 = j$ and $e_3 = k$, and

$$e_i \times e_j = \sum_{k=1}^{3} c_{ij}^k e_k, \quad i, j = 1, 2, 3.$$

Construct a matrix, called the structure matrix as

$$M = \begin{pmatrix} c_{11}^1 & c_{12}^1 & c_{13}^1 & c_{21}^1 & c_{22}^1 & c_{23}^1 & c_{31}^1 & c_{32}^1 & c_{33}^1 \\ c_{11}^2 & c_{12}^2 & c_{13}^2 & c_{21}^2 & c_{22}^2 & c_{23}^2 & c_{31}^2 & c_{32}^2 & c_{33}^2 \\ c_{11}^3 & c_{12}^3 & c_{13}^3 & c_{21}^3 & c_{22}^3 & c_{23}^3 & c_{31}^3 & c_{32}^3 & c_{33}^3 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Consider $X_s = (a_s, b_s, c_s)^T \in R^3$, $s = 1, \cdots, n$. Then it is easy to see that

$$X_1 \times X_2 = M \ltimes X_1 \ltimes X_2 = M X_1 X_2.$$

Since the semi-tensor product is associative, unlike the cross product we don't need to worry about the order of the product. In general

$$(\cdots (X_1 \times X_2) \times \cdots) \times X_n = M^n X_1 X_2 \cdots X_n.$$

$\square$

The last example gives an additional reason for introducing the left semi-tensor product.

EXAMPLE 3.4. Let $X, Y, Z, W \in \mathbb{R}^n$. Then

$$(XY^T)(ZW^T) \in M_{n \times n}.$$

Let's do the following transformation (within conventional matrix product) :

$$(XY^T)(ZW^T) = X(Y^T Z)W^T = (Y^T Z)XW^T = Y^T(ZX)W^T. \tag{3.1}$$

The above transformation seems legal because $Y^T Z$ is a scaler and the conventional matrix product is associative. But finally we meet ZX, which is not defined. Now if we generalize the conventional matrix product to left semi-tensor product, the puzzle is solved completely. Not only ZX is well defined but also the equality (3.1) holds perfectly. $\square$

4. Commutation Matrix. The commutation matrix was introduced in [8]. We give a constructive definition:

DEFINITION 4.1. *An $mn \times mn$ matrix $W_{[m,n]}$ is called a commutation matrix if we label its columns by $(11, 12, \cdots, 1n, \cdots, m1, m2, \cdots, mn)$ and its rows by $(11, 21, \cdots, m1, \cdots, 1n, 2n, \cdots, mn)$, and set its entries in the (I, J)-th row and (i, j)-th column as*

$$w_{(IJ),(ij)} = \delta_{i,j}^{I,J} = \begin{cases} 1, & I = i \text{ and } J = j, \\ 0, & \text{otherwise.} \end{cases} \tag{4.1}$$

When $m = n$ we denote $W_{[n,n]} = W_{[n]}$.

EXAMPLE 4.2. Let $m = 2$ and $n = 3$, $W_{[2,3]}$ is expressed as

$$W_{[2,3]} = \begin{array}{c} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array} \begin{array}{l} (11) \\ (21) \\ (12) \\ (22) \\ (13) \\ (23) \end{array}.$$

with column labels $(11)\ (12)\ (13)\ (21)\ (22)\ (23)$

While $W_{[3,2]}$ is

$$
W_{[3,2]} \quad = \quad
\begin{array}{c}
\begin{array}{cccccc}
(11) & (12) & (21) & (22) & (31) & (32)
\end{array} \\
\left(\begin{array}{cccccc}
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{array}\right)
\begin{array}{c}
(11) \\ (21) \\ (31) \\ (12) \\ (22) \\ (32)
\end{array}
\end{array}
\quad .
$$

$\Box$

We refer to [7] and [9] for the following properties.

PROPOSITION 4.3.   *1. The inverse and the transpose of a commutation matrix are*

$$W_{[n,m]} = W_{[m,n]}^{T} = W_{[m,n]}^{-1}. \tag{4.2}$$

2. When $n = m$ (4.2) becomes

$$W_{[n]} = W_{[n]}^{T} = W_{[n]}^{-1}. \tag{4.3}$$

Given a matrix $A = (a_{ij}) \in M_{m \times n}$. We use $V_c(A)$ and $V_r(A)$ for its column stacking form and row stacking form respectively. That is,

$$V_c(A) = (a_{11} \cdots a_{m1}, \cdots, a_{1n} \cdots a_{mn})^{T},$$

$$V_r(A) = (a_{11} \cdots a_{1n}, \cdots, a_{m1} \cdots a_{mn})^{T}.$$

The commutation matrix can realize the swap between row and column stacking forms of a matrix:

PROPOSITION 4.4.   *Let $A \in M_{m \times n}$. Then*

$$\begin{cases} W_{[m,n]} V_r(A) = V_c(A), \\ W_{[n,m]} V_c(A) = V_r(A). \end{cases} \tag{4.4}$$

The following factorization property is very useful for simplifying the product of swap matrices etc.

PROPOSITION 4.5 [5].

$$W_{[p,qr]} = (I_q \otimes W_{[p,r]})(W_{[p,q]} \otimes I_r) = (I_r \otimes W_{[p,q]})(W_{[p,r]} \otimes I_q), \tag{4.5}$$

$$W_{[pq,r]} = (W_{[p,r]} \otimes I_q)(I_p \otimes W_{[q,r]}) = (W_{[q,r]} \otimes I_p)(I_q \otimes W_{[p,r]}). \tag{4.6}$$

5. Tensor Form of Polynomials. Denote by B_n^k the set of k-th homogeneous polynomials in $\mathbb{R}^n$. It is obvious that x^k is a basis. That is, let $p(x_1, \cdots, x_n) \in B_n^k$. Then there exists a row vector $F \in \mathbb{R}^{n^k}$ such that

$$p(x) = Fx^k \ (= F \ltimes \underbrace{x \ltimes \cdots \ltimes x}_{k}). \tag{5.1}$$

Of course x^k is a redundant basis. So the coefficient array F is not unique. F is a symmetric coefficient array, if for same monomials (such as $x_1^2 x_2$ and $x_2 x_1^2$), the corresponding coefficients are equal.

A natural basis, denoted by $x_{(k)}$, is

$$x_{(k)} = (x_1^k, x_1^{k-1}x_2, \cdots, x_{n-1}x_n^{k-1}, x_n^k)^T,$$

that is, arrange the set of k-th degree monic monomials in alphabetic order.

It can be proved by mathematical induction that the dimension of $x_{(k)}$ is

$$s := dim(x_{(k)}) = \frac{(n+k-1)!}{k!(n-1)!}; \quad k \geq 0, \quad n \geq 1.$$

Then there exist two matrices $T_N(n, k) \in M_{n^k \times s}$ and $T_B(n, k) \in M_{s \times n^k}$, such that

$$x^k = T_N(n,k)x_{(k)}, \quad x_{(k)} = T_B(n,k)x^k.$$

Moreover, it is an immediate consequence of the definition that

$$T_B(n,k)T_N(n,k) = I_s.$$

Now given a $p(x) \in B_n^k$, which is expressed as

$$p(x) = Fx^k = Gx_{(k)},$$

then $FT_N = G$.

But since x^k is redundant, GT_B is only the symmetric expression of F.

Next, we consider the differential of a matrix with differentiable function entries.

DEFINITION 5.1. *Let $H = (h_{ij}(x))$ be a $p \times q$ matrix with the entries $h_{ij}(x)$ as smooth functions of $x \in \mathbb{R}^n$. Then the differential of H is defined as a $p \times nq$ matrix obtained by replacing each element h_{ij} by its differential $dh_{ij} = (\frac{\partial h_{ij}(x)}{\partial x_1}, \cdots, \frac{\partial h_{ij}(x)}{\partial x_n})$.*

Our goal is to apply it to polynomials. We construct an $n^{k+1} \times n^{k+1}$ matrix Φ_k as

$$\Phi_k = \sum_{s=0}^{k} I_{n^s} \otimes W_{[n^{k-s}, n]}. \tag{5.2}$$

Then we have the following differential form of x^k, which is fundamental in the further approach.

PROPOSITION 5.2 [5].

$$D(x^{k+1}) = \Phi_k \ltimes x^k. \tag{5.3}$$

6. Some Properties. In this section we give a few properties of the semi-tensor product. Most of the proofs can be obtained by definitions and straightforward computations.

PROPOSITION 6.1.   *Assume $A \in M_{m \times n}$ is given.*
1. Let $Z \in R^t$ be a row vector. Then

$$A \ltimes Z = Z \ltimes W_{[m,t]} \ltimes A \ltimes W_{[t,n]} = Z \ltimes (I_t \otimes A). \qquad (6.1)$$

2. Let $Z \in R^t$ be a column vector. Then

$$Z \ltimes A = W_{[m,t]} \ltimes A \ltimes W_{[t,n]} \ltimes Z = (I_t \otimes A) \ltimes Z. \qquad (6.2)$$

3. Let $X \in R^t$ be a column, $Y \in R^s$ be a row. Then

$$XY = Y \ltimes W_{[t,s]} \ltimes X. \qquad (6.3)$$

PROPOSITION 6.2.   *Let $A \in M_{m \times n}$ and $B \in M_{p \times q}$. Then*

$$W_{[m,p]}(A \otimes B)W_{[q,n]} = (B \otimes A). \qquad (6.4)$$

Particularly,

$$(I_p \otimes A)W_{[n,p]} = W_{[m,p]}(A \otimes I_p), \qquad (6.5)$$

The following property of swap is very useful.

PROPOSITION 6.3.   *Let $X_i \in \mathbb{R}^{n_i}$, $i = 1, \cdots, m$. Then*

$$\left(I_{n_1} \otimes \cdots \otimes I_{n_{i-1}} \otimes W_{[n_i, n_{i+1}]} \otimes I_{n_{i+2}} \otimes \cdots \otimes I_{n_m}\right) X_1 \cdots X_{i-1} X_i X_{i+1} X_{i+2} \cdots X_m$$
$$= X_1 \cdots X_{i-1} X_{i+1} X_i X_{i+2} \cdots X_m. \qquad (6.6)$$

A natural question is: why we need semi-tensor product? Before ending this section, we should like to answer this question. It can be seen easily that a semi-tensor product can be expressed by conventional matrix product with Kronecker product. But the point is: if an expression has both conventional and Kronecker products, the associativity doesn't exist. But since conventional product is a particular semi-tensor product, the associativity remains true between them. This advantage makes many manipulations of matrix products possible. In later discussion it can be seen from time to time. Without semi-tensor product, some formulas are just impossible (not only difficult) to be deduced.

7. Carleman Linearization. In this section the tensor product form will be used to analyze the Carleman linearization. Using the linearization form, a kind of first integrals are investigated and a general formula is obtained.

Consider a dynamical system

$$\dot{x} = f(x), \quad x \in \mathbb{R}^n, \qquad (7.1)$$

where $f(x)$ is an analytic vector field with $f(0) = 0$.

J. Carleman has proposed a method to embed it into an infinite dimensional linear system. We refer to [11], [1] and the references there in for details and later development.

Choosing $x, x^2, \cdots$ as a basis, the system (7.1) can be expressed as

$$\dot{x} = F_1 x + F_2 x^2 + F_3 x^3 + \cdots, \tag{7.2}$$

where F_1 is an $n \times n$ matrix and F_2 is of $n \times n^2$ and so on.

We may consider x^2, x^3 etc. as a set of independent arguments and calculate their derivatives to get the Carleman linearization form as

$$\begin{pmatrix} \dot{x} \\ \dot{x^2} \\ \dot{x^3} \\ \vdots \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} & \cdots \\ 0 & A_{22} & A_{23} & A_{24} & \cdots \\ 0 & 0 & A_{33} & A_{34} & \cdots \\ \vdots & & & & \end{pmatrix} \begin{pmatrix} x \\ x^2 \\ x^3 \\ \vdots \end{pmatrix}. \tag{7.3}$$

THEOREM 7.1. *In Carleman linearization form (7.3) the coefficients A_{ij} are determined by the following equations.*

$$\begin{cases} A_{1i} = F_i, \quad i \geq 1, \\ A_{k,k+s} = \sum\limits_{i=0}^{k-1} I_{n^i} \otimes F_{s+1} \otimes I_{n^{k-1-i}}. \end{cases} \tag{7.4}$$

Proof. According to chain rule, we have

$$\frac{d}{dt}(x^k) = \sum_{i=0}^{k-1} x^i \dot{x} x^{k-i-1} = \sum_{s=0}^{\infty} \sum_{i=0}^{k-1} x^i F_{s+1} x^{k-i+s}.$$

Using equation (6.2), we have

$$x^i F_{s+1} x^{k-i+s} = (I_{n^i} \otimes F_{s+1}) \ltimes x^{k+s} = (I_{n^i} \otimes F_{s+1} \otimes I_{n^{k-i-1}}) x^{k+s}.$$

The equation (7.4) follows. □

We can formally express (7.3) in a linear form as

$$\dot{X} = AX, \tag{7.5}$$

where A is an infinite dimensional block upper triangular matrix.

An infinite dimensional block upper triangular matrix has some special properties, which make the expression (7.5) meaningful. We give a brief discussion here:

Denote the upper left (leading) k blocks minor of an infinite dimensional block upper triangular matrix A by A_k. That is

$$A_k = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ 0 & A_{22} & \cdots & A_{2k} \\ \vdots & & & \\ 0 & 0 & \cdots & A_{kk} \end{pmatrix}.$$

We say a block upper triangular matrix with square diagonal blocks to have a set of structure constants $(k_1, k_2, \cdots)$ if the dimensions of the diagonal blocks are $\dim(A_{ii}) = k_i \times k_i$. For instance, for equation (7.3) the structure constants are $(n, n^2, n^3, \cdots)$. For statement ease, we identify A_k with it extension: an infinite dimensional block upper triangular matrix, A_k^e with A_k as its upper-left minor and zero for all other elements. Using this convention, the coefficient matrix, A, can be considered as

$$A = \lim_{k \to \infty} A_k.$$

This limit is well defined because if we denote the (ij)-th element of A_k as a_{ij}^k, then the sequence of $\{a_{ij}^k, \ k = 1, 2, \cdots\}$ has the form as

$$(a_{ij}^1, a_{ij}^2, \cdots, a_{ij}^k, \cdots) = (0, \cdots, 0, c_{ij}, c_{ij}, c_{ij}, \cdots),$$

i.e., it is a constant sequence except the first finite terms. Based on the same argument the following notations are well defined.

DEFINITION 7.2.   *1. Let A and B be two block upper triangular infinite dimensional matrices with same structure constants. Then we define the product of A and B as*

$$AB := \lim_{k \to \infty} A_k B_k.$$

2. Assume A_{ii}, $i = 1, 2, \cdots$ are invertible, then we define

$$A^{-1} := \lim_{k \to \infty} A_k^{-1}.$$

3.

$$e^A := \lim_{k \to \infty} e^{A_k}.$$

Now it is natural to use linear solution

$$X = e^{At} X_0$$

as the solution of the non-linear system (7.3). In fact, we can use only finite term to approximate the solution.

Denote the (i, j)-th block of $e^{A_k t}$ by $E_{ij}^k(t)$. Then it is easy to see that

$$E_{ij}^s(t) = E_{ij}^k(t), \quad s > k, \quad i, j \le k.$$

Hence we can define

$$X^n(t) = \sum_{k=1}^{n} E_{1k}^k(t) X_0^k.$$

From (7.3) it is clear that if

$$X(t) = \lim_{n \to \infty} X^n(t)$$

exists, then it is the solution of (7.1) with $X(0) = X_0$.

We are particularly interested in the upper triangular matrices obtained by Carleman linearization. In the Carleman linearization form (7.3) suppose $F_1 = A_{11}$ is stable (anti-stable), i.e., $\mathrm{Re}\sigma(A_{11}) < 0$ ($\mathrm{Re}\sigma(A_{11}) > 0$), then A is invertible. In fact, we have

THEOREM 7.3.   *Assume $F_1 = A_{11}$ has eigenvalues as $\sigma(A_{11}) = \{\lambda_1, \cdots, \lambda_n\}$, then A_{ii}, $i \geq 2$ have eigenvalues*

$$\sigma(A_{ii}) = \{\lambda_{k_1} + \cdots + \lambda_{k_i} \mid k_1, \cdots, k_i = 1, \cdots, n\}.$$

Proof. First, we assume the eigenvalues of A_{11} are distinct and their corresponding eigenvectors are

$$\{\xi_1, \cdots, \xi_n\}.$$

Then a straightforward computation shows that

$$A_{ii}(\xi_{k_1} \ltimes \cdots \ltimes \xi_{k_i}) = (\lambda_{k_1} + \cdots + \lambda_{k_i})(\xi_{k_1} \ltimes \cdots \ltimes \xi_{k_i}).$$

To avoid notational mess, we show it for only $i = 2$. By (7.4) we have

$$A_{22} = I_n \otimes A_{11} + A_{11} \otimes I_n.$$

Then

$$\begin{aligned}
A_{22}(\xi_i \ltimes \xi_j) &= (I_n \otimes A_{11} + A_{11} \otimes I_n)(\xi_i \ltimes \xi_j) \\
&= (I_n \otimes A_{11})(\xi_i \ltimes \xi_j) + (A_{11} \otimes I_n)(\xi_i \ltimes \xi_j) \\
&= \lambda_j \xi_i \ltimes \xi_j + \lambda_i \xi_i \ltimes \xi_j.
\end{aligned}$$

Since all λ_i are distinct, we can claim that all

$$\{\xi_i \ltimes \xi_j \mid i, j = 1, \cdots, n\}$$

are linearly independent. In fact, since λ_i are distinct, all ξ_i are linearly independent. Now a straightforward computation shows that

$$(\xi_1, \cdots, \xi_n) \otimes (\xi_1, \cdots, \xi_n) = (\xi_1\xi_1, \cdots, \xi_1\xi_n, \cdots, \xi_n\xi_1, \cdots, \xi_n\xi_n).$$

The claim follows.

Hence the eigenvalues of A_{22} are

$$\sigma(A_{22}) = \{\lambda_i + \lambda_j \mid i, j = 1, \cdots, n\}.$$

By continuity, the eigenvalue structure is also true even the multi-fold eigenvalues exist. $\square$

Next, we consider a polynomial system

$$\dot{x} = F_1 x + F_2 x^2 + \cdots + F_k x^k. \tag{7.6}$$

The Carleman linearization technique is used to investigating its first integrals of the form

$$H(t, x) = e^{-\xi t} P(x).$$

Particularly, this kind of first integrals of Lorenz systems were discussed in [11] , of Lotka-Volterra systems were discussed in [1].

Assume $P(x) = P_0 + P_1 x + \cdots + P_s x^s$, with symmetric coefficients $P_1, \cdots, P_s$. (Where "symmetric" means the coefficients for the same monomial terms with different factor orders are the same. Say, the coefficients for $x_1^2 x_2$, $x_1 x_2 x_1$ and $x_2 x_1^2$ are the same.) It is easy to see that if $\xi \neq 0$, then $P_0 = 0$. So we simply assume $P_0 = 0$.

Setting $dH(t,x)/dt = 0$, we get

$$
(P_1, \cdots, P_s)
\begin{pmatrix}
A_{11} & \cdots & \cdots & A_{1k} & & \\
 & \ddots & & & \ddots & \\
 & & A_{ss} & \cdots & \cdots & A_{s,s+k-1}
\end{pmatrix}
\begin{pmatrix}
x \\ x^2 \\ \vdots \\ x^{s+k-1}
\end{pmatrix}
$$
$$
= \xi (P_1, \cdots, P_s)
\begin{pmatrix}
x \\ x^2 \\ \vdots \\ x^s
\end{pmatrix}.
\tag{7.7}
$$

Since the basis, x^k, is redundant, the coefficients are not unique. Using this form to search first integral may be too conservative. In other words, the conditions obtained may not be necessary, because under another equivalent coefficients another kind of first integral may also be obtained.

To get necessary and sufficient condition we have to convert the system into the natural basis. Set

$$
P_i = \tilde{P}_i T_B(n, i), \quad \tilde{A}_{ij} = T_B(n, i) A_{ij} T_N(n, j).
$$

Putting them into (7.7), it turns out to be

$$
(\tilde{P}_1, \cdots, \tilde{P}_s)
\begin{pmatrix}
\tilde{A}_{11} & \cdots & \cdots & \tilde{A}_{1k} & & \\
 & \ddots & & & \ddots & \\
 & & \tilde{A}_{ss} & \cdots & \cdots & \tilde{A}_{s,s+k-1}
\end{pmatrix}
\begin{pmatrix}
x \\ x_{(2)} \\ \vdots \\ x_{(s+r-1)}
\end{pmatrix}
$$
$$
= \xi (\tilde{P}_1, \cdots, \tilde{P}_s)
\begin{pmatrix}
x \\ x_{(2)} \\ \vdots \\ x_{(s)}
\end{pmatrix}.
$$

THEOREM 7.4. *Denote* $h_i = \tilde{P}_i^T$, $B_{ij} = \tilde{A}_{ji}^T$. *Then system (7.6) has first integral* $H(t, x)$, *iff, there exists* ξ *such that the following equations have non-zero solution*

$(h_1, \cdots, h_s)$.

$$
\begin{cases}
\begin{pmatrix}
B_{11} & 0 & 0 & \cdots & 0 \\
B_{21} & B_{22} & 0 & \cdots & 0 \\
\vdots & & & & \\
B_{s1} & B_{s2} & & \cdots & B_{s,s}
\end{pmatrix}
\begin{pmatrix}
h_1 \\ h_2 \\ \vdots \\ h_s
\end{pmatrix}
= \xi
\begin{pmatrix}
h_1 \\ h_2 \\ \vdots \\ h_s
\end{pmatrix}, \\[2em]
\begin{pmatrix}
B_{s+1\,1} & B_{s+1\,2} & \cdots & B_{s+1\,s} \\
\vdots & & & \\
B_{k\,1} & B_{k\,2} & \cdots & B_{k\,s} \\
0 & B_{k+1\,2} & \cdots & B_{k+1\,s} \\
\vdots & & & \\
0 & \cdots & 0 & B_{s+k-1\,s}
\end{pmatrix}
\begin{pmatrix}
h_1 \\ h_2 \\ \vdots \\ h_s
\end{pmatrix}
= 0.
\end{cases}
\tag{7.8}
$$

Next, we consider the solution form of equation (7.8). We need some a new notation: Let $Id(I; n^k)$ be the set of k-th fold indexes. That is, $I = (i_1, \cdots, i_k)$ and let i_k goes from 1 to n first, then i_{k-1} and so on. e.g.,

$$Id(I; 2^3) = (111, 112, 121, 122, 211, 212, 221, 222).$$

A row of vector $h \in \mathbb{R}^{n^k}$ is said to be symmetric with respect to $Id(I; n^k)$, if

$$h_{i_1, \cdots, i_k} = h_{i_{\sigma(1)}, \cdots, i_{\sigma(k)}}, \quad \forall \sigma \in S_k.$$

Where the S_k is the permutation group. e.g.,

$$h = (2, 1, 1, 1, 3, 1, 3, 3, -2)$$

is symmetric with respect to $Id(I; 2^3)$. because

$$h_{112} = h_{121} = h_{211} = 1, \quad h_{122} = h_{212} = h_{221} = 3.$$

The following lemma itself is interesting.

LEMMA 7.5. *Assume the row vector $h \in R^{n^k}$ is indexed by $Id(I; n^k)$ and is symmetric with respect to $Id(I; n^k)$. $F \in M_{n \times n}$. Set*

$$A = F \otimes I_{n^{k-1}} + I_n \otimes F \otimes I_{n^{k-2}} + \cdots + I_{n^{k-1}} \otimes F.$$

Then hA is also symmetric with respect to $Id(I; n^k)$.

Proof. Interchanging any two indexes can be realized by interchanging adjacent indexes. Hence we have only to show that hA is invariant under the interchange of two adjacent indexes. Set

$$\Phi = I_{n^{j-1}} \otimes W_{[n]} \otimes I_{n^{k-j-1}}.$$

It is obvious that swapping the j-th and the $(j + 1)$-th indexes of h yields the new vector $h\Phi$. Since h is symmetric with respect to $Id(I; n^k)$, then $h\Phi = h$, and hence $h\Phi A = hA$. To see hA is symmetric, i.e., $hA\Phi = hA$, it suffices to show that

$$A\Phi = \Phi A.$$

Note that the terms in A have the following form:

$$I \otimes \cdots \otimes I \otimes F \otimes I \otimes \cdots \otimes I. \tag{7.9}$$

If F does not locate on j or $j+1$ position, then it is obvious that Φ commutes with equation (7.9). So we have only to consider the related two terms in A. That is to show that

$$W_{[n]}(F \otimes I + I \otimes F) = (F \otimes I + I \otimes F)W_{[n]}.$$

Note that $W_{[n]}^{-1} = W_{[n]}$, equation (6.5) implies the above equality. $\square$

PROPOSITION 7.6. *Let the eigenvalues of F_1 in (7.6) be $\sigma = \{\lambda_1, \cdots, \lambda_n\}$. Then in equation (7.8) the eigenvalues of B_{kk} are*

$$\sigma(B_{kk}) = \{\lambda_{i_1} + \cdots + \lambda_{i_k} \mid i_1, \cdots, i_k = 1, \cdots, n\}.$$

Proof. Since $B_{kk} = \tilde{A}_{kk}^T$ and the eigenvalues of A_{kk} is σ_k, it is enough to show that $\tilde{A}_{kk}$ and A_{kk} have same eigenvalues. Let μ be an eigenvalue of $\tilde{A}_{kk}$. Then there exists a $\tilde{P} \neq 0$ such that

$$\tilde{P}\tilde{A}_{kk} = \mu\tilde{P}.$$

By definition, $\tilde{A}_{kk} = T_B(n,k)A_{kk}T_N(n,k)$. Note that $T_B(n,k)T_N(n,k) = I$, then

$$\tilde{P}T_B(n,k)A_{kk}T_N(n,k)\begin{pmatrix} x_{(1)} \\ x_{(2)} \\ \vdots \\ x_{(k)} \end{pmatrix} = \mu\tilde{P}T_B(n,k)T_N(n,k)\begin{pmatrix} x_{(1)} \\ x_{(2)} \\ \vdots \\ x_{(k)} \end{pmatrix}. \tag{7.10}$$

Set $P = \tilde{P}T_B(n,k)$, then $P \neq 0$ is a symmetric set. For P, (7.10) becomes

$$PA_{kk}\begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^k \end{pmatrix} = \mu P\begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^k \end{pmatrix}. \tag{7.11}$$

According to Lemma 7.5, PA_{kk} is still a symmetric set. By the uniqueness of symmetric coefficients, we have

$$PA_{kk} = \mu P.$$

Hence, μ is also an eigenvalue of A_{kk}.

Conversely, assume μ is an eigenvalue of A_{kk}. Then $\mu = \lambda_{i_1} + \cdots + \lambda_{i_k}$. Denote by Y_j the eigenvector of F_1 with respect to λ_{i_j}, then we construct

$$Y = \sum_{\sigma \in S_k} Y_{\sigma(1)} \otimes \cdots \otimes Y_{\sigma(k)},$$

where S_k is the k-th order symmetric group. It is obvious that

$$YA_{kk} = \mu Y.$$

Since Y is symmetric, there exists $\tilde{Y} \neq 0$ such that $Y = \tilde{Y}T_B(n,k)$. It follows that

$$\tilde{Y}T_B(n,k)A_{kk} = \mu\tilde{Y}T_B(n,k).$$

Right multiply both sides of the above equation by $T_N(n,k)$ yields

$$\tilde{Y}\tilde{A}_{kk} = \mu\tilde{Y}.$$

That is, μ is an eigenvalue of $\tilde{A}_{kk}$. $\square$

PROPOSITION 7.7. *1. If equation (7.8) has solution $h \neq 0$, then*

$$\xi = c_1\lambda_{i_1} + \cdots + c_s\lambda_{i_s}.$$

where $\lambda_{i_1}, \cdots, \lambda_{i_s} \in \sigma(F_1)$; $c_1, \cdots, c_s$ are either 1 or 0.
2. If h has a component $h_j \neq 0$, then $\xi \in \sigma^j$. If h has t non-zero components, $h_{i_1} \neq 0, \cdots, h_{i_t} \neq 0$, then σ^s has at least a t fold elements. Here $\sigma^t = \{c_1\lambda_{i_1} + \cdots + c_t\lambda_{i_t} \mid c_1, \cdots, c_t \in \{0,1\}\}$.
3. If (7.6) has a linear first integral $H(t,x) = e^{-\xi t}h^T x$, then for any integer $j > 0$, $H_j(t,x) = e^{-j\xi t}(h^T)^j x^j$ is a first integral of (7.6).

Proof. 1 and 2 are the immediate consequence of the Proposition 7.6. We prove 3. If (7.6) has a linear first integral $H(t,x) = e^{-\xi t}h^T X$, then

$$\begin{cases} F_1 h = \xi h, \\ F_i h = 0, \quad i = 2, \cdots, k. \end{cases}$$

Let $p = (0_n, 0_{n^2}, \cdots, 0_{n^{j-1}}, h^j)$, where 0_k is the zero vector in R^k. Since

$$A_{j,j+s-1} = I_{n^{j-1}} \otimes F_s + I_{n^{j-2}} \otimes F_s \otimes I + \cdots + F_s \otimes I_{n^{j-1}},$$

then

$$\begin{cases} A_{jj}h^j = j\xi h^j, \\ A_{jt}h^j = 0, \quad t = j+1, \cdots, j+k-1, \end{cases}$$

which implies that p satisfies (7.7) with ξ being replaced by $j\xi$. $\square$

Proposition 6.7 provides a convenient tool for searching the first integrals. In fact, after fixed the ξ, the problem becomes a problem of solving linear algebraic system. For Lorenz system, 1–2 of Proposition 7.7 are known [1]. So the current result stated here is a generalization of their work.

EXAMPLE 7.8. Lotka-Volterra equations model the interactions between biological species and chemical reactions. Lotka-Volterra equation can be written as

$$\dot{x}_i = x_i\left(a_i + \sum_{j=1}^n b_{ij}x_j\right), \quad i = 1, \cdots, n. \tag{7.12}$$

Let $n = 2$ [11]. Set

$$A_{11} = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}, \quad A_{12} = \begin{pmatrix} b_{11} & b_{12} & 0 & 0 \\ 0 & 0 & b_{21} & b_{22} \end{pmatrix}.$$

Then

$$A_{22} = A_{11} \otimes I_2 + I_2 \otimes A_{11} = \begin{pmatrix} 2a_1 & 0 & 0 & 0 \\ 0 & a_1+a_2 & 0 & 0 \\ 0 & 0 & a_1+a_2 & 0 \\ 0 & 0 & 0 & 2a_2 \end{pmatrix},$$

$$A_{23} = A_{12} \otimes I_2 + I_2 \otimes A_{12}$$
$$= \begin{pmatrix} 2b_{11} & b_{12} & b_{12} & 0 & 0 & 0 & 0 & 0 \\ 0 & b_{11} & b_{21} & b_{12}+b_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & b_{11}+b_{21} & b_{12} & b_{22} & 0 \\ 0 & 0 & 0 & 0 & 0 & b_{21} & b_{21} & 2b_{22} \end{pmatrix}.$$

The Carleman linearized form becomes

$$\begin{pmatrix} \dot{x} \\ \dot{x}^2 \\ \dot{x}^3 \\ \vdots \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & 0 & 0 & \cdots \\ 0 & A_{22} & A_{23} & 0 & \cdots \\ 0 & 0 & A_{33} & A_{34} & \cdots \\ \vdots & & & & \end{pmatrix} \begin{pmatrix} x \\ x^2 \\ x^3 \\ \vdots \end{pmatrix}.$$

Say, we search for the first integral of the form: $H(t,x) = e^{-\xi t}(P_1 x + P_2 x^2)$. Then

$$T_B(2,2) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad T_N(2,2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$T_N(2,3) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$B_{11} = A_{11}^T, \quad B_{21} = \tilde{A}_{12}^T = T_B(2,1)A_{12}T_N(2,2) = \begin{pmatrix} b_{11} & b_{12} & 0 \\ 0 & b_{21} & b_{22} \end{pmatrix},$$

$$B_{22}^T = T_B(2,2)A_{22}T_N(2,2) = \begin{pmatrix} 2a_1 & 0 & 0 \\ 0 & a_1+a_2 & 0 \\ 0 & 0 & 2a_2 \end{pmatrix},$$

$$B_{32}^T = T_B(2,2)A_{23}T_N(2,3) = \begin{pmatrix} 2b_1 & 2b_{12} & 0 & 0 \\ 0 & b_{11}+b_{21} & b_{12}+b_{22} & 0 \\ 0 & 0 & 2b_{21} & 2b_{22} \end{pmatrix}.$$

We conclude that the second degree integral exists iff the following equation has non-zero solution:

$$\begin{cases} \begin{pmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \xi \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}, \\ B_{32}h_2 = 0, \end{cases}$$

where the only possible ξ should be $\xi \in \{a_1, a_2, 2a_1, 2a_2, a_1+a_2\}$. $\square$

8. Invariants of Planar Polynomial Systems. In this section we consider the invariants of planar polynomial systems with non-polynomial invariants. Consider a polynomial system

$$\dot{x} = F_0 + F_1 x + \cdots + F_k x^k, \quad x \in \mathbb{R}^2. \tag{8.1}$$

The invariants considered are of the form

$$H(x,t) = e^{\xi t} x_1^{\alpha} x_2^{\beta} (P_0 + P_1 x + \cdots + P_l x^l). \tag{8.2}$$

This kind of invariants was investigated in [1]. Recently, the Darboux method has been used in searching such invariants [2].

Our purpose is to convert the problem into a set of algebraic equations. Using (5.3), the time derivative of $H(x,t)$ is

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + DH \cdot \dot{x}$$
$$= \xi e^{\xi t} x_1^{\alpha} x_2^{\beta} (P_0 + P_1 x + \cdots + P_l x^l)$$
$$\quad + e^{\xi t} (\alpha x_1^{\alpha-1} x_2^{\beta}, \ \beta x_1^{\alpha} x_2^{\beta-1})(P_0 + P_1 x + \cdots + P_l x^l)(F_0 + F_1 x + \cdots + F_k x^k)$$
$$\quad + e^{\xi t} x_1^{\alpha} x_2^{\beta} (P_1 + P_2 \Phi_1 x + \cdots + P_l \Phi_{l-1} x^{l-1})(F_0 + F_1 x + \cdots + F_k x^k).$$

Setting

$$\frac{dH}{dt} = 0$$

and noting that

$$x_1 x_2 = (0 \ 1 \ 0 \ 0)x^2, \quad (\alpha x_2, \ \beta x_1) = (0 \ \beta \ \alpha \ 0)x,$$

we have

$$\xi(0 \ 1 \ 0 \ 0)x^2 (P_0 + P_1 x + \cdots + P_l x^l)$$
$$\quad + (0 \ \beta \ \alpha \ 0)x(P_0 + P_1 x + \cdots + P_l x^l)(F_0 + F_1 x + \cdots + F_k x^k)$$
$$\quad + (0 \ 1 \ 0 \ 0)x^2 (P_1 + P_2 \Phi_1 x + \cdots + P_l \Phi_{l-1} x^{l-1})(F_0 + F_1 x + \cdots + F_k x^k) \tag{8.3}$$
$$= 0.$$

Using (6.2), it can be expressed as

$$\xi(0 \ 1 \ 0 \ 0)[(I_4 \otimes P_0)x^2 + (I_4 \otimes P_1)x^3 + \cdots + (I_4 \otimes P_l)x^{l+2}]$$
$$\quad + (0 \ \beta \ \alpha \ 0)[(I_2 \otimes P_0)x + (I_2 \otimes P_1)x^2 + \cdots + (I_2 \otimes P_l)x^{l+1}]$$
$$\quad \times (F_0 + F_1 x + \cdots + F_k x^k)$$
$$\quad + (0 \ 1 \ 0 \ 0)[(I_4 \otimes P_1)x^2 + (I_4 \otimes P_2 \Phi_1)x^3 + \cdots + (I_4 \otimes P_l \Phi_{l-1})x^{l+1}]$$
$$\quad \times (F_0 + F_1 x + \cdots + F_k x^k)$$
$$= 0.$$

Using (6.2) again, we can multiply the products out to get

$$\sum_{s=1}^{l+1} \xi(0 \ 1 \ 0 \ 0)[I_4 \otimes P_{s-1}]x^{s+1}$$
$$\quad + \sum_{s=0}^{k+l} (0 \ \beta \ \alpha \ 0) \sum_{i=0,j=0}^{i+j=s} [(I_2 \otimes P_i)(I_{2^{i+1}} \otimes F_{s-i})]x^{s+1}$$
$$\quad + \sum_{s=1}^{k+l} (0 \ 1 \ 0 \ 0) \sum_{i=1,j=0}^{i+j=s} [(I_4 \otimes P_i \Phi_{i-1})(I_{2^{i+1}} \otimes F_{s-i})]x^{s+1}$$
$$= 0.$$

Converting term by term to natural basis, we have the following result.

THEOREM 8.1. *The system (8.1) has invariant of the form of (8.2), iff the following algebraic equations have solution* $(\xi, \alpha, \beta, P_0, \cdots, P_l)$:

$$(0 \; \beta \; \alpha \; 0)[(I_2 \otimes P_0)(I_2 \otimes F_0)] = 0$$

$$
\begin{cases}
\xi(0 \; 1 \; 0 \; 0)(I_4 \otimes P_{s-1}) + (0 \; \beta \; \alpha \; 0) \sum_{i=0,j=0}^{i+j=s} [(I_2 \otimes P_i)(I_{2^{i+1}} \otimes F_{s-i})] \\
\quad + (0 \; 1 \; 0 \; 0) \sum_{i=1,j=0}^{i+j=s} [(I_4 \otimes P_i \Phi_{i-1})(I_{2^{i+1}} \otimes F_{s-i})]
\end{cases} T_N(2, s+1) = 0,
$$

$$s = 1, \cdots, l+1 \tag{8.4}$$

$$
\begin{cases}
(0 \; \beta \; \alpha \; 0) \sum_{i=0,j=0}^{i+j=s} [(I_2 \otimes P_i)(I_{2^{i+1}} \otimes F_{s-i})] \\
\quad + (0 \; 1 \; 0 \; 0) \sum_{i=1,j=0}^{i+j=s} [(I_4 \otimes P_i \Phi_{i-1})(I_{2^{i+1}} \otimes F_{s-i})]
\end{cases} T_N(2, s+1) = 0,
$$

$$s = l+2, \cdots, l+k.$$

REMARK. *The advantage of this approach lies on: 1. It can be solved numerically by computer. 2. The approach can be easily extended to higher dimensional case.*

To depict the second item in above remark, consider the case of $n = 3$. $3D$ Lotka-Volterra system has been discussed in [3]. Using our approach, assume the system considered is (8.1) with $x \in \mathbb{R}^3$, and the invariants are of the form:

$$H(x, t) = e^{\xi t} x_1^\alpha x_2^\beta x_3^\gamma (P_0 + P_1 x + \cdots + P_l x^l). \tag{8.5}$$

Then one sees easily that

$$x_1 x_2 x_3 = (\underbrace{00000}_{5} 1 \underbrace{0 \cdots 0}_{21}) x^3 := \delta x^3;$$

and

$$(\alpha x_2 x_3 \;\; \beta x_1 x_3 \;\; \gamma x_1 x_2) = (0 \;\; \gamma \;\; \beta \;\; 0 \;\; 0 \;\; \alpha \;\; 0 \;\; 0 \;\; 0) x^2 := \eta x^2.$$

Then the corresponding (8.3) becomes

$$
\begin{aligned}
&\xi \delta x^3 (P_0 + P_1 x + \cdots + P_l x^l) \\
&\quad + \eta x^2 (P_0 + P_1 x + \cdots + P_l x^l)(F_0 + F_1 x + \cdots + F_k x^k) \\
&\quad + \delta x^3 (P_1 + P_2 \Phi_1 x + \cdots + P_l \Phi_{l-1} x^{l-1})(F_0 + F_1 x + \cdots + F_k x^k) \\
&= 0.
\end{aligned} \tag{8.6}
$$

Then the rest argument in the above remains available for producing the set of algebraic equations.

This approach provides only the algebraic equations for the solutions. It doesn't provide all detailed solutions as in [3].

9. Contraction of Tensor Field. The semi-tensor product is a powerful tool for treating multi-dimensional data in matrix form. This is one of its major contributions.

The following statement is cited from [6]: "It is also sometimes convenient to use matrix methods to handle the summations over repeated suffixes. These methods are restricted to quantities carrying either one or two suffixes, enabling them to be arranged as either one-dimensional arrays (row vectors or column vectors) or two-dimensional arrays (matrices)." In the above statement one sees easily that in conventional way only one or two dimensional data can be treated in matrix form. But semi-tensor product, collaborated with tensor product and swap matrix etc., can treat higher dimensional data in matrix way easily.

For instance, in this section we consider the contraction of a tensor field [6, 10]. Let $\sigma \in T^r_s(V)$ be a tensor field of the type (r, s) on an n dimensional vector space V, $1 \leq p \leq r$, $1 \leq q \leq s$. We define a contraction, $\pi^p_q : T^r_s(V) \rightarrow T^{r-1}_{s-1}(V)$ in the following way: Fix a basis $(d_1, \cdots, d_n)$ for V and its dual basis $(e^1, \cdots, e^n)$ for V^* respectively. For $\omega \in T^r_s(V)$, denoted by

$$\omega^{i_1 \cdots i_r}_{j_1 \cdots j_s} = \omega(d_{i_1}, \cdots, d_{i_r}; e^{j_1}, \cdots, e^{j_s}), \quad i_1, \cdots, i_r, j_1, \cdots, j_s = 1, \cdots, n.$$

Then we get an n^{r+s}-dimensional data, and arrange them into a matrix, M_ω, as

$$M_\omega = \begin{pmatrix} \omega^{1\cdots11}_{1\cdots11} & \omega^{1\cdots12}_{1\cdots11} & \cdots & \omega^{n\cdots nn}_{1\cdots11} \\ \omega^{1\cdots11}_{1\cdots12} & \omega^{1\cdots12}_{1\cdots12} & \cdots & \omega^{n\cdots nn}_{1\cdots12} \\ \vdots & & & \\ \omega^{1\cdots11}_{n\cdots nn} & \omega^{1\cdots12}_{n\cdots nn} & \cdots & \omega^{n\cdots nn}_{n\cdots nn} \end{pmatrix}. \tag{9.1}$$

Matrix M_ω is called the structure array of the tensor ω. Now it is not difficult to verify the following formula:

$$\begin{aligned} \omega(\sigma_1, \cdots, \sigma_s; X_1, \cdots, X_r) \\ = \sigma_s \ltimes \cdots \ltimes \sigma_1 \ltimes M_\omega \ltimes X_1 \ltimes \cdots \ltimes X_r. \end{aligned} \tag{9.2}$$

Next we define the contraction, $\pi^p_q(\sigma)$ by its structure matrix, with the entries determined by

$$\omega^{i_1 \cdots \hat{i}_p \cdots i_r}_{j_1 \cdots \hat{j}_q \cdots j_s} = \sum_{i_p = j_q} \omega^{i_1 \cdots i_p \cdots i_r}_{j_1 \cdots j_q \cdots j_s}. \tag{9.3}$$

Where the "$\hat{}$" means the corresponding index is missed.

Since the definition depends on the basis (generally, in tensor field case, it depends on the coordinate change), we have to show that this definition is independent of the coordinate change. It was said that under the suffix form, the proof is "cumbersome" [6]. We will give an elegant matrix proof. In addition, the structure matrix of the contracted tensor will be obtained.

First of all, we give the structure matrix of $\pi^p_q(\sigma)$. Let $\xi = n^{s-1}$ and $\eta = n^{r-1}$. Then the structure matrix M_σ can be split as

$$M_\sigma = \begin{pmatrix} M_{11} & \cdots & M_{1\eta} \\ \vdots & & \\ M_{\xi 1} & \cdots & M_{\xi\eta} \end{pmatrix}, \tag{9.4}$$

where each block M_{ij} is an $n \times n$ matrix.

Then a straightforward computation shows that

LEMMA 9.1.   *Assume $p = r$ and $q = s$. Then we have*

$$M_{\pi_s^r(\sigma)} = \begin{pmatrix} \mathrm{tr}(M_{11}) & \cdots & \mathrm{tr}(M_{1\eta}) \\ \vdots & & \\ \mathrm{tr}(M_{\xi 1}) & \cdots & \mathrm{tr}(M_{\xi\eta}) \end{pmatrix} := TR(M_\sigma). \tag{9.5}$$

The operator TR takes trace for all split $n \times n$ blocks.

Now for general case, we have to swap the index p with r, and the index q with s. The swap of any two elements can be realized by a sequence of the swaps of two adjacent elements. Using (6.6), we have the swapped structure matrix as the follows:

$$\begin{aligned} \tilde{M}_\sigma &= \textstyle\prod_{t=0}^{s-q-1}(I_{n^{s-2-t}} \otimes W_{[n]} \otimes I_{n^t})M_\sigma \prod_{t=0}^{r-p-1}(I_{n^{r-2-t}} \otimes W_{[n]} \otimes I_{n^t}) \\ &:= \Pi_1 M_\sigma \Pi_2. \end{aligned} \tag{9.6}$$

Similar to M, we can split $\tilde{M}$ into $\xi \times \eta$ blocks of $n \times n$ matrices, denoted by $\tilde{M}_{ij}$. Then we have

PROPOSITION 9.2.   *The structure matrix of $\pi_q^p(\sigma)$ is*

$$M_{\pi_q^p(\sigma)} = TR(\tilde{M}_\sigma) = TR(\Pi_1 M_\sigma \Pi_2). \tag{9.7}$$

We give a simple example to show this contraction.

EXAMPLE 9.3.   Let $n = 2$, $r = 2$, and $s = 3$. We consider $\pi_1^1(\sigma)$. Denote

$$M_\sigma = \begin{pmatrix} a_{111}^{11} & a_{111}^{12} & a_{111}^{21} & a_{111}^{22} \\ a_{112}^{11} & a_{112}^{12} & a_{112}^{21} & a_{112}^{22} \\ a_{121}^{11} & a_{121}^{12} & a_{121}^{21} & a_{121}^{22} \\ a_{122}^{11} & a_{122}^{12} & a_{122}^{21} & a_{122}^{22} \\ a_{211}^{11} & a_{211}^{12} & a_{211}^{21} & a_{211}^{22} \\ a_{212}^{11} & a_{212}^{12} & a_{212}^{21} & a_{212}^{22} \\ a_{221}^{11} & a_{221}^{12} & a_{221}^{21} & a_{221}^{22} \\ a_{222}^{11} & a_{222}^{12} & a_{222}^{21} & a_{222}^{22} \end{pmatrix}.$$

$$\begin{aligned} \Pi_1 &= \textstyle\prod_{t=0}^{1} I_{2^{1-t}} \otimes W_{[2]} \otimes I_{2^t} = (I_2 \otimes W_{[2]})(W_{[2]} \otimes I_2) \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

$$\Pi_2 = W_{[2]} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Then

$$\Pi_1 M_\sigma \Pi_2 = \begin{pmatrix} a^{11}_{111} & a^{21}_{111} & a^{12}_{111} & a^{22}_{111} \\ a^{11}_{211} & a^{21}_{211} & a^{12}_{211} & a^{22}_{211} \\ a^{11}_{112} & a^{21}_{112} & a^{12}_{112} & a^{22}_{112} \\ a^{11}_{212} & a^{21}_{212} & a^{12}_{212} & a^{22}_{212} \\ a^{11}_{121} & a^{21}_{121} & a^{12}_{121} & a^{22}_{121} \\ a^{11}_{221} & a^{21}_{221} & a^{12}_{221} & a^{22}_{221} \\ a^{11}_{122} & a^{21}_{122} & a^{12}_{122} & a^{22}_{122} \\ a^{11}_{222} & a^{21}_{222} & a^{12}_{222} & a^{22}_{222} \end{pmatrix}.$$

Using it, the Proposition 9.2 yields

$$M_{\pi^1_1(\sigma)} = \begin{pmatrix} a^{11}_{111} + a^{21}_{211} & a^{12}_{111} + a^{22}_{211} \\ a^{11}_{112} + a^{21}_{212} & a^{12}_{112} + a^{22}_{212} \\ a^{11}_{121} + a^{21}_{221} & a^{12}_{121} + a^{22}_{221} \\ a^{11}_{122} + a^{21}_{222} & a^{12}_{122} + a^{22}_{222} \end{pmatrix}.$$

Next, we prove that the structure matrix defined by (9.6) is independent of the coordinate change. Now assume we have a coordinate change as $z = z(x)$ with the Jacobian matrix as $J = \frac{\partial z}{\partial x}$.

The following Lemma can be verified by straightforward computation.

LEMMA 9.4. *1. Let $P \in M_{s \times m}$, $Q \in M_{n \times n}$, and $A_i \in M_{n \times rn}$, $i = 1, \cdots, m$. Set*

$$\tilde{A} = \begin{pmatrix} \tilde{A}_1 \\ \vdots \\ \tilde{A}_m \end{pmatrix} = (P \otimes Q) \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix} = (P \otimes Q)A.$$

Then

$$\begin{pmatrix} \tilde{A}_1 \\ \vdots \\ \tilde{A}_m \end{pmatrix} = P \ltimes \begin{pmatrix} QA_1 \\ \vdots \\ QA_m \end{pmatrix}.$$

Moreover,

$$TR(\tilde{A}) = P \cdot TR \begin{pmatrix} QA_1 \\ \vdots \\ QA_m \end{pmatrix}.$$

2. Let $P \in M_{m \times s}$, $Q \in M_{n \times n}$, and $A_i \in M_{nr \times n}$, $i = 1, \cdots, m$. Set

$$\tilde{A} = \left(\tilde{A}_1, \cdots, \tilde{A}_m \right) = (A_1, \cdots, A_m)(P \otimes Q) = A(P \otimes Q).$$

Then

$$\left(\tilde{A}_1, \cdots, \tilde{A}_m \right) = (A_1 Q, \cdots, A_m Q) P.$$

Moreover,

$$TR(\tilde{A}) = TR\left(A_1 Q, \cdots, A_m Q\right) P.$$

Now we are ready to prove the main result: the contraction is well defined.

THEOREM 9.5. *The contraction defined by (9.3) is independent of the coordinates.*

Proof. Let

$$M_\sigma = \begin{pmatrix} M_{11} & \cdots & M_{1\eta} \\ \vdots & & \\ M_{\xi 1} & \cdots & M_{\xi\eta} \end{pmatrix}.$$

Using Proposition 9.2, we have

$$M_{\pi_q^p(\sigma)} = TR\left(\Pi_1 M_\sigma \Pi_2\right).$$

Now consider a coordinate change $z = z(x)$, then M_σ becomes $\bar{M}_\sigma$, which is

$$\bar{M}_\sigma = \underbrace{J^{-1} \otimes \cdots \otimes J^{-1}}_{s} M_\sigma \underbrace{J \otimes \cdots \otimes J}_{t}.$$

Notice that Π_1 is commutative with $\underbrace{J^{-1} \otimes \cdots \otimes J^{-1}}_{s}$, and Π_2 is commutative with $\underbrace{J \otimes \cdots \otimes J}_{t}$. Applying Proposition 9.2 to $\bar{M}_\sigma$, we have

$$\begin{aligned}
\bar{M}_{\pi_q^p(\sigma)} &= TR\left(\Pi_1 \underbrace{(J^{-1} \otimes \cdots \otimes J^{-1})}_{s} M_\sigma \underbrace{(J \otimes \cdots \otimes J)}_{t} \Pi_2\right) \\
&= TR\left(\underbrace{(J^{-1} \otimes \cdots \otimes J^{-1})}_{s}(\Pi_1 M_\sigma \Pi_2)\underbrace{(J \otimes \cdots \otimes J)}_{t}\right) \\
&= TR\left(\underbrace{((J^{-1} \otimes \cdots \otimes J^{-1})}_{s-1} \otimes J^{-1})(\tilde{M}_\sigma)(\underbrace{(J \otimes \cdots \otimes J)}_{t-1} \otimes J)\right) \\
&= \underbrace{(J^{-1} \otimes \cdots \otimes J^{-1})}_{s-1} TR(J^{-1}(\tilde{M}_\sigma)J)\underbrace{(J \otimes \cdots \otimes J)}_{t-1} \\
&= \underbrace{(J^{-1} \otimes \cdots \otimes J^{-1})}_{s-1} TR(\tilde{M}_\sigma)\underbrace{(J \otimes \cdots \otimes J)}_{t-1} \\
&= \underbrace{(J^{-1} \otimes \cdots \otimes J^{-1})}_{s-1} M_{\pi_q^p(\sigma)}\underbrace{(J \otimes \cdots \otimes J)}_{t-1}.
\end{aligned}$$

Note that the last third equality is from Lemma 9.4. The proof is completed. □

10. Conclusion. In this paper the left and right semi-tensor products of any two matrices were defined and several basic properties were obtained. Then under the assumption A1 a few further properties, which are necessary for the later discussions, were investigated.

As applications of the semi-tensor products to physics, some physical problems were considered. First of all, the general formula for the Carlman's linearization was presented. The necessary and sufficient condition for the existence of polynomial type of first integrals was proved. Then for planar polynomial systems a non-polynomial type of invariants was investigated. The problem was converted to the solvability of a set of algebraic equations. Another problem considered is the contraction of tensor fields, which is useful in relativity etc. We gave a rigorous proof and provided a matrix structure for the contracted tensor field.

REFERENCES

[1] L. CAIRO, M. R. FEIX, *Families of invariants of the motion for the Lotka-Volterra 1st family*, J. Math. Phys., 33:7 (1992),pp. 2240–2455.
[2] L. CAIRO, M. R. FEIX, J. LLIBRE, *Darboux method and search of invariats for the Lotka-Volterra and Complex Quadratic Systems*, J. Math. Phys., 40:4 (1999), pp. 2074–2091.
[3] L. CAIRO, *Darboux first integral conditions and integrability of the 3D Lotka-Volterra system*, J. Nnolinear Math. Phys., 7:4 (2000), pp. 511–531.
[4] D. CHENG, *Semi-tensor product of matrices and its application to Morgen's problem*, Chinese Science, series F, 44:3 (2001), pp. 195–212.
[5] D. CHENG, X. HU, Y. WANG, *Non-regular Feedback Linearization of Nonlinear Systems via a Normal Form Algorithm*, Automatica, 40:3 (2004), pp. 439–447.
[6] J. FOSTER, J. D. NIGHTNGALE, *A Short Course in General Relativity*, Springer-Verlag, 1995.
[7] R. HORN, C. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
[8] J. R. MAGNUS, H. NEUDECKER, *The commutation matrix: some properties and applications*, Annals of Statistics, 7 (1979), pp. 381–394.
[9] J. R. MAGNUS, H. NEUDECKER, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Revised Ed., John Wiley & Sons, 1999.
[10] R. K. SACHS, H. WU, *General Relativity for Mathematicians*, Springer-Verlag, 1977.
[11] W. M. STEEB, F. WILHELM, *Non-linear autonomous systems of differential equations and Carleman linearization procedure*, J. Math. Anal. Appl., 77 (1980), pp. 601–611.