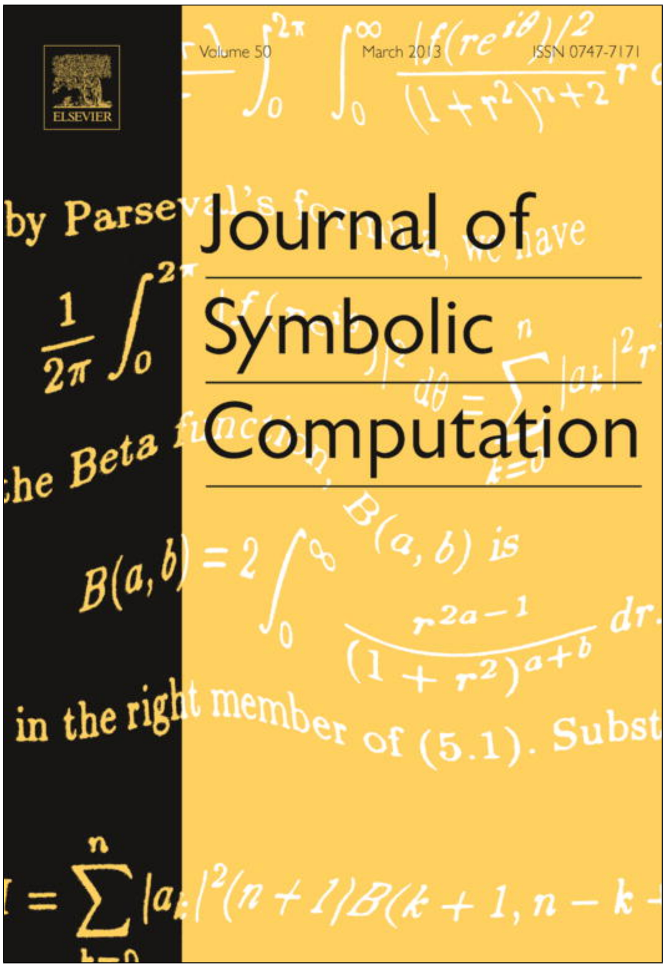
ELSEVIER
Volume 50
March 2013
ISSN 0747-7171
Journal of
Symbolic
Computation

Journal of Symbolic Computation 50 (2013) 508–531

Contents lists available at SciVerse ScienceDirect

Journal of Symbolic Computation

www.elsevier.com/locate/jsc

On determinants and eigenvalue theory of tensors

Shenglong Hu [a,1], Zheng-Hai Huang [b,2], Chen Ling [c,3], Liqun Qi [a,4]

[a] *Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong*
[b] *Department of Mathematics, School of Science, Tianjin University, Tianjin, China*
[c] *School of Science, Hangzhou Dianzi University, Hangzhou 310018, China*

ARTICLE INFO

Article history:
Received 12 August 2012
Accepted 15 October 2012
Available online 29 October 2012

Keywords:
Tensor
Eigenvalue
Determinant
Characteristic polynomial

ABSTRACT

We investigate properties of the determinants of tensors, and their applications in the eigenvalue theory of tensors. We show that the determinant inherits many properties of the determinant of a matrix. These properties include: solvability of polynomial systems, product formula for the determinant of a block tensor, product formula of the eigenvalues and Geršgorin's inequality. As a simple application, we show that if the leading coefficient tensor of a polynomial system is a triangular tensor with nonzero diagonal elements, then the system definitely has a solution in the complex space. We investigate the characteristic polynomial of a tensor through the determinant and the higher order traces. We show that the k-th order trace of a tensor is equal to the sum of the k-th powers of the eigenvalues of this tensor, and the coefficients of its characteristic polynomial are recursively generated by the higher order traces. Explicit formula for the second order trace of a tensor is given.

E-mail addresses: Tim.Hu@connect.polyu.hk (S. Hu), huangzhenghai@tju.edu.cn (Z.-H. Huang), macling@hdu.edu.cn (C. Ling), maqilq@polyu.edu.hk (L. Qi).

[1] This author's work was partially supported by the National Natural Science Foundation of China (Grant No. 11101303).

[2] This author's work was supported by the National Natural Science Foundation of China (Grant No. 11171252). Fax: +86 22 27403615.

[3] This author's work was supported by the National Natural Science Foundation of China (Grant No. 10971187 and Grant No. 11171083), and the Zhejiang Provincial Natural Science Foundation of China (Grant No. Y6100366).

[4] This author's work was supported by the Hong Kong Research Grant Council (Grant No. PolyU 501808, 501909, 502510 and 502111).

0747-7171/$ – see front matter
http://dx.doi.org/10.1016/j.jsc.2012.10.001

1. Introduction

Eigenvalues of tensors, proposed by Qi (2005) and Lim (2005) independently, have attracted much attention in the literature and found various applications in science and engineering (see Balan, 2012; Cartwright and Sturmfels, 2011; Chang et al. 2008, 2011; Comon et al., 2008; Hu and Qi, 2012; Hu et al., 2012; Li et al., 2013; Lim, 2007; Ng et al., 2009; Pearson, 2010; Qi 2006, 2007, 2011; Qi et al., 2010; Ragnarsson and Van Loan, 2011; Rota Bulò and Pelillo, 2009; Yang and Yang, 2010; Zhang et al., 2012, and references therein). There are several generalizations of eigenvalues, singular values and decompositions from symmetric matrices to higher order symmetric tensors (see Brachat et al., 2010; Cartwright and Sturmfels, 2011; Chang et al., 2009; Comon et al., 2008; Lim, 2005; Oeding and Ottaviani, 2011; Qi 2005, 2006, and references therein). In any case, not all the properties of the eigenvalues of matrices are preserved for higher order tensors. In this paper, we concentrate on the eigenvalues introduced by Qi (2005) (see Definition 2.1).

The concept of symmetric hyperdeterminant was introduced by Qi (2005) to investigate the eigenvalues of a symmetric tensor. It is based on the resultant of a homogeneous polynomial system, which is defined in Definition 1.1 (Theorem 3.2.3 in Cox et al., 1998; Morozov and Shakirov 2010, 2011).

For a positive integer n, an n-tuple $\alpha = (\alpha_1, \dots, \alpha_n)$ of nonnegative integers, and an n-tuple $\mathbf{x} := (x_1, \dots, x_n)^T$ of indeterminate variables, denote by $\mathbf{x}^\alpha$ the monomial $\prod_{i=1}^n x_i^{\alpha_i}$. Denote by $\mathbb{C}$ the field of complex numbers, and $\mathbb{Z}$ the ring of integers.

Definition 1.1. For fixed positive degrees $d_1, \dots, d_n$, let $f_i := \sum_{|\alpha|=d_i} c_{i,\alpha}\mathbf{x}^\alpha$ be a homogeneous polynomial of degree d_i in $\mathbb{C}[\mathbf{x}]$ for $i \in \{1, \dots, n\}$. Then the unique polynomial $\mathrm{RES}_{d_1,\dots,d_n} \in \mathbb{Z}[\{u_{i,\alpha}\}]$, which has the following properties, is called the *resultant* of degrees $(d_1, \dots, d_n)$.

(i) The system of polynomial equations $f_1 = \cdots = f_n = 0$ has a nontrivial solution in $\mathbb{C}^n$ if and only if $\mathrm{Res}_{d_1,\dots,d_n}(f_1, \dots, f_n) = 0$.
(ii) $\mathrm{Res}_{d_1,\dots,d_n}(x_1^{d_1}, \dots, x_n^{d_n}) = 1$.
(iii) $\mathrm{RES}_{d_1,\dots,d_n}$ is an irreducible polynomial in $\mathbb{C}[\{u_{i,\alpha}\}]$.

The differences between the capital notation $\mathrm{RES}_{d_1,\dots,d_n}$ and the notation $\mathrm{Res}_{d_1,\dots,d_n}(f_1, \dots, f_n)$ for a specific system $(f_1, \dots, f_n)$ are: the former is understood as a polynomial in the variables $\{u_{i,\alpha} \mid |\alpha| = d_i,\ i \in \{1, \dots, n\}\}$ and the latter is understood as the evaluation of $\mathrm{RES}_{d_1,\dots,d_n}$ at the point $\{u_{i,\alpha} = c_{i,\alpha}\}$ with $\{c_{i,\alpha}\}$ being given by f_i. Consequently, $\mathrm{Res}_{d_1,\dots,d_n}(f_1, \dots, f_n)$ is a number in $\mathbb{C}$. When $d_1 = \cdots = d_n = d$, we simplify $\mathrm{RES}_{d,\dots,d}$ (respectively $\mathrm{Res}_{d,\dots,d}$) as RES (respectively Res). The value of d will be clear from the content.

Let $\mathcal{T} = (t_{i_1\dots i_m})$ be an m-th order n-dimensional tensor, $\mathbf{x} = (x_i) \in \mathbb{C}^n$ (the n-dimensional complex space) and $\mathcal{T}\mathbf{x}^{m-1}$ be an n-dimensional vector with its i-th element being

$$\sum_{i_2=1}^{n} \cdots \sum_{i_m=1}^{n} t_{ii_2\dots i_m} x_{i_2} \dots x_{i_m}.$$

It is actually a tensor contraction (Landsberg, 2011). The symmetric hyperdeterminant for symmetric tensors of order m is defined as the resultant RES of degrees $(m-1, \dots, m-1)$ such that the value of the symmetric hyperdeterminant for a specific symmetric tensor $\mathcal{T}$, which is denoted by $\mathrm{Res}(\mathcal{T}\mathbf{x}^{m-1})$, is the resultant of the polynomial system $\mathcal{T}\mathbf{x}^{m-1} = \mathbf{0}$. The symmetric hyperdeterminant of a symmetric tensor is equal to the product of all of the eigenvalues of that tensor (Qi, 2005).

Recently, Li et al. (2013) proved that the constant term of the E-characteristic polynomial[5] of tensor $\mathcal{T}$ (not necessarily symmetric) is a power of the resultant $\mathrm{Res}(\mathcal{T}\mathbf{x}^{m-1})$ of the polynomial system $\mathcal{T}\mathbf{x}^{m-1} = \mathbf{0}$. They further found that $\mathrm{Res}(\mathcal{T}\mathbf{x}^{m-1})$ is an invariant of $\mathcal{T}$ under the orthogonal linear transformation group. Li et al. (2013) pointed out that $\mathrm{Res}(\mathcal{T}\mathbf{x}^{m-1})$ deserves further study, since it

[5] The characteristic polynomial for another type of eigenvalues, E-eigenvalues, proposed by Qi (2005).

has close relation to the eigenvalue theory of tensors. In this paper, we generalize the definition of symmetric hyperdeterminant to nonsymmetric tensors and study it systematically. The following is the definition.

Definition 1.2. Let RES be the resultant of degrees $(m-1,\dots,m-1)$ which is a polynomial in variables $\{u_{i,\alpha} \mid |\alpha| = m-1,\ i \in \{1,\dots,n\}\}$. Let tensor $\mathcal{T} = (t_{ii_2\dots i_m}) \in \mathbb{T}(\mathbb{C}^n, m)$ (the space of m-th order n-dimensional tensors). The determinant DET of m-th order n-dimensional tensors is defined as the polynomial with variables $\{v_{ii_2\dots i_m} \mid i, i_2,\dots,i_m \in \{1,\dots,n\}\}$ through replacing $u_{i,\alpha}$ in the polynomial RES by $\sum_{(i_2,\dots,i_m)\in\mathbb{X}(\alpha)} v_{ii_2\dots i_m}$. Here $\mathbb{X}(\alpha) := \{(i_2,\dots,i_m) \in \{1,\dots,n\}^{m-1} \mid x_{i_2}\dots x_{i_m} = \mathbf{x}^\alpha\}$. The value of the determinant $\mathrm{Det}(\mathcal{T})$ of the specific tensor $\mathcal{T}$ is defined as the evaluation of DET at the point $\{v_{ii_2\dots i_m} = t_{ii_2\dots i_m}\}$.

For the convenience of the subsequent analysis, we define $\mathrm{DET}(\mathcal{T})$ as the polynomial with variables $\{t_{ii_2\dots i_m} \mid i, i_2,\dots,i_m \in \{1,\dots,n\}\}$ through replacing $v_{ii_2\dots i_m}$ in DET by $t_{ii_2\dots i_m}$. There can be some specific relations on the variables $\{t_{ii_2\dots i_m}\}$, such as some being zero. In this case, $\mathcal{T}$ is considered as a tensor of indeterminate variables, while it is considered as a tensor of numbers in $\mathbb{C}$ when we talk about $\mathrm{Det}(\mathcal{T})$.

Given a tensor $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$, we can associate to it a multilinear function $f : \mathbb{C}^n \times \dots \times \mathbb{C}^n \to \mathbb{C}$ as $f(\mathbf{x}^{(1)},\dots,\mathbf{x}^{(m)}) := \sum_{1\leqslant i_1,\dots,i_m\leqslant n} t_{i_1\dots i_m} x^{(1)}_{i_1}\dots x^{(m)}_{i_m}$. The hyperdeterminant is defined as the unique irreducible polynomial (up to a scalar factor) HDET such that its evaluation $\mathrm{Hdet}(\mathcal{T})$ at $\mathcal{T}$ is zero if and only if there are nonzero $\mathbf{x}^{(j)}$ for all $j \in \{1,\dots,m\}$ such that $\frac{\partial f}{\partial x^{(j)}_i} = 0$ for all $i \in \{1,\dots,n\}$ and $j \in \{1,\dots,m\}$. Then, the tensor determinant is different from the hyperdeterminant investigated in Bézout (1779); Cayley (1843, 1845); Cox et al. (1998); Dixon (1906); Gelfand et al. (1992, 1994); Macaulay (1902); de Silva and Lim (2008); Sylvester (1840).

It is easy to see from Definition 1.2 that the tensor determinant generalizes the matrix determinant (Horn and Johnson 1985, 1994; Strang, 1993) and the symmetric hyperdeterminant (Qi, 2005). Consequently, the notation $\mathrm{Det}(\cdot)$ is meaningful with both a matrix and a tensor as arguments. It should be pointed out that the same thing under the notion resultant[6] has been extensively studied in the monograph by Dolotin and Morozov (2007). In this paper, we give some new developments of the tensor determinant, and especially investigate some properties related to the eigenvalue theory of tensors proposed by Qi (2005) and Lim (2005).

The rest of this paper is organized as follows.

In the next section, we present some basic properties of the determinant. Then, in Section 3, we show that the solvability of a polynomial system is characterized by the determinant of the leading coefficient tensor of that polynomial system.

Block tensors are discussed in Section 4. We give an expression of the determinant of a tensor, which has an "upper block triangular" structure, in terms of the determinants of its two diagonal sub-tensors.

As a simple application of the determinant theory, we show, in Section 5, that if the leading coefficient tensor of a polynomial system is a triangular tensor with nonzero diagonal elements, then the system definitely has a solution in the complex space.

Based on a result of Morozov and Shakirov (2011), in Subsection 6.1, we give a trace formula for the determinant. This formula involves some differential operators. Using this formula, we establish an explicit formula for the determinant when the dimension is two. As this needs to use some results in Subsection 6.2, we do this in Section 7.

The determinant contributes to the characteristic polynomial theory of tensors. In Subsections 6.2 and 6.3, we analyze various related properties of the characteristic polynomial and the determinant. Especially, a trace formula for the characteristic polynomial is presented, which has potential applications in various areas, such as scientific computing and geometrical analysis of eigenvalues. We also

---

6 As resultant is created for a general polynomial system (Definition 1.1), we prefer to Definition 1.2 which is unambiguous as well.

generalize the eigenvalue representation for the determinant of a matrix to the determinant of a tensor. We show that the k-th order trace of a tensor is equal to the sum of the k-th powers of the eigenvalues of this tensor, and the coefficients of its characteristic polynomial are recursively generated by the higher order traces. Based on the characteristic polynomial, a sufficient and a necessary condition for the positive semidefiniteness problem of an even order tensor are presented (Proposition 6.14 and Remark 6.15).

In Section 7, we give explicit formulae for the second order trace of a tensor, and the determinant and the characteristic polynomial of a tensor when the dimension is two.

We generalize Geršgorin's inequality for the determinant of a matrix to that for a tensor in Section 8. It gives estimation for the determinant in terms of the entries of the underlying tensor.

Some final remarks are given in Section 9.

The following is the notation that is used in the sequel. Scalars are written as lowercase letters $(\lambda, a, \ldots)$; vectors are written as bold lowercase letters $(\mathbf{x} = (x_i), \ldots)$; matrices are written as italic capitals $(A = (a_{ij}), \ldots)$; tensors are written as calligraphic letters $(\mathcal{T} = (t_{i_1 \ldots i_m}), \ldots)$; and, sets are written as blackboard bold letters $(\mathbb{T}, \mathbb{S}, \ldots)$.

Given a ring $\mathbb{K}$ (hereafter, we mean a commutative ring with 1 (see Lang, 2002), e.g., $\mathbb{C}$), we denote by $\mathbb{K}[\mathbb{E}]$ the polynomial ring consisting of polynomials in the set $\mathbb{E}$ of indeterminate variables with coefficients in $\mathbb{K}$. Especially, we denote by $\mathbb{K}[\mathcal{T}]$ the polynomial ring consisting of polynomials in indeterminate variables $\{t_{i_1 \ldots i_m}\}$ with coefficients in $\mathbb{K}$, and similarly for $\mathbb{K}[\lambda]$, $\mathbb{K}[A]$, $\mathbb{K}[\lambda, \mathcal{T}]$, etc.

For a matrix A, A^T denotes its transpose and $\mathrm{Tr}(A)$ denotes its trace. We denote by $\mathbb{N}_+$ the set of all positive integers and $\mathbf{e}_i$ the i-th identity vector, i.e., the i-th column vector of the identity matrix I. Throughout this paper, unless stated otherwise, integers $m, n \geqslant 2$ and tensors refer to m-th order n-dimensional tensors with entries in $\mathbb{C}$. We use $\mathcal{T} = (t_{i_1 i_2 \ldots i_m})$ to mean both a tensor of indeterminate variables $t_{i_1 i_2 \ldots i_m}$ and a specific tensor in $\mathbb{T}(\mathbb{C}^n, m)$, which are clear from the content.

2. Basic properties of the determinant

Let $\mathcal{I}$ be the identity tensor of appropriate order and dimension, e.g., $i_{i_1 \ldots i_m} = 1$ if and only if $i_1 = \cdots = i_m \in \{1, \ldots, n\}$, and zero otherwise. The following definitions were introduced by Qi (2005).

Definition 2.1. Let $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$. For some $\lambda \in \mathbb{C}$, if polynomial system $(\lambda \mathcal{I} - \mathcal{T})\mathbf{x}^{m-1} = \mathbf{0}$ has a solution $\mathbf{x} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$, then λ is called an eigenvalue of the tensor $\mathcal{T}$ and $\mathbf{x}$ an eigenvector of $\mathcal{T}$ associated with λ.

We denote by $\sigma(\mathcal{T})$ the set of all eigenvalues of the tensor $\mathcal{T}$.

Definition 2.2. Let $\mathcal{T}$ be an m-th order n-dimensional tensor of indeterminate variables and λ be an indeterminate variable. The determinant $\mathrm{DET}(\lambda \mathcal{I} - \mathcal{T})$ of $\lambda \mathcal{I} - \mathcal{T}$ which is a polynomial in $(\mathbb{C}[\mathcal{T}])[\lambda]$, denoted by $\chi_{\mathcal{T}}(\lambda)$, is called the *characteristic polynomial* of the tensor $\mathcal{T}$.

For a specific $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$, $\chi_{\mathcal{T}}(\lambda) \in \mathbb{C}[\lambda]$. When there is no confusion, we simplify $\chi_{\mathcal{T}}(\lambda)$ as $\chi(\lambda)$. Denote by $\mathbb{V}(f)$ the algebraic set associated to the principal ideal $\langle f \rangle$ generated by f (Cox et al. 1998, 2006; Lang, 2002). By Definitions 1.1, 1.2, 2.1 and 2.2, we have the following result.

Theorem 2.3. *Let $\mathcal{T}$ be an m-th order n-dimensional tensor of indeterminate variables. Then $\chi(\lambda) \in \mathbb{C}[\lambda, \mathcal{T}]$ is homogeneous of degree $n(m-1)^{n-1}$, and for a specific tensor $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$,*

$$\mathbb{V}\big(\chi(\lambda)\big) = \sigma(\mathcal{T}). \tag{1}$$

When $\mathcal{T}$ is symmetric, Qi proved (1), see Theorem 1(a) in Qi (2005). If λ is a root of $\chi(\lambda)$ of multiplicity s, then we call s the *algebraic multiplicity* of eigenvalue λ.

For $f \in \mathbb{K}[\mathbf{x}]$, we denote by $\deg(f)$ the degree of f. If every monomial in f has degree $\deg(f)$, then f is called homogeneous of degree $\deg(f)$.

Proposition 2.4. *Let $\mathcal{T}$ be an m-th order n-dimensional tensor of indeterminate variables $t_{ii_2\ldots i_m}$. Then*:

(i) *For every $i \in \{1, \ldots, n\}$, define $\mathbb{K}_i$ as the polynomial ring*

$$\mathbb{C}\big[\big\{t_{ji_2\ldots i_m} \,\big|\, j, i_2, \ldots, i_m \in \{1, \ldots, n\},\ j \neq i\big\}\big].$$

Then $\mathrm{DET}(\mathcal{T}) \in \mathbb{K}_i[\{t_{ii_2\ldots i_m} \mid i_2, \ldots, i_m \in \{1, \ldots, n\}\}]$ *is homogeneous of degree* $(m-1)^{n-1}$.

(ii) $\mathrm{DET}(\mathcal{T}) \in \mathbb{C}[\mathcal{T}]$ *is irreducible and homogeneous of degree* $n(m-1)^{n-1}$.

(iii) $\mathrm{Det}(\mathcal{I}) = 1$.

Proof. Denote by RES the resultant of degrees $(m-1, \ldots, m-1)$ which is a polynomial in the variables $\{u_{i,\alpha} \mid |\alpha| = m-1,\ i \in \{1, \ldots, n\}\}$ by Definition 1.1. Then, by Proposition 13.1.1 in Gelfand et al. (1994) (see also p. 713 in Morozov and Shakirov (2011)), RES is homogeneous of degree $(m-1)^{n-1}$ in the variables $\{u_{i,\alpha} \mid |\alpha| = m-1\}$ for every $i \in \{1, \ldots, n\}$. Consequently, by the replacement in Definition 1.2, the determinant $\mathrm{DET}(\mathcal{T})$ is homogeneous of degree $(m-1)^{n-1}$ in the variables $\{t_{ii_2\ldots i_m} \mid i_2, \ldots, i_m \in \{1, \ldots, n\}\}$ for every $i \in \{1, \ldots, n\}$. This is exactly statement (i).

By (i), we immediately get that $\mathrm{DET}(\mathcal{T}) \in \mathbb{C}[\mathcal{T}]$ is homogeneous of degree $n(m-1)^{n-1}$. We claim that $\mathrm{DET}(\mathcal{T}) \in \mathbb{C}[\mathcal{T}]$ is irreducible. Suppose on the contrary that $\mathrm{DET}(\mathcal{T}) \in \mathbb{C}[\mathcal{T}]$ can be reduced as the product of two homogeneous polynomials in $\mathbb{C}[\mathcal{T}]$ as

$$\mathrm{DET}(\mathcal{T}) = f(\mathcal{T})g(\mathcal{T}) \tag{2}$$

with $\deg(f) \geqslant 1$ and $\deg(g) \geqslant 1$. If we replace the indeterminate variable $t_{ii_2\ldots i_m}$ with $(i_2, \ldots, i_m) \in \mathbb{X}(\alpha)$ in the tensor $\mathcal{T}$ by the variable $\frac{u_{i,\alpha}}{|\mathbb{X}(\alpha)|}$ and denote the resulting tensor by $\mathcal{U}$, then we get that

$$\mathrm{RES} = \mathrm{DET}(\mathcal{U}) = f(\mathcal{U})g(\mathcal{U}).$$

Here the first equality follows from Definition 1.2 and the second from (2). Obviously, $f(\mathcal{U}), g(\mathcal{U}) \in \mathbb{C}[\{u_{i,\alpha}\}]$ are nonzero and of degrees $\deg(f)$ and $\deg(g)$ respectively. Then, RES is reduced as a product of polynomials of positive degrees. This contradicts Definition 1.1(iii). Hence, $\mathrm{RES}(\mathcal{T}) \in \mathbb{C}[\mathcal{T}]$ is irreducible.

(iii) follows from Definitions 1.1(ii) and 1.2. □

By Proposition 2.4, we have the following corollary.

Corollary 2.5. *Let $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$. If for some i, $t_{ii_2\ldots i_m} = 0$ for all $i_2, \ldots, i_m \in \{1, \ldots, n\}$, then* $\mathrm{Det}(\mathcal{T}) = 0$. *In particular, the determinant of the zero tensor is zero.*

Proof. Let $\mathcal{V}$ be an m-th order n-dimensional tensor of indeterminate variables $v_{ii_2\ldots i_m}$ and $\mathbb{K}_i := \mathbb{C}[\{v_{ji_2\ldots i_m} \mid j, i_2, \ldots, i_m \in \{1, \ldots, n\},\ \ j \neq i\}]$. Then by Proposition 2.4(i), $\mathrm{DET}(\mathcal{V})$ is a homogeneous polynomial in the variable set $\{v_{ii_2\ldots i_m} \mid i_2, \ldots, i_m \in \{1, \ldots, n\}\}$ with coefficients in the ring $\mathbb{K}_i$. Moreover, $\mathrm{Det}(\mathcal{T})$ is just the evaluation of $\mathrm{DET}(\mathcal{V})$ at the point $\mathcal{V} = \mathcal{T}$. As $t_{ii_2\ldots i_m} = 0$ for all $i_2, \ldots, i_m \in \{1, \ldots, n\}$ by the assumption, $\mathrm{Det}(\mathcal{T}) = 0$ as desired. □

By Proposition 2.4(ii), we have another corollary as follows.

Corollary 2.6. *Let $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$ and $\gamma \in \mathbb{C}$. Then*

$$\mathrm{Det}(\gamma\mathcal{T}) = \gamma^{n(m-1)^{n-1}}\mathrm{Det}(\mathcal{T}).$$

3. Solvability of polynomial equations

Letting matrix $A \in \mathbb{T}(\mathbb{C}^n, 2)$, we know that (Lang, 1987; Strang, 1993):

(i) $\mathrm{Det}(A) = 0$ if and only if $A\mathbf{x} = \mathbf{0}$ has a solution in $\mathbb{C}^n \setminus \{\mathbf{0}\}$, and

(ii) $\mathrm{Det}(A) \neq 0$ if and only if $A\mathbf{x} = \mathbf{b}$ has a unique solution in $\mathbb{C}^n$ for every $\mathbf{b} \in \mathbb{C}^n$.

We generalize such a result to the tensor determinant and the nonlinear polynomial system in this section.

Theorem 3.1. *Let $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$. Then,*

(i) $\operatorname{Det}(\mathcal{T}) = 0$ *if and only if* $\mathcal{T}\mathbf{x}^{m-1} = \mathbf{0}$ *has a solution in* $\mathbb{C}^n \setminus \{\mathbf{0}\}$.
(ii) *If* $\operatorname{Det}(\mathcal{T}) \neq 0$, *then for any* $\mathbf{b} \in \mathbb{C}^n$, $A \in \mathbb{T}(\mathbb{C}^n, 2)$, *and* $\mathcal{B}^j \in \mathbb{T}(\mathbb{C}^n, j)$ *for* $j \in \{3, \dots, m-1\}$, $\mathcal{T}\mathbf{x}^{m-1} = (\mathcal{B}^{m-1})\mathbf{x}^{m-2} + \cdots + (\mathcal{B}^3)\mathbf{x}^2 + A\mathbf{x} + \mathbf{b}$ *has a solution in* $\mathbb{C}^n$.

Proof. (i) It follows from Definitions 1.1 and 1.2 immediately.

(ii) Suppose that $\operatorname{Det}(\mathcal{T}) \neq 0$. For any $\mathbf{b} \in \mathbb{C}^n$, $A \in \mathbb{T}(\mathbb{C}^n, 2)$, and $\mathcal{B}^j \in \mathbb{T}(\mathbb{C}^n, j)$ for $j \in \{3, \dots, m-1\}$, we define tensor $\mathcal{U} \in \mathbb{T}(\mathbb{C}^{n+1}, m)$ as follows:

$$u_{i_1 i_2 \dots i_m} := \begin{cases} t_{i_1 i_2 \dots i_m} & \forall i_j \in \{1, \dots, n\},\ j \in \{1, \dots, m\}, \\ -b_{i_1} & \forall i_1 \in \{1, \dots, n\} \text{ and } i_2 = \cdots = i_m = n+1, \\ -a_{i_1 i_2} & \forall i_1, i_2 \in \{1, \dots, n\} \text{ and } i_3 = \cdots = i_m = n+1, \\ -b^k_{i_1 \dots i_k} & \forall i_1, \dots, i_k \in \{1, \dots, n\} \text{ and } i_{k+1} = \cdots = i_m = n+1, \\ & \forall k \in \{3, \dots, m-1\}, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

Actually, the tensor $\mathcal{U}$ is the tensor corresponding to the homogeneous polynomial in $n+1$ variables by homogenizing $\mathcal{T}\mathbf{x}^{m-1} = (\mathcal{B}^{m-1})\mathbf{x}^{m-2} + \cdots + (\mathcal{B}^3)\mathbf{x}^2 + A\mathbf{x} + \mathbf{b}$. By Corollary 2.5, we have that $\operatorname{Det}(\mathcal{U}) = 0$ since $u_{i_1 i_2 \dots i_m} = 0$ whenever $i_1 = n+1$. Hence, by (i), there exists $\mathbf{y} := (\mathbf{x}^T, \alpha)^T \in \mathbb{C}^{n+1} \setminus \{\mathbf{0}\}$ such that $\mathcal{U}\mathbf{y}^{m-1} = \mathbf{0}$. Consequently, by (3) and the first n equations in $\mathcal{U}\mathbf{y}^{m-1} = \mathbf{0}$, we know that

$$\mathcal{T}\mathbf{x}^{m-1} - \alpha\big(\mathcal{B}^{m-1}\big)\mathbf{x}^{m-2} - \cdots - \alpha^{m-3}\big(\mathcal{B}^3\big)\mathbf{x}^2 - \alpha^{m-2}A\mathbf{x} - \alpha^{m-1}\mathbf{b} = \mathbf{0}. \tag{4}$$

Furthermore, we claim that $\alpha \neq 0$. Otherwise, from (4), $\mathcal{T}\mathbf{x}^{m-1} = \mathbf{0}$ which means $\operatorname{Det}(\mathcal{T}) = 0$ by (i). It is a contradiction. Hence, from (4) we know that $\frac{\mathbf{x}}{\alpha}$ is a solution to

$$\mathcal{T}\mathbf{x}^{m-1} = \big(\mathcal{B}^{m-1}\big)\mathbf{x}^{m-2} + \cdots + \big(\mathcal{B}^3\big)\mathbf{x}^2 + A\mathbf{x} + \mathbf{b}.$$

The proof is complete. □

So, like the matrix determinants of linear equations, the tensor determinants are criteria for the solvability of nonlinear polynomial equations. It is interesting to investigate whether $\mathcal{T}\mathbf{x}^{m-1} = (\mathcal{B}^{m-1})\mathbf{x}^{m-2} + \cdots + (\mathcal{B}^3)\mathbf{x}^2 + A\mathbf{x} + \mathbf{b}$ has only finitely many solutions whenever $\operatorname{Det}(\mathcal{T}) \neq 0$.

4. Block tensors

In the content of matrices, if a square matrix A can be partitioned as

$$A = \begin{pmatrix} B & D \\ 0 & C \end{pmatrix}$$

with square sub-matrices B and C, and sub-matrix D, then $\operatorname{Det}(A) = \operatorname{Det}(B)\operatorname{Det}(C)$ (Lang, 1987; Strang, 1993). We now generalize this property to tensors. The following definition is straightforward.

Definition 4.1. Let $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$ and $1 \leqslant k \leqslant n$. Tensor $\mathcal{U} \in \mathbb{T}(\mathbb{C}^k, m)$ is called a *sub-tensor* of $\mathcal{T}$ associated to the index set $\{j_1, \dots, j_k\} \subseteq \{1, \dots, n\}$ if and only if $u_{i_1 \dots i_m} = t_{j_{i_1} \dots j_{i_m}}$ for all $i_1, \dots, i_m \in \{1, \dots, k\}$.

Though Definition 4.1 is for a specific tensor, the generalization to tensors of indeterminate variables is straightforward. Given a set $\mathbb{E} \subseteq \mathbb{C}^n$, we denote by $\mathbb{I}(\mathbb{E}) \subseteq \mathbb{C}[\mathbf{x}]$ the ideal of polynomials in $\mathbb{C}[\mathbf{x}]$ which vanish identically on $\mathbb{E}$. Given a set of polynomials $\mathbb{F} := \{f_1, \dots, f_s\colon\ f_i \in \mathbb{C}[\mathbf{x}]\}$, we denote by $\mathbb{V}(\mathbb{F}) \subseteq \mathbb{C}^n$ the algebraic set associated to $\mathbb{F}$, i.e., the set of the common roots of polynomials in $\mathbb{F}$ (Cox et al., 1998; Lang, 2002).

Theorem 4.2. *Let $\mathcal{T}$ be an m-th order n-dimensional tensor of indeterminate variables such that there exists an integer $k \in \{1, \dots, n-1\}$ satisfying $t_{ii_2\dots i_m} \equiv 0$ for every $i \in \{k+1, \dots, n\}$ and all indices $i_2, \dots, i_m$ such that $\{i_2, \dots, i_m\} \cap \{1, \dots, k\} \neq \emptyset$. Denote by $\mathcal{U}$ and $\mathcal{V}$ the* sub-tensors *of $\mathcal{T}$ associated to $\{1, \dots, k\}$ and $\{k+1, \dots, n\}$, respectively. Then, it holds that*

$$\mathrm{DET}(\mathcal{T}) = \left[\mathrm{DET}(\mathcal{U})\right]^{(m-1)^{n-k}} \left[\mathrm{DET}(\mathcal{V})\right]^{(m-1)^k}. \tag{5}$$

A word on the notation is necessary before the proof. Though with the same notation, implicitly, $\mathrm{DET}(\mathcal{T})$ is understood as the determinant for m-th order n-dimensional tensors, $\mathrm{DET}(\mathcal{U})$ for m-th order k-dimensional tensors and $\mathrm{DET}(\mathcal{V})$ for m-th order $(n-k)$-dimensional tensors. The actual meanings are clear from the content. The notation Det is similar.

Proof of Theorem 4.2. We first show that for any specific tensor $\mathcal{T}$ satisfying the hypothesis,

$$\mathrm{Det}(\mathcal{T}) = 0 \quad \Longleftrightarrow \quad \mathrm{Det}(\mathcal{U})\mathrm{Det}(\mathcal{V}) = 0. \tag{6}$$

Suppose that $\mathrm{Det}(\mathcal{T}) = 0$. Then there exists $\mathbf{x} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$ such that $\mathcal{T}\mathbf{x}^{m-1} = \mathbf{0}$ by Theorem 3.1(i). Denote by $\mathbf{u} \in \mathbb{C}^k$ the vector consisting of $x_1, \dots, x_k$, and $\mathbf{v} \in \mathbb{C}^{n-k}$ the vector consisting of $x_{k+1}, \dots, x_n$. If $\mathbf{v} \neq \mathbf{0}$, then from $\mathcal{T}\mathbf{x}^{m-1} = \mathbf{0}$ we get that $\mathcal{V}\mathbf{v}^{m-1} = \mathbf{0}$. Consequently, $\mathrm{Det}(\mathcal{V}) = 0$ by Theorem 3.1(i). Otherwise, $\mathbf{u} \neq \mathbf{0}$ and $\mathbf{v} = \mathbf{0}$. This, together with $\mathcal{T}\mathbf{x}^{m-1} = \mathbf{0}$, implies that $\mathcal{U}\mathbf{u}^{m-1} = \mathbf{0}$. Thus, $\mathrm{Det}(\mathcal{U}) = 0$ by Theorem 3.1(i). Hence, we have

$$\mathrm{Det}(\mathcal{T}) = 0 \quad \Longrightarrow \quad \mathrm{Det}(\mathcal{U})\mathrm{Det}(\mathcal{V}) = 0.$$

Conversely, suppose that $\mathrm{Det}(\mathcal{U})\mathrm{Det}(\mathcal{V}) = 0$. If $\mathrm{Det}(\mathcal{U}) = 0$, then there exists $\mathbf{u} \in \mathbb{C}^k \setminus \{\mathbf{0}\}$ such that $\mathcal{U}\mathbf{u}^{m-1} = \mathbf{0}$. Denote $\mathbf{x} := (\mathbf{u}^T, \mathbf{0})^T \in \mathbb{C}^n \setminus \{\mathbf{0}\}$, then $\mathcal{T}\mathbf{x}^{m-1} = \mathbf{0}$, which implies $\mathrm{Det}(\mathcal{T}) = 0$ by Theorem 3.1(i). If $\mathrm{Det}(\mathcal{U}) \neq 0$, then $\mathrm{Det}(\mathcal{V}) = 0$, which implies that there exists $\mathbf{v} \in \mathbb{C}^{n-k} \setminus \{\mathbf{0}\}$ such that $\mathcal{V}\mathbf{v}^{m-1} = \mathbf{0}$. Now, by the vector $\mathbf{v}$ and the tensor $\mathcal{T}$, we construct the vector $\mathbf{b} \in \mathbb{C}^k$ as

$$b_i := \sum_{j_2, \dots, j_m = k+1}^{n} t_{ij_2\dots j_m} v_{j_2-k} \dots v_{j_m-k}, \quad \forall i \in \{1, \dots, k\}; \tag{7}$$

the matrix $A \in \mathbb{T}(\mathbb{C}^k, 2)$ as

$$a_{ij} := \sum_{(q_2, \dots, q_m) \in \mathbb{D}(j)} t_{iq_2\dots q_m} \prod_{q_w > k} v_{q_w - k}, \quad \forall i, j \in \{1, \dots, k\} \tag{8}$$

with $\mathbb{D}(j) := \{(q_2, \dots, q_m) \mid j = q_p \text{ for some } p \in \{2, \dots, m\},\ \text{and } q_l \in \{k+1, \dots, n\},\ l \neq p\}$; and, the tensors $\mathcal{B}^s \in \mathbb{T}(\mathbb{C}^k, s)$ for $s \in \{3, \dots, m-1\}$ as

$$b^s_{ij_2\dots j_s} := \sum_{(q_2, \dots, q_m) \in \mathbb{D}^s(j_2, \dots, j_s)} t_{iq_2\dots q_m} \prod_{q_w > k} v_{q_w - k}, \quad \forall i, j_2, \dots, j_s \in \{1, \dots, k\} \tag{9}$$

with

$$\mathbb{D}^s(j_2, \dots, j_s) := \big\{(q_2, \dots, q_m) \bigm| \{q_{t_2}, \dots, q_{t_s}\} = \{j_2, \dots, j_s\} \text{ for some pairwise}$$
$$\text{different } t_2, \dots, t_s \text{ in } \{2, \dots, m\},\ \text{and } q_l \in \{k+1, \dots, n\},\ l \notin \{t_2, \dots, t_s\}\big\}.$$

Since $\mathrm{Det}(\mathcal{U}) \neq 0$, by Theorem 3.1(ii),

$$\mathcal{U}\mathbf{u}^{m-1} + \left(\mathcal{B}^{m-1}\right)\mathbf{u}^{m-2} + \dots + \left(\mathcal{B}^3\right)\mathbf{u}^2 + A\mathbf{u} + \mathbf{b} = \mathbf{0}$$

has a solution $\mathbf{u} \in \mathbb{C}^k$. Let $\mathbf{x} := (\mathbf{u}^T, \mathbf{v}^T)^T \in \mathbb{C}^n \setminus \{\mathbf{0}\}$ as $\mathbf{v} \in \mathbb{C}^{n-k} \setminus \{\mathbf{0}\}$. By (7), (8) and (9), we have that

$$\left(\mathcal{T}\mathbf{x}^{m-1}\right)_i = \left(\mathcal{U}\mathbf{u}^{m-1} + \left(\mathcal{B}^{m-1}\right)\mathbf{u}^{m-2} + \cdots + \left(\mathcal{B}^3\right)\mathbf{u}^2 + A\mathbf{u} + \mathbf{b}\right)_i = 0, \quad \forall i \in \{1, \ldots, k\}.$$

Furthermore,

$$\left(\mathcal{T}\mathbf{x}^{m-1}\right)_i = \left(\mathcal{V}\mathbf{v}^{m-1}\right)_i = 0, \quad \forall i \in \{k+1, \ldots, n\}.$$

Consequently, $\mathcal{T}\mathbf{x}^{m-1} = \mathbf{0}$ which implies $\mathrm{Det}(\mathcal{T}) = 0$ by Theorem 3.1(i).

Hence, we proved (6). In the following, we show that (5) holds. Note that the dimension of $\mathbb{T}(\mathbb{C}^n, m)$ is n^m. The set of tensors satisfying the hypothesis of this theorem forms a vector subspace $\mathbb{S}$ of $\mathbb{T}(\mathbb{C}^n, m)$ with dimension $kn^{m-1} + (n-k)^{m-1}$. Consequently, the number of variables of the polynomial $\mathrm{DET}(\mathcal{T})$ is $kn^{m-1} + (n-k)^{m-1}$. In the following, the ambient space for the algebraic sets is understood as $\mathbb{S}$. As sets of variables, the sets of entries of $\mathcal{U}$ and $\mathcal{V}$ are subsets of the set of entries of $\mathcal{T}$. Hence, we can view $\mathrm{DET}(\mathcal{U}), \mathrm{DET}(\mathcal{V}) \in \mathbb{C}[\mathcal{T}]$. By (6), we have

$$\mathbb{V}\left(\mathrm{DET}(\mathcal{U})\mathrm{DET}(\mathcal{V})\right) = \mathbb{V}\left(\mathrm{DET}(\mathcal{T})\right),$$

which implies that

$$\mathbb{I}\left(\mathbb{V}\left(\mathrm{DET}(\mathcal{T})\right)\right) = \mathbb{I}\left(\mathbb{V}\left(\mathrm{DET}(\mathcal{U})\mathrm{DET}(\mathcal{V})\right)\right).$$

By Proposition 2.4(ii), both $\mathrm{DET}(\mathcal{U}) \in \mathbb{C}[\mathcal{U}]$ and $\mathrm{DET}(\mathcal{V}) \in \mathbb{C}[\mathcal{V}]$ are irreducible. Consequently,

$$\mathbb{I}\left(\mathbb{V}\left(\mathrm{DET}(\mathcal{T})\right)\right) = \mathbb{I}\left(\mathbb{V}\left(\mathrm{DET}(\mathcal{U})\mathrm{DET}(\mathcal{V})\right)\right) = \left\langle \mathrm{DET}(\mathcal{U})\mathrm{DET}(\mathcal{V})\right\rangle.$$

Let $\sqrt{\langle \mathrm{DET}(\mathcal{T})\rangle}$ be the radical ideal of the ideal $\langle \mathrm{DET}(\mathcal{T})\rangle$ (Lang, 2002). Then, Hilbert's Nullstellensatz (see Theorem 4.2 in Cox et al. (2006)) implies that

$$\sqrt{\left\langle \mathrm{DET}(\mathcal{T})\right\rangle} = \mathbb{I}\left(\mathbb{V}\left(\mathrm{DET}(\mathcal{T})\right)\right) = \left\langle \mathrm{DET}(\mathcal{U})\mathrm{DET}(\mathcal{V})\right\rangle.$$

Since both $\sqrt{\langle \mathrm{DET}(\mathcal{T})\rangle}$ and $\langle \mathrm{DET}(\mathcal{U})\mathrm{DET}(\mathcal{V})\rangle$ are principal ideals and $\mathbb{C}[\mathcal{T}]$ is a unique factorization domain, we have that

$$\mathrm{DET}(\mathcal{T}) = \left(\mathrm{DET}(\mathcal{U})\right)^{r_1}\left(\mathrm{DET}(\mathcal{V})\right)^{r_2} \tag{10}$$

for some $r_1, r_2 \in \mathbb{N}_+$.

By Proposition 2.4(i), $\mathrm{DET}(\mathcal{T})$ is homogeneous of degree $(m-1)^{n-1}$ in the variables $\{t_{1i_2\ldots i_m} \mid i_2, \ldots, i_m \in \{1, \ldots, n\}\}$. By the hypothesis, $\mathrm{DET}(\mathcal{V})$ is independent of the variables $\{t_{1i_2\ldots i_m} \mid i_2, \ldots, i_m \in \{1, \ldots, n\}\}$. By Proposition 2.4(i) again, $\mathrm{DET}(\mathcal{U})$ is homogeneous of degree $(m-1)^{k-1}$ in the variables $\{t_{1i_2\ldots i_m} \mid i_2, \ldots, i_m \in \{1, \ldots, k\}\}$. Consequently, $r_1 = (m-1)^{n-k}$ by (10). Now, comparing the degrees of the both sides of (10) with Proposition 2.4(ii), we get $r_2 = (m-1)^k$ and hence (5). The proof is complete. □

5. A simple application: Triangular tensors

Let $\mathcal{T} = (t_{i_1\ldots i_m}) \in \mathbb{T}(\mathbb{C}^n, m)$. If $t_{i_1\ldots i_m} \equiv 0$ whenever $\min\{i_2, \ldots, i_m\} < i_1$, then $\mathcal{T}$ is called an upper triangular tensor. If $t_{i_1\ldots i_m} \equiv 0$ whenever $\max\{i_2, \ldots, i_m\} > i_1$, then $\mathcal{T}$ is called a lower triangular tensor. If $\mathcal{T}$ is either upper or lower triangular, then $\mathcal{T}$ is called a triangular tensor. In particular, a diagonal tensor is a triangular tensor.

Theorem 4.2 presents the determinant formula for "upper block triangular" tensors. By similar proof of Theorem 4.2, it can be proved that the result holds true for lower block triangular tensors as well. Consequently, we have the following proposition.

Proposition 5.1. *Suppose that $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$ is a triangular tensor. Then*

$$\mathrm{Det}(\mathcal{T}) = \prod_{i=1}^{n} (t_{i\ldots i})^{(m-1)^{n-1}}.$$

It is a generalization of Section 3.1.5 in Dolotin and Morozov (2007).

Corollary 5.2. *Suppose that $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$ is a triangular tensor. Then*

$$\sigma(\mathcal{T}) = \{t_{i\dots i} \mid i \in \{1, \dots, n\}\},$$

and the algebraic multiplicity of $t_{i\dots i}$ is $(m-1)^{n-1}$ for all $i \in \{1, \dots, n\}$.

Proof. By Definition 2.2, $\chi(\lambda)$ is defined as the determinant of tensor $\lambda\mathcal{I} - \mathcal{T}$. Since $\mathcal{T}$ is a triangular tensor, $\lambda\mathcal{I} - \mathcal{T}$ is a triangular tensor as well. The diagonal elements of $\lambda\mathcal{I} - \mathcal{T}$ are $\{\lambda - t_{i\dots i} \mid i \in \{1, \dots, n\}\}$. Consequently, $\chi(\lambda) = \mathrm{Det}(\lambda\mathcal{I} - \mathcal{T}) = \prod_{i=1}^n (\lambda - t_{i\dots i})^{(m-1)^{n-1}}$ by Proposition 5.1. By Theorem 2.3, $\sigma(\mathcal{T})$ consists of the roots of the characteristic polynomial $\chi(\lambda)$ of the tensor $\mathcal{T}$. Then, the proof is complete. □

With Theorem 3.1, we have the following simple application of the determinant theory.

Theorem 5.3. *Suppose that $\mathcal{T}$ is a triangular tensor with nonzero diagonal elements. Then for any $\mathbf{b} \in \mathbb{C}^n$, $A \in \mathbb{T}(\mathbb{C}^n, 2)$, and $\mathcal{B}^j \in \mathbb{T}(\mathbb{C}^n, j)$ for $j \in \{3, \dots, m-1\}$, $\mathcal{T}\mathbf{x}^{m-1} = (\mathcal{B}^{m-1})\mathbf{x}^{m-2} + \dots + (\mathcal{B}^3)\mathbf{x}^2 + A\mathbf{x} + \mathbf{b}$ has a solution in $\mathbb{C}^n$.*

6. The characteristic polynomial

By Definition 2.2, for any $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$, its characteristic polynomial is $\chi(\lambda) = \mathrm{Det}(\lambda\mathcal{I} - \mathcal{T})$. In this section, we discuss some properties of the characteristic polynomial of a tensor related to the determinant. To this end, we give a trace formula for the determinant in Subsection 6.1 first. This result, due to Morozov and Shakirov (2011), is a corner stone for the subsequent analysis of the characteristic polynomials.

6.1. A trace formula of the determinant

Let $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$. Define the following differential operators:

$$\hat{g}_i := \sum_{i_2=1}^n \dots \sum_{i_m=1}^n t_{ii_2\dots i_m} \frac{\partial}{\partial a_{ii_2}} \dots \frac{\partial}{\partial a_{ii_m}}, \quad \forall i \in \{1, \dots, n\}, \tag{11}$$

where A is an auxiliary $n \times n$ matrix consists of indeterminate variables a_{ij}'s. It is clear that for every i, $\hat{g}_i$ is a differential operator which belongs to the operator algebra $\mathbb{C}[\partial A]$, here ∂A is the $n \times n$ matrix with elements $\frac{\partial}{\partial a_{ij}}$'s. The Schur polynomials are defined as:

$$p_0(t_0) = 1, \quad \text{and} \quad p_k(t_1, \dots, t_k) := \sum_{i=1}^k \sum_{d_j > 0,\, \sum_{j=1}^i d_j = k} \frac{\prod_{j=1}^i t_{d_j}}{i!}, \quad \forall k \geqslant 1, \tag{12}$$

where $\{t_0, t_1, \dots\}$ are variables. Motivated by Cooper and Dutle (2012) and Morozov and Shakirov (2011), we define the *d-th order trace* of the tensor $\mathcal{T}$ as

$$\mathrm{Tr}_d(\mathcal{T}) := (m-1)^{n-1} \left[\sum_{\sum_{i=1}^n k_i = d} \prod_{i=1}^n \frac{(\hat{g}_i)^{k_i}}{((m-1)k_i)!} \right] \mathrm{Tr}\left(A^{(m-1)d}\right). \tag{13}$$

We show in Proposition 6.4 that $\mathrm{Tr}_1(\mathcal{T}) = (m-1)^{n-1} \sum_{i=1}^n t_{i\dots i}$. Hence, it is a generalization of the trace of a matrix. $\mathrm{Tr}_d(\mathcal{T})$'s are called higher order traces for $d > 1$.

We now have the following proposition.

Proposition 6.1. *Let $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$ and the notation be defined as above. Then,*

$$\mathrm{DET}(\mathcal{T}) = 1 + \sum_{k=1}^{\infty} p_k\left(-\frac{\mathrm{Tr}_1(\mathcal{I}-\mathcal{T})}{1}, \ldots, -\frac{\mathrm{Tr}_k(\mathcal{I}-\mathcal{T})}{k}\right). \tag{14}$$

Proof. This result follows from Proposition II in Morozov and Shakirov (2011), the identity $\log(\mathrm{DET}(I - A)) = \mathrm{Tr}(\log(I - A))$ for the matrix A, and the definitions of the Schur polynomials and the higher order traces. As the proof is a restatement of those from Sections 4–8 in Morozov and Shakirov (2011) in the language of tensors, we omit it. □

The following proposition is useful in the sequel, which also helps to give an expression of $\mathrm{DET}(\mathcal{T})$ with only finitely many terms.

Proposition 6.2. *Let $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$ and the notation be defined as above. Then, the following hold*:

(i) *for every $d \in \mathbb{N}_+$, $\mathrm{Tr}_d(\mathcal{T}) \in \mathbb{C}[\mathcal{T}]$ is homogeneous of degree d*;
(ii) *for every $k \in \mathbb{N}_+$, $p_k(-\frac{\mathrm{Tr}_1(\mathcal{T})}{1}, \ldots, -\frac{\mathrm{Tr}_k(\mathcal{T})}{k}) \in \mathbb{C}[\mathcal{T}]$ is homogeneous of degree k; and,*
(iii) *for any integer $k > n(m-1)^{n-1}$, $p_k(-\frac{\mathrm{Tr}_1(\mathcal{T})}{1}, \ldots, -\frac{\mathrm{Tr}_k(\mathcal{T})}{k}) \in \mathbb{C}[\mathcal{T}]$ is zero.*

Proof. (i) By the formulae of $\hat{g}_i$'s as in (11), it is easy to see that

$$\sum_{\sum_{i=1}^n k_i = d} \prod_{i=1}^n \frac{(\hat{g}_i)^{k_i}}{((m-1)k_i)!} \in \mathbb{C}[\mathcal{T}, \partial A]$$

is homogeneous, and more explicitly, homogeneous of degree d in the variable $\mathcal{T}$ and homogeneous of degree $(m-1)d$ in the variable ∂A. It is also known that

$$\mathrm{Tr}\big(A^k\big) = \sum_{i_1=1}^n \ldots \sum_{i_k=1}^n a_{i_1 i_2} a_{i_2 i_3} \ldots a_{i_{k-1} i_k} a_{i_k i_1} \in \mathbb{C}[A] \tag{15}$$

is homogeneous of degree k. These, together with (13), imply that $\mathrm{Tr}_d(\mathcal{T}) \in \mathbb{C}[\mathcal{T}]$ is homogeneous of degree d as desired.

(ii) It follows from (i) and the definitions of the Schur polynomials as in (12) directly.

(iii) From Proposition 2.4(ii), it is clear that $\mathrm{DET}(\mathcal{B})$ is an irreducible polynomial which is homogeneous of degree $n(m-1)^{n-1}$ in the variables $\{b_{i_1 \ldots i_m}\}$. In the following, let $\mathcal{B} := \mathcal{I} - \mathcal{T}$. Since the entries of $\mathcal{B}$ consist of 1 and the entries of the tensor $\mathcal{T}$, the highest degree of $\mathrm{DET}(\mathcal{I} - \mathcal{T})$ viewed as a polynomial in $\mathbb{C}[\mathcal{T}]$ is not greater than $n(m-1)^{n-1}$. This, together with (14) and (ii) which asserts that $p_k(-\frac{\mathrm{Tr}_1(\mathcal{T})}{1}, \ldots, -\frac{\mathrm{Tr}_k(\mathcal{T})}{k}) \in \mathbb{C}[\mathcal{T}]$ is homogeneous of degree k, implies the result (iii).

The proof is complete. □

By Proposition 6.2 and (14), we immediately get

$$\mathrm{DET}(\mathcal{T}) = 1 + \sum_{k=1}^{n(m-1)^{n-1}} p_k\left(-\frac{\mathrm{Tr}_1(\mathcal{I}-\mathcal{T})}{1}, \ldots, -\frac{\mathrm{Tr}_k(\mathcal{I}-\mathcal{T})}{k}\right). \tag{16}$$

This is a trace formula for the tensor determinant. It provides a way to approach the computation of the tensor determinant. However, it involves the higher order traces of tensors, and hence the differential operators $\hat{g}_i$'s. It is very hard to compute them (Dolotin and Morozov, 2007; Morozov and Shakirov, 2011). In Section 7, we give an explicit formulae for the second order trace of a tensor of arbitrary dimension and the determinant of a tensor when $n = 2$.

6.2. Basic properties of the characteristic polynomial

Some basic properties of the characteristic polynomial are derived in this subsection.

Theorem 6.3. *Let $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$ and the notation be defined as above. Then*

$$\begin{aligned}\chi(\lambda) &= \mathrm{Det}(\lambda\mathcal{I} - \mathcal{T}) \\ &= \lambda^{n(m-1)^{n-1}} + \sum_{k=1}^{n(m-1)^{n-1}} \lambda^{n(m-1)^{n-1}-k} p_k\left(-\frac{\mathrm{Tr}_1(\mathcal{T})}{1}, \ldots, -\frac{\mathrm{Tr}_k(\mathcal{T})}{k}\right) \\ &= \prod_{\lambda_i \in \sigma(\mathcal{T})} (\lambda - \lambda_i)^{m_i},\end{aligned}$$

where m_i is the algebraic multiplicity of the eigenvalue λ_i.

Proof. The first equality follows from Definition 2.2, and the last one from Theorem 2.3.

By Proposition 6.2 and (16), we can get that

$$\chi(1) = \mathrm{Det}(\mathcal{I} - \mathcal{T}) = 1 + \sum_{k=1}^{n(m-1)^{n-1}} p_k\left(-\frac{\mathrm{Tr}_1(\mathcal{T})}{1}, \ldots, -\frac{\mathrm{Tr}_k(\mathcal{T})}{k}\right). \tag{17}$$

Consequently, when $\lambda \neq 0$,

$$\begin{aligned}\chi(\lambda) &= \mathrm{Det}(\lambda\mathcal{I} - \mathcal{T}) \\ &= \lambda^{n(m-1)^{n-1}} \mathrm{Det}\left(\mathcal{I} - \frac{\mathcal{T}}{\lambda}\right) \\ &= \lambda^{n(m-1)^{n-1}} \left[1 + \sum_{k=1}^{n(m-1)^{n-1}} p_k\left(-\frac{\mathrm{Tr}_1(\frac{\mathcal{T}}{\lambda})}{1}, \ldots, -\frac{\mathrm{Tr}_k(\frac{\mathcal{T}}{\lambda})}{k}\right)\right] \\ &= \lambda^{n(m-1)^{n-1}} \left[1 + \sum_{k=1}^{n(m-1)^{n-1}} \frac{1}{\lambda^k} p_k\left(-\frac{\mathrm{Tr}_1(\mathcal{T})}{1}, \ldots, -\frac{\mathrm{Tr}_k(\mathcal{T})}{k}\right)\right] \\ &= \lambda^{n(m-1)^{n-1}} + \sum_{k=1}^{n(m-1)^{n-1}} \lambda^{n(m-1)^{n-1}-k} p_k\left(-\frac{\mathrm{Tr}_1(\mathcal{T})}{1}, \ldots, -\frac{\mathrm{Tr}_k(\mathcal{T})}{k}\right).\end{aligned}$$

Here the second equality comes from Corollary 2.6, the third from (17), and the fourth from Proposition 6.2. Hence, the result follows from the fact that the field $\mathbb{C}$ is of characteristic zero. The proof is complete. □

Theorem 6.3 gives a trace formula for the characteristic polynomial of the tensor $\mathcal{T}$ as well as an eigenvalue representation for it.

Here are some properties concerning the coefficients of $\chi(\lambda)$.

Proposition 6.4. *Let $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$ and the notation be defined as above. Then,*

(i) $p_1(-\mathrm{Tr}_1(\mathcal{T})) = -\mathrm{Tr}_1(\mathcal{T}) = -(m-1)^{n-1}\sum_{i=1}^n t_{ii\ldots i}$,
(ii) $p_2(-\frac{\mathrm{Tr}_1(\mathcal{T})}{1}, -\frac{\mathrm{Tr}_2(\mathcal{T})}{2}) = \frac{1}{2}([\mathrm{Tr}_1(\mathcal{T})]^2 - \mathrm{Tr}_2(\mathcal{T}))$, *and*
(iii) $p_{n(m-1)^{n-1}}(-\frac{\mathrm{Tr}_1(\mathcal{T})}{1}, \ldots, -\frac{\mathrm{Tr}_{n(m-1)^{n-1}}(\mathcal{T})}{n(m-1)^{n-1}}) = (-1)^{n(m-1)^{n-1}}\mathrm{Det}(\mathcal{T})$.

Proof. (i) By (12), we know that $p_1(-\mathrm{Tr}_1(\mathcal{T})) = -\mathrm{Tr}_1(\mathcal{T})$. Furthermore, by (13), it is easy to see that

$$\begin{aligned}
\mathrm{Tr}_1(\mathcal{T}) &= (m-1)^{n-1}\sum_{i=1}^{n}\frac{\hat{g}_i}{(m-1)!}\mathrm{Tr}\big(A^{m-1}\big)\\
&= \frac{(m-1)^{n-1}}{(m-1)!}\sum_{i=1}^{n}\left[\sum_{i_2=1}^{n}\cdots\sum_{i_m=1}^{n}t_{ii_2\ldots i_m}\frac{\partial}{\partial a_{ii_2}}\cdots\frac{\partial}{\partial a_{ii_m}}\right]\mathrm{Tr}\big(A^{m-1}\big)\\
&= \frac{(m-1)^{n-1}}{(m-1)!}\sum_{i=1}^{n}\left[\sum_{i_2=1}^{n}\cdots\sum_{i_m=1}^{n}t_{ii_2\ldots i_m}\frac{\partial}{\partial a_{ii_2}}\cdots\frac{\partial}{\partial a_{ii_m}}\right]\\
&\quad\cdot\left(\sum_{i_1=1}^{n}\cdots\sum_{i_{m-1}=1}^{n}a_{i_1i_2}a_{i_2i_3}\ldots a_{i_{m-2}i_{m-1}}a_{i_{m-1}i_1}\right)\\
&= \frac{(m-1)^{n-1}}{(m-1)!}\sum_{i=1}^{n}\left[t_{ii\ldots i}\frac{\partial}{\partial a_{ii}}\cdots\frac{\partial}{\partial a_{ii}}(a_{ii})^{m-1}\right]\\
&= (m-1)^{n-1}\sum_{i=1}^{n}t_{ii\ldots i}.
\end{aligned}$$

Here, the fourth equality follows from the fact that: (a) the differential operator in the right hand side of the third equality contains only items $\frac{\partial}{\partial a_{i\star}}$'s for $\star\in\{1,\ldots,n\}$ and the total degree is $m-1$, and (b) only terms in $\mathrm{Tr}(A^{m-1})$ that contain the same $\frac{\partial}{\partial a_{i\star}}$'s of total degree $m-1$ can contribute to the result and this case occurs only when every $\star = i$ by (15). Consequently, the result (i) follows.

(ii) follows from the definition (12) by direct calculation.

(iii) By Theorem 6.3, it is clear that

$$\chi(0) = \mathrm{Det}(-\mathcal{T}) = p_{n(m-1)^{n-1}}\left(-\frac{\mathrm{Tr}_1(\mathcal{T})}{1},\ldots,-\frac{\mathrm{Tr}_{n(m-1)^{n-1}}(\mathcal{T})}{n(m-1)^{n-1}}\right).$$

Moreover, $\mathrm{DET}(-\mathcal{T})\in\mathbb{C}[\mathcal{T}]$ is homogeneous of degree $n(m-1)^{n-1}$ by Proposition 2.4(ii), which implies that $\mathrm{Det}(-\mathcal{T}) = (-1)^{n(m-1)^{n-1}}\mathrm{Det}(\mathcal{T})$. Consequently, the result follows. □

Corollary 6.5. *Let $\mathcal{T}\in\mathbb{T}(\mathbb{C}^n, m)$ and the notation be defined as above. Then,*

(i) $\sum_{\lambda_i\in\sigma(\mathcal{T})} m_i\lambda_i = (m-1)^{n-1}\sum_{i=1}^{n} t_{ii\ldots i} = \mathrm{Tr}_1(\mathcal{T})$,
(ii) $\sum_{\lambda_i\in\sigma(\mathcal{T})} m_i\lambda_i^2 = \mathrm{Tr}_2(\mathcal{T})$, *and*
(iii) $\prod_{\lambda_i\in\sigma(\mathcal{T})}\lambda_i^{m_i} = \mathrm{Det}(\mathcal{T})$.

Here m_i is the algebraic multiplicity of the eigenvalue λ_i.

Proof. The results (i) and (iii) follow from the eigenvalue representation of $\chi(\lambda)$ in Theorem 6.3 and the coefficients of $\chi(\lambda)$ in Proposition 6.4 immediately. For (ii), by Proposition 6.4(ii) and Newton's identities for the roots and the coefficients of a polynomial, we get that $\sum_{i<j,\lambda_i,\lambda_j\in\sigma(\mathcal{T})} m_im_j\lambda_i\lambda_j = p_2(-\frac{\mathrm{Tr}_1(\mathcal{T})}{1}, -\frac{\mathrm{Tr}_2(\mathcal{T})}{2}) = \frac{1}{2}([\mathrm{Tr}_1(\mathcal{T})]^2 - \mathrm{Tr}_2(\mathcal{T}))$. Consequently, (ii) follows from (i) and the perfect square formula. □

Remark 6.6. In Qi (2005), Qi proved the results in Corollary 6.5(i) and (iii) for $\mathcal{T}\in\mathbb{S}(\mathbb{R}^n, m)$ (the space of real symmetric tensors of order m and dimension n). By Theorem 3.1 and Corollary 6.5, we see that the solvability of a homogeneous polynomial equation is characterized by the zero eigenvalue of the underlying tensor.

In the following, we generalize Corollary 6.5(i) and (ii) to $\mathrm{Tr}_k(\mathcal{T})$ for all $k \in \{1, \ldots, n(m-1)^{n-1}\}$. To this end, we need the following lemmas.

Lemma 6.7. *Let $p_k(t_1, \ldots, t_k)$ be the Schur polynomials defined as (12). Then, for all $k \in \mathbb{N}_+$,*

$$\frac{\partial}{\partial t_i} p_k = p_{k-i}, \quad \forall i \in \{1, \ldots, k\}. \tag{18}$$

Proof. The case for $i = k$ is easy to see, since $p_0 = 1$ and the only monomial in p_k having the variable t_k is t_k by (12).

Now, we show the cases $i \in \{1, \ldots, k-1\}$. For each fixed i, we have that $p_{k-i}(t_1, \ldots, t_{k-i}) = \sum_{s=1}^{k-i} \sum_{d_j>0, \sum_{j=1}^{s} d_j = k-i} \frac{\prod_{j=1}^{s} t_{d_j}}{s!}$ by (12). To prove (18), it is sufficient to show that there is a one-to-one correspondence between the monomials in p_{k-i} and these in $p_k(t_1, \ldots, t_k) = \sum_{w=1}^{k} \sum_{d_j>0, \sum_{j=1}^{w} d_j = k} \frac{\prod_{j=1}^{w} t_{d_j}}{w!}$ having variable t_i, and their coefficients satisfying the derivative relation.

First, the one-to-one correspondence between the monomials in p_{k-i} and these in p_k having variable t_i is obvious: for any monomial $c \prod_{j=1}^{s} t_{d_j}$ with nonzero coefficient c in p_{k-i}, there is the monomial $d t_i \prod_{j=1}^{s} t_{d_j}$ with nonzero coefficient d in p_k, and vice verse.

Second, suppose that $\frac{c}{s!} \prod_{j=1}^{s} t_{d_j}$ with nonzero coefficient c is a monomial in the polynomial p_{k-i} for some $s \in \{1, \ldots, k-i\}$ and $d_1, \ldots, d_s$. Then, by (12), we see that the number of cases of the ordered s-tuples $(q_1, \ldots, q_s)$ such that $\sum_{j=1}^{s} q_j = k - i$ and $q_j > 0$, $j \in \{1, \ldots, s\}$ resulting in $\prod_{j=1}^{s} t_{d_j}$ is c. For any such ordered $(q'_1, \ldots, q'_s)$, we get $s+1$ ordered $(s+1)$-tuples $(i, q'_1, \ldots, q'_s), (q'_1, i, \ldots, q'_s) \ldots, (q'_1, \ldots, q'_s, i)$ such that every $(s+1)$-tuple results in $t_i \prod_{j=1}^{s} t_{d_j}$. Note that some of the $(s+1)$-tuples may be the same. Let r be the degree of the variable t_i in the monomial $t_i \prod_{j=1}^{s} t_{d_j}$. Consequently, the number of cases of the ordered $(s+1)$-tuples $(q_1, \ldots, q_s, q_{s+1})$ such that $\sum_{j=1}^{s+1} q_j = k$ and $q_j > 0$, $j \in \{1, \ldots, s+1\}$ resulting in $t_i \prod_{j=1}^{s} t_{d_j}$ is $\frac{(s+1)c}{r}$. These, together with (12), imply that the monomial $\frac{c}{s!} \prod_{j=1}^{s} t_{d_j}$ in p_{k-i} corresponds to the following monomial in p_k:

$$\frac{(s+1)c}{r} \frac{1}{(s+1)!} t_i \prod_{j=1}^{s} t_{d_j}.$$

The derivative of this monomial with respective to t_i is exactly $\frac{c}{s!} \prod_{j=1}^{s} t_{d_j}$. The proof is complete. □

Lemma 6.8. *Let $\mathbf{x} = (x_1, \ldots, x_n)$ and $h_1, \ldots, h_n \in \mathbb{C}[\mathbf{x}]$ be polynomials. If there is some $f \in \mathbb{C}[\mathbf{x}]$ satisfying the following system of differential equations*

$$\frac{\partial}{\partial x_i} f = h_i, \quad \forall i \in \{1, \ldots, n\}$$

and $f(\mathbf{0}) = 0$, then f is unique.

Proof. First, $f(\mathbf{0}) = 0$ implies that the constant term of f is zero. Then, as $\mathbb{C}$ is algebraically closed and of characteristic zero, it is sufficient to prove that every monomial of positive degree in f is uniquely determined by the differential equations. This is easy to see: (i) every monomial of f containing the variable x_i is uniquely determined by the i-th differential equation in the hypothesis, and (ii) every monomial of positive degree of f has at least one variable in the set $\{x_i \mid i \in \{1, \ldots, n\}\}$. The proof is complete. □

Lemma 6.9. *Let $p_k(t_1, \ldots, t_k)$ be the Schur polynomials defined as (12). Then, for all $k \in \mathbb{N}_+$,*

$$k p_k = k t_k + \sum_{i=0}^{k-1} i t_i p_{k-i}. \tag{19}$$

Proof. On the one side, by Lemmas 6.7 and 6.8, we know that for all $k \in \mathbb{N}_+$, p_k defined as (12) is the unique polynomial satisfying

$$\frac{\partial}{\partial t_i} p_k = p_{k-i}, \quad \forall i \in \{1, \ldots, k\}. \tag{20}$$

On the other side, we show that polynomials $q_k(t_1, \ldots, t_k)$ defined through the recursive formulae

$$q_0 = 1, \qquad kq_k = kt_k + \sum_{i=0}^{k-1} it_i q_{k-i}, \quad \forall k = 1, 2, \ldots \tag{21}$$

satisfy (20) through replacing p_k's by q_k's as well. Consequently, $q_k = p_k$ for all $k \in \mathbb{N}$ and (19) follows.

The proof is by induction, the first step for $k = 1$ is obvious, since $\frac{\partial}{\partial t_1} q_1 = 1 = q_0$. Second, suppose that all of $\{q_1, \ldots, q_k\}$ satisfy (20) for some $k \geqslant 1$, we prove that q_{k+1} satisfies (20) as well. It is easy to see that $\frac{\partial}{\partial t_{k+1}} q_{k+1} = 1 = q_0$ by (21) and the fact that q_s is independent of t_{k+1} for $s \leqslant k$. For $s \in \{1, \ldots, k\}$, by (21)

$$\begin{aligned}
(k+1)\frac{\partial}{\partial t_s} q_{k+1} &= \frac{\partial}{\partial t_s}\left(\sum_{i=0}^{k} it_i q_{k+1-i}\right) \\
&= \sum_{0 \leqslant i \leqslant k,\, i \neq s} it_i \frac{\partial}{\partial t_s} q_{k+1-i} + st_s \frac{\partial}{\partial t_s} q_{k+1-s} + sq_{k+1-s} \\
&= \sum_{i=0}^{k} it_i \frac{\partial}{\partial t_s} q_{k+1-i} + sq_{k+1-s} \\
&= \sum_{i=1}^{k+1-s} it_i \frac{\partial}{\partial t_s} q_{k+1-i} + sq_{k+1-s} \\
&= \sum_{i=1}^{k+1-s} it_i q_{k+1-i-s} + sq_{k+1-s} \\
&= \sum_{i=1}^{k-s} it_i q_{k+1-s-i} + (k+1-s)t_{k+1-s} + sq_{k+1-s} \\
&= (k+1-s)q_{k+1-s} + sq_{k+1-s} \\
&= (k+1)q_{k+1-s},
\end{aligned}$$

where the fourth equality follows from the fact that q_w is independent of t_s for $w \leqslant s-1$, the fifth from the inductive hypothesis, the sixth from the fact that $q_0 = 1$, and the seventh from (21). Therefore, q_{k+1} satisfies (20). Then, q_k satisfies (20) for all $k \in \mathbb{N}_+$ by induction. The proof is complete. □

Now, we are in the position to give the main theorem in this section.

Theorem 6.10. *Let $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$. Denote by p_i the codegree i coefficient of the characteristic polynomial of the tensor $\mathcal{T}$. Then, for all $k \in \{1, \ldots, n(m-1)^{n-1}\}$,*

$$\mathrm{Tr}_k(\mathcal{T}) = -kp_k - \sum_{i=1}^{k-1} p_i \,\mathrm{Tr}_{k-i}(\mathcal{T}).$$

Moreover, for all $k \in \{1, \ldots, n(m-1)^{n-1}\}$,

$$\mathrm{Tr}_k(\mathcal{T}) = \sum_{\lambda_i \in \sigma(\mathcal{T})} m_i \lambda_i^k,$$

where m_i is the algebraic multiplicity of the eigenvalue λ_i.

Proof. The first half of this theorem follows from Theorem 6.3 and Lemma 6.9 by inserting t_i with $-\frac{\mathrm{Tr}_i(\mathcal{T})}{i}$. The second half follows from the first half and Newton's identities on the roots and the coefficients of a polynomial: for a univariate polynomial equation

$$t^k + a_1 t^{k-1} + \cdots + a_k = 0,$$

let s_i be the sum of the i-th powers of its roots with multiplicity. Then,

$$s_i = -i a_i - \sum_{j=1}^{i-1} s_{i-j} a_j.$$

The proof is complete. □

Remark 6.11. Theorem 6.10 reveals two facts: (i) the coefficients of the characteristic polynomial of a tensor are recursively generated by the higher order traces of the tensor, and (ii) the higher order traces of a tensor are elementary symmetric functions of powers of the eigenvalues of the tensor. It is a generalization of Newton's identities on the characteristic polynomial for a matrix to a tensor. It also indicates the fundamental roles of the higher order traces in the eigenvalue theory of tensors.

6.3. *Positive semidefinite tensors*

In this subsection, we present a sufficient and a necessary condition, based on the characteristic polynomial, for an even order tensor to be positive semidefinite.

A tensor $\mathcal{T} \in \mathbb{S}(\mathbb{R}^n, m)$ is called positive semidefinite if and only if $\mathbf{x}^T(\mathcal{T}\mathbf{x}^{m-1}) \geqslant 0$ for all $\mathbf{x} \in \mathbb{R}^n$. Obviously, m being even is necessary for positive semidefinite tensors.

Lemma 6.12. *Let $\mathcal{T} \in \mathbb{S}(\mathbb{R}^n, m)$ and m be even. Then, $\mathcal{T}$ is positive semidefinite if all the real eigenvalues of $\mathcal{T}$ are nonnegative.*

Proof. The result follows from Theorem 5 in Qi (2005). □

The following classical result on real roots of a polynomial is Déscartes' Rule of Signs, see Theorem 1.5 in Sturmfels (2002).

Lemma 6.13. *The number of positive real roots of a polynomial is at most the number of sign changes in its coefficients.*

Let $\mathrm{sgn}(\cdot)$ be the sign function for scalars, i.e., $\mathrm{sgn}(\gamma) = 1$ if $\gamma > 0$, $\mathrm{sgn}(0) = 0$ and $\mathrm{sgn}(\gamma) = -1$ if $\gamma < 0$.

Proposition 6.14. *Let m be even, $\mathcal{T} \in \mathbb{T}(\mathbb{R}^n, m)$, and*

$$\chi(\lambda) = \lambda^{n(m-1)^{n-1}} + \sum_{k=1}^{n(m-1)^{n-1}} \lambda^{n(m-1)^{n-1}-k} p_k\left(-\frac{\mathrm{Tr}_1(\mathcal{T})}{1}, \ldots, -\frac{\mathrm{Tr}_k(\mathcal{T})}{k}\right).$$

If

$$\operatorname{sgn}\left(p_k\left(-\frac{\operatorname{Tr}_1(\mathcal{T})}{1},\ldots,-\frac{\operatorname{Tr}_k(\mathcal{T})}{k}\right)\right)=(-1)^k \tag{22}$$

for all $p_k(-\frac{\operatorname{Tr}_1(\mathcal{T})}{1},\ldots,-\frac{\operatorname{Tr}_k(\mathcal{T})}{k})\neq 0$ *with* $1\leqslant k\leqslant n(m-1)^{n-1}$, *then* $\mathcal{T}$ *is positive semidefinite.*

Proof. Suppose that m is even and n is odd. Then, $n(m-1)^{n-1}$ is odd. Consequently,

$$\begin{aligned}\phi(\lambda)&:=\chi(-\lambda)\\&=-\lambda^{n(m-1)^{n-1}}+\sum_{k=1}^{n(m-1)^{n-1}}(-1)^{k+1}\lambda^{n(m-1)^{n-1}-k}p_k\left(-\frac{\operatorname{Tr}_1(\mathcal{T})}{1},\ldots,-\frac{\operatorname{Tr}_k(\mathcal{T})}{k}\right).\end{aligned}$$

Then, by Lemma 6.13, ϕ defined as above has no positive real root, since the sign of the coefficient of ϕ is negative when it is nonzero. Hence, χ has no negative real root. Consequently, $\mathcal{T}$ is positive semidefinite by Lemma 6.12.

The proof for the other case for m and n is similar. Consequently, the result follows. The proof is complete. □

Remark 6.15. If tensor $\mathcal{T}\in\mathbb{S}(\mathbb{R}^n,m)$ is nonzero and positive semidefinite, then $-\mathcal{T}$ isn't positive semidefinite. Consequently, (22) with $\mathcal{T}$ being replaced by $-\mathcal{T}$ fails to hold by Proposition 6.14. Hence, Proposition 6.14 provides a necessary as well as a sufficient condition for a nonzero tensor to be positive semidefinite. In the next section, the coefficients of the characteristic polynomial are discussed. When $n=2$, Theorem 7.7 gives explicit formulae for them. Then, it is easy to check the corresponding hypothesis in Proposition 6.14 in this case. While, for many tensors, both $\mathcal{T}$ and $-\mathcal{T}$ will fail to satisfy the hypotheses of Proposition 6.14, then the criteria will be inconclusive about the tensor's semidefiniteness.

7. The second order trace

We discuss more on the higher order traces of a tensor $\mathcal{T}\in\mathbb{T}(\mathbb{C}^n,m)$ in this section. Note that the trace formulae of both the characteristic polynomial and the determinant depend on the d-th order traces of the underlying tensor for all $d\in\{1,\ldots,n(m-1)^{n-1}\}$. However, it is very complicated (Cooper and Dutle, 2012; Morozov and Shakirov, 2011). In this section, we give preliminary results on the computation of the d-th order traces of a tensor. In particular, we give explicit formulae of $\operatorname{Tr}_2(\mathcal{T})$ for a tensor $\mathcal{T}$ of arbitrary order and dimension, and the characteristic polynomial $\chi(\lambda)$ and the determinant $\operatorname{Det}(\mathcal{T})$ of a tensor $\mathcal{T}$ when $n=2$.

The following lemma is a generalization of Proposition 6.4(i).

Lemma 7.1. *Let* $\mathcal{T}\in\mathbb{T}(\mathbb{C}^n,m)$ *and the notation be defined as above. We have*

$$\frac{(\hat{g}_i)^k}{((m-1)k)!}\operatorname{Tr}\left(A^{(m-1)k}\right)=t^k_{ii\ldots i} \tag{23}$$

for all $k\geqslant 0$ *and* $i\in\{1,\ldots,n\}$. *So,*

$$\sum_{i=1}^n\frac{(\hat{g}_i)^k}{((m-1)k)!}\operatorname{Tr}\left(A^{(m-1)k}\right)=\sum_{i=1}^n t^k_{ii\ldots i}. \tag{24}$$

Proof. By (11) and (15), similar to the proof of Proposition 6.14(i), we have

$$(\hat{g}_i)^k\operatorname{Tr}\left(A^{(m-1)k}\right)=\left[\sum_{i_2=1}^n\cdots\sum_{i_m=1}^n t_{ii_2\ldots i_m}\frac{\partial}{\partial a_{ii_2}}\cdots\frac{\partial}{\partial a_{ii_m}}\right]^k\operatorname{Tr}\left(A^{(m-1)k}\right)$$

$$= \left[\sum_{i_2=1}^{n} \cdots \sum_{i_m=1}^{n} t_{ii_2\ldots i_m} \frac{\partial}{\partial a_{ii_2}} \cdots \frac{\partial}{\partial a_{ii_m}} \right]^k$$
$$\cdot \left(\sum_{i_1=1}^{n} \cdots \sum_{i_{(m-1)k}=1}^{n} a_{i_1 i_2} a_{i_2 i_3} \ldots a_{i_{(m-1)k-1} i_{(m-1)k}} a_{i_{(m-1)k} i_1} \right)$$
$$= \left[t_{ii\ldots i} \frac{\partial}{\partial a_{ii}} \cdots \frac{\partial}{\partial a_{ii}} \right]^k (a_{ii})^{(m-1)k}$$
$$= \big((m-1)k\big)! t_{ii\ldots i}^k,$$

which implies (23), and hence (24). □

Before further analysis, we need the following combinatorial result.

Lemma 7.2. *Let $i \neq j$, $k \geqslant 1$, $h \geqslant 1$ and $s \in \{1, \ldots, \min\{h,k\}(m-1)\}$ be arbitrary but fixed. Then, the coefficient of the following term*

$$(a_{ii})^{k(m-1)-s} (a_{ij})^s (a_{ji})^s (a_{jj})^{h(m-1)-s}$$

in the expansion of $\mathrm{Tr}(A^{(k+h)(m-1)})$ *is*

$$\binom{k(m-1)}{s} \binom{h(m-1)-1}{s-1} + \binom{h(m-1)}{s} \binom{k(m-1)-1}{s-1}.$$

Proof. For the convenience of the subsequent analysis, we define a *packaged element* of i as an ordered collection of a_{ij}, a_{jj}'s and a_{ji} in the form:

$$a_{ij} \underbrace{a_{jj} \ldots a_{jj}}_{p} a_{ji}.$$

The number p of a_{jj}'s in a packaged element of i can vary from 0 to the maximal number. A packaged element of j can be defined similarly.

Note that any term in

$$\mathrm{Tr}\big(A^{(k+h)(m-1)}\big) = \sum_{i_1=1}^{n} \cdots \sum_{i_{(k+h)(m-1)}=1}^{n} a_{i_1 i_2} a_{i_2 i_3} \ldots a_{i_{(k+h)(m-1)-1} i_{(k+h)(m-1)}} a_{i_{(k+h)(m-1)} i_1}$$

which results in $(a_{ii})^{k(m-1)-s} (a_{ij})^s (a_{ji})^s (a_{jj})^{h(m-1)-s}$ has either the packaged elements of i or the packaged elements of j, not both, if we count from the left most in the expression, and is totally determined by the numbers of a_{jj}'s in the packaged elements and the positions of the packaged elements in the expression

$$a_{i_1 i_2} a_{i_2 i_3} \ldots a_{i_{(k+h)(m-1)-1} i_{(k+h)(m-1)}} a_{i_{(k+h)(m-1)} i_1}. \tag{25}$$

So, the coefficient of the term $(a_{ii})^{k(m-1)-s} (a_{ij})^s (a_{ji})^s (a_{jj})^{h(m-1)-s}$ in the expansion

$$\mathrm{Tr}\big(A^{(k+h)(m-1)}\big) = \sum_{i_1=1}^{n} \cdots \sum_{i_{(k+h)(m-1)}=1}^{n} a_{i_1 i_2} a_{i_2 i_3} \ldots a_{i_{(k+h)(m-1)-1} i_{(k+h)(m-1)}} a_{i_{(k+h)(m-1)} i_1}$$

is totally determined by the number of cases how the packaged elements are arranged multiplying the number of cases of the positions of the packaged elements in the expression (25).

In the following, we consider only the situation of packaged elements of i. The other situation is similar. Note that there are altogether s packaged elements of i in every expression (25) which result in

$$(a_{ii})^{k(m-1)-s}(a_{ij})^s(a_{ji})^s(a_{jj})^{h(m-1)-s}. \tag{26}$$

Firstly, note that there are $h(m-1)-s$ a_{jj}'s in (26). Then we have

$$\binom{h(m-1)-s+(s-1)}{s-1} = \binom{h(m-1)-1}{s-1}$$

different cases of s packaged elements of i with every case resulting in

$$(a_{ij})^s(a_{ji})^s(a_{jj})^{h(m-1)-s}.$$

Secondly, for an arbitrary but fixed case of s packaged elements of i in the first step, we get $k(m-1)$ "mixed" elements consisting of the s packaged elements and the rest $k(m-1)-s$ a_{ii}'s. Consequently, there are exactly

$$\binom{k(m-1)}{s}$$

cases of the expression (25), under which the expression (25) results in the term

$$(a_{ii})^{k(m-1)-s}(a_{ij})^s(a_{ji})^s(a_{jj})^{h(m-1)-s}.$$

Therefore, the number of times of the term (26) occurs in the expansion $\mathrm{Tr}(A^{(k+h)(m-1)})$ is

$$\binom{k(m-1)}{s}\binom{h(m-1)-1}{s-1}$$

in the situation of packaged elements of i.

By the symmetry of i and j, we can prove similarly that, in the situation of packaged elements of j, the number of times of the term (26) that occurs in the expansion $\mathrm{Tr}(A^{(k+h)(m-1)})$ is

$$\binom{h(m-1)}{s}\binom{k(m-1)-1}{s-1}.$$

Hence, we have that the coefficient of term $(a_{ii})^{k(m-1)-s}(a_{ij})^s(a_{ji})^s(a_{jj})^{k(m-1)-s}$ in the expansion $\mathrm{Tr}(A^{(k+h)(m-1)})$ is

$$\binom{k(m-1)}{s}\binom{h(m-1)-1}{s-1} + \binom{h(m-1)}{s}\binom{k(m-1)-1}{s-1}.$$

The proof is complete. $\square$

In the sequel, in order to make the operators in (11) more convenient to use and the resulting formulae more tidy, we reformulate $\hat{g}_i$ in the following way:

$$\hat{g}_i =: \sum_{1 \leqslant i_2 \leqslant i_3 \leqslant \cdots \leqslant i_m \leqslant n} w_{ii_2\ldots i_m} \frac{\partial}{\partial a_{ii_2}} \cdots \frac{\partial}{\partial a_{ii_m}}, \quad \forall i \in \{1, \ldots, n\}. \tag{27}$$

Lemma 7.3. *Let $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$. For arbitrary $i < j$, and $h, k \geqslant 1$, we have*

$$\begin{aligned} &\frac{(\hat{g}_i)^h(\hat{g}_j)^k}{(h(m-1))!(k(m-1))!}\mathrm{Tr}\big(A^{(h+k)(m-1)}\big) \\ &\quad = \left(\frac{h+k}{hk(m-1)}\right) \sum_{s=1}^{\min\{h,k\}(m-1)} \sum_{\substack{(a_1,\ldots,a_h)\in\mathbb{D}^s \\ (b_1,\ldots,b_k)\in\mathbb{E}^s}} s \prod_{p=1}^{h}\prod_{q=1}^{k} w_{ii\ldots i\underbrace{j\ldots j}_{a_p}} w_{j\underbrace{i\ldots i}_{b_q}j\ldots j}, \end{aligned} \tag{28}$$

with $\mathbb{D}^s := \{(a_1, \ldots, a_h) \mid a_1 + \cdots + a_h = s,\ 0 \leqslant a_p \leqslant m-1,\ \forall p \in \{1, \ldots, h\}\}$ *and* $\mathbb{E}^s := \{(b_1, \ldots, b_k) \mid b_1 + \cdots + b_k = s,\ 0 \leqslant b_q \leqslant m-1,\ \forall q \in \{1, \ldots, k\}\}$.

Proof. Let $w := \min\{h, k\}(m-1)$, $\mathbb{D}^s := \{(a_1, \ldots, a_h) \mid a_1 + \cdots + a_h = s,\ 0 \leqslant a_p \leqslant m-1,\ \forall p \in \{1, \ldots, h\}\}$ and $\mathbb{E}^s := \{(b_1, \ldots, b_k) \mid b_1 + \cdots + b_k = s,\ 0 \leqslant b_q \leqslant m-1,\ \forall q \in \{1, \ldots, k\}\}$ for all $s \in \{1, \ldots, w\}$. By (27) and (15), we have

$$\begin{aligned}
&(\hat{g}_i)^h (\hat{g}_j)^k \operatorname{Tr}\big(A^{(h+k)(m-1)}\big) \\
&= \left[\sum_{i_2 \leqslant \cdots \leqslant i_m} w_{ii_2\ldots i_m} \frac{\partial}{\partial a_{ii_2}} \cdots \frac{\partial}{\partial a_{ii_m}}\right]^h \left[\sum_{j_2 \leqslant \cdots \leqslant j_m} w_{jj_2\ldots j_m} \frac{\partial}{\partial a_{jj_2}} \cdots \frac{\partial}{\partial a_{jj_m}}\right]^k \\
&\quad \cdot \operatorname{Tr}\big(A^{(h+k)(m-1)}\big) \\
&= \left[\sum_{i_2 \leqslant \cdots \leqslant i_m} w_{ii_2\ldots i_m} \frac{\partial}{\partial a_{ii_2}} \cdots \frac{\partial}{\partial a_{ii_m}}\right]^h \left[\sum_{j_2 \leqslant \cdots \leqslant j_m} w_{jj_2\ldots j_m} \frac{\partial}{\partial a_{jj_2}} \cdots \frac{\partial}{\partial a_{jj_m}}\right]^k \\
&\quad \cdot \left(\sum_{i_1=1}^{n} \cdots \sum_{i_{(h+k)(m-1)}=1}^{n} a_{i_1 i_2} a_{i_2 i_3} \ldots a_{i_{(h+k)(m-1)-1} i_{(h+k)(m-1)}} a_{i_{(h+k)(m-1)} i_1}\right) \\
&= \left(\sum_{s=1}^{w} \sum_{(a_1,\ldots,a_h)\in\mathbb{D}^s,\ (b_1,\ldots,b_k)\in\mathbb{E}^s} \prod_{p=1}^{h} \prod_{q=1}^{k} w_{ii\ldots i\underbrace{j\ldots j}_{a_p}} w_{j\underbrace{i\ldots i}_{b_q} j\ldots j} \right. \\
&\quad \left. \cdot \left(\frac{\partial}{\partial a_{ii}}\right)^{h(m-1)-s} \left(\frac{\partial}{\partial a_{ij}}\right)^{s} \left(\frac{\partial}{\partial a_{ji}}\right)^{s} \left(\frac{\partial}{\partial a_{jj}}\right)^{k(m-1)-s}\right) \\
&\quad \cdot \left\{\sum_{s=1}^{w} \left[\binom{k(m-1)}{s}\binom{h(m-1)-1}{s-1} + \binom{h(m-1)}{s}\binom{k(m-1)-1}{s-1}\right]\right. \\
&\quad \left. \cdot (a_{ii})^{h(m-1)-s} (a_{ij})^{s} (a_{ji})^{s} (a_{jj})^{k(m-1)-s}\right\} \\
&= \sum_{s=1}^{w} \sum_{(a_1,\ldots,a_h)\in\mathbb{D}^s,\ (b_1,\ldots,b_k)\in\mathbb{E}^s} \prod_{p=1}^{h} \prod_{q=1}^{k} w_{ii\ldots i\underbrace{j\ldots j}_{a_p}} w_{j\underbrace{i\ldots i}_{b_q} j\ldots j} \\
&\quad \cdot s\Big(\big(k(m-1)\big)!\big(h(m-1)-1\big)! + \big(h(m-1)\big)!\big(k(m-1)-1\big)!\Big).
\end{aligned}$$

Here, the third equality follows from Lemma 7.2. Consequently, (28) follows. □

Especially, the following is a direct corollary of Lemma 7.3.

Corollary 7.4. *Let $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$. For arbitrary $i < j$, we have*

$$\frac{\hat{g}_i \hat{g}_j}{[(m-1)!]^2} \operatorname{Tr}\big(A^{2(m-1)}\big) = \sum_{s=1}^{m-1} \left(\frac{2s}{m-1}\right) w_{ii\ldots i\underbrace{j\ldots j}_{s}} w_{ji\ldots i\underbrace{j\ldots j}_{m-1-s}}.$$

Given an index set $\mathbb{L} := \{k_1, \ldots, k_l\}$ with k_s taking value in $\{1, \ldots, n\}$ for $s \in \{1, \ldots, l\}$, denote by $\mathbb{H}_i(\mathbb{L})$ the set of indices in $\mathbb{L}$ taking value i for $i \in \{1, \ldots, n\}$. We denote by $|\mathbb{E}|$ the cardinality of a set $\mathbb{E}$.

Theorem 7.5. *Let $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$. We have*

$$\begin{aligned}
\mathrm{Tr}_2(\mathcal{T}) &= (m-1)^{n-1}\left[\sum_{i=1}^{n} \frac{(\hat{g}_i)^2}{(2(m-1))!} + \sum_{i<j} \frac{\hat{g}_i \hat{g}_j}{[(m-1)!]^2}\right] \mathrm{Tr}\big(A^{2(m-1)}\big) \\
&= (m-1)^{n-1}\left[\sum_{i=1}^{n} t_{ii\ldots i}^2 + \sum_{i<j}\sum_{s=1}^{m-1}\left(\frac{2s}{m-1}\right)\right. \\
&\quad \cdot \left(\sum_{|\mathbb{H}_i(\{i_2,\ldots,i_m\})|=m-1-s,\ |\mathbb{H}_j(\{i_2,\ldots,i_m\})|=s} t_{ii_2\ldots i_m}\right) \\
&\quad \left. \cdot \left(\sum_{|\mathbb{H}_i(\{j_2,\ldots,j_m\})|=s,\ |\mathbb{H}_j(\{j_2,\ldots,j_m\})|=m-1-s} t_{jj_2\ldots j_m}\right)\right].
\end{aligned}$$

Proof. The result follows from Lemma 7.1, Corollary 7.4, (11) and (27) immediately. □

Remark 7.6. By Proposition 6.4, Theorem 7.5, and Corollary 6.5, we see that

$$\begin{aligned}
\sum_{i<j,\ \lambda_i,\lambda_j\in\sigma(\mathcal{T})} m_i m_j \lambda_i \lambda_j &= \frac{1}{2}\left(\left[\mathrm{Tr}_1(\mathcal{T})\right]^2 - \mathrm{Tr}_2(\mathcal{T})\right) \\
&= (m-1)^{n-2}\sum_{i<j}\sum_{s=1}^{m-1} s\left(\sum_{|\mathbb{H}_i(\{i_2,\ldots,i_m\})|=m-1-s,\ |\mathbb{H}_j(\{i_2,\ldots,i_m\})|=s} t_{ii_2\ldots i_m}\right) \\
&\quad \cdot \left(\sum_{|\mathbb{H}_i(\{j_2,\ldots,j_m\})|=s,\ |\mathbb{H}_j(\{j_2,\ldots,j_m\})|=m-1-s} t_{jj_2\ldots j_m}\right),
\end{aligned}$$

where m_i is the algebraic multiplicity of λ_i. The left most expression is a degree two elementary symmetric polynomial on the eigenvalues, the middle expression is the codegree two coefficient of the characteristic polynomial, and the right most expression is a polynomial in the entries of the tensor which is invariant under action of the permutation group on n elements to its indices.

When $n=2$, we can get the coefficients of the characteristic polynomial $\chi(\lambda)$ explicitly in terms of the entries of the underlying tensor by using Theorem 6.3, Lemmas 7.1 and 7.3. It is an alternate to Sylvester's formula (Sylvester, 1840). In the following theorem, w, $\mathbb{D}^s$'s and $\mathbb{E}^s$'s are defined as those in Lemma 7.3.

Theorem 7.7. *Let $\mathcal{T} \in \mathbb{T}(\mathbb{C}^2, m)$. We have*

$$\chi(\lambda) = \lambda^{2(m-1)} + \sum_{k=1}^{2(m-1)} \lambda^{2(m-1)-k} \sum_{i=1}^{k} \frac{1}{i!} \sum_{d_j>0,\ \sum_{j=1}^{i} d_j = k} \prod_{j=1}^{i} \frac{-\mathrm{Tr}_{d_j}(\mathcal{T})}{d_j}$$

with

$$\begin{aligned}
\mathrm{Tr}_d(\mathcal{T}) = (m-1)\Bigg\{&\left(t_{11\ldots 1}^d + t_{22\ldots 2}^d\right) + \sum_{h+k=d,h,k\geqslant 1}\sum_{s=1}^{w} \frac{s(h+k)}{hk(m-1)} \\
&\cdot \left[\sum_{\substack{(a_1,\ldots,a_h)\in\mathbb{D}^s \\ (b_1,\ldots,b_k)\in\mathbb{E}^s}} \prod_{p=1}^{h}\prod_{q=1}^{k}\left(\sum_{|\mathbb{H}_1(\{i_2,\ldots,i_m\})|=m-1-a_p,\ |\mathbb{H}_2(\{i_2,\ldots,i_m\})|=a_p} t_{1i_2\ldots i_m}\right)\right.
\end{aligned}$$

$$\cdot\Biggl(\sum_{|\mathbb{H}_1(\{j_2,\ldots,j_m\})|=b_p,\,|\mathbb{H}_2(\{j_2,\ldots,j_m\})|=m-1-b_p} t_{2j_2\ldots j_m}\Biggr)\Biggr]\Biggr]\Biggr\}$$

for $d \in \{1,\ldots,2(m-1)\}$.

It follows from Theorem 6.3 and Corollary 2.6 that $\mathrm{Det}(\mathcal{T}) = (-1)^{n(m-1)^{n-1}}\chi(0)$. Thus, when $n=2$, we get an explicit formula for $\mathrm{Det}(\mathcal{T})$ as

$$\mathrm{Det}(\mathcal{T}) = \sum_{i=1}^{2(m-1)} \frac{1}{i!} \sum_{d_j>0,\,\sum_{j=1}^{i} d_j=2(m-1)} \prod_{j=1}^{i} \frac{-\mathrm{Tr}_{d_j}(\mathcal{T})}{d_j}.$$

When $m=3$, we have the following corollary.

Corollary 7.8. *Let* $\mathcal{T} \in \mathbb{T}(\mathbb{C}^2,3)$. *We have*

$$\begin{aligned}
\chi(\lambda) = {} & \lambda^4 - \lambda^3\,\mathrm{Tr}_1(\mathcal{T}) + \frac{1}{2}\lambda^2\bigl(\bigl[\mathrm{Tr}_1(\mathcal{T})\bigr]^2 - \mathrm{Tr}_2(\mathcal{T})\bigr) \\
& + \frac{1}{12}\lambda\bigl(-2\bigl[\mathrm{Tr}_1(\mathcal{T})\bigr]^3 + 6\,\mathrm{Tr}_1(\mathcal{T})\,\mathrm{Tr}_2(\mathcal{T}) - 4\,\mathrm{Tr}_3(\mathcal{T})\bigr) \\
& + \frac{1}{24}\bigl(\bigl[\mathrm{Tr}_1(\mathcal{T})\bigr]^4 - 6\bigl[\mathrm{Tr}_1(\mathcal{T})\bigr]^2\,\mathrm{Tr}_2(\mathcal{T}) \\
& + 8\,\mathrm{Tr}_1(\mathcal{T})\,\mathrm{Tr}_3(\mathcal{T}) + 3\bigl[\mathrm{Tr}_2(\mathcal{T})\bigr]^2 - 6\,\mathrm{Tr}_4(\mathcal{T})\bigr)
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Det}(\mathcal{T}) = {} & \frac{1}{24}\bigl(\bigl[\mathrm{Tr}_1(\mathcal{T})\bigr]^4 - 6\bigl[\mathrm{Tr}_1(\mathcal{T})\bigr]^2\,\mathrm{Tr}_2(\mathcal{T}) \\
& + 8\,\mathrm{Tr}_1(\mathcal{T})\,\mathrm{Tr}_3(\mathcal{T}) + 3\bigl[\mathrm{Tr}_2(\mathcal{T})\bigr]^2 - 6\,\mathrm{Tr}_4(\mathcal{T})\bigr)
\end{aligned}$$

with

$$\begin{aligned}
\mathrm{Tr}_1(\mathcal{T}) = {} & 2(t_{111}+t_{222}), \\
\mathrm{Tr}_2(\mathcal{T}) = {} & 2\bigl(t_{111}^2+t_{222}^2\bigr) + 2(t_{112}+t_{121})(t_{212}+t_{221}) + 4(t_{122}t_{211}), \\
\mathrm{Tr}_3(\mathcal{T}) = {} & 2\bigl(t_{111}^3+t_{222}^3\bigr) \\
& + \frac{3}{2}(t_{112}+t_{121})\bigl[(t_{212}+t_{221})t_{222}\bigr] + 3t_{122}\bigl[(t_{212}+t_{221})^2 + t_{211}t_{222}\bigr] \\
& + \frac{3}{2}(t_{212}+t_{221})\bigl[(t_{112}+t_{121})t_{111}\bigr] + 3t_{211}\bigl[(t_{112}+t_{121})^2 + t_{122}t_{111}\bigr], \\
\mathrm{Tr}_4(\mathcal{T}) = {} & 2\bigl(t_{111}^4+t_{222}^4\bigr) \\
& + \frac{4}{3}(t_{112}+t_{121})\bigl[t_{222}^2(t_{212}+t_{221})\bigr] + \frac{8}{3}t_{122}\bigl[t_{222}(t_{212}+t_{221})^2 + t_{222}^2t_{211}\bigr] \\
& + \frac{4}{3}(t_{212}+t_{221})\bigl[t_{122}^2(t_{112}+t_{121})\bigr] + \frac{8}{3}t_{211}\bigl[t_{111}(t_{112}+t_{121})^2 + t_{111}^2t_{122}\bigr] \\
& + t_{111}(t_{121}+t_{121})t_{222}(t_{212}+t_{221}) + 3\bigl[t_{111}(t_{121}+t_{112})\bigr]\bigl[t_{211}(t_{221}+t_{212})\bigr] \\
& + 2\bigl[(t_{121}+t_{112})^2 + t_{122}t_{111}\bigr]\bigl[t_{211}t_{222} + (t_{212}+t_{221})^2\bigr] + 4t_{122}^2t_{211}^2.
\end{aligned}$$

8. Geršgorin's inequality of the determinant

We generalize Geršgorin's inequality for matrices (see Problem 6.1.3 in Horn and Johnson (1985)) to tensors in this section.

Lemma 8.1. *Let $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$ and $\rho(\mathcal{T}) := \max_{\lambda \in \sigma(\mathcal{T})} |\lambda|$ be its spectral radius. Then,*

$$\rho(\mathcal{T}) \leqslant \max_{1 \leqslant i \leqslant n} \left(\sum_{i_2,\ldots,i_m=1}^{n} |t_{ii_2\ldots i_m}| \right).$$

Proof. The result follows from Theorem 6 in Qi (2005) immediately. □

Proposition 8.2. *Let $\mathcal{T} \in \mathbb{T}(\mathbb{C}^n, m)$. Then,*

$$\left|\mathrm{Det}(\mathcal{T})\right| \leqslant \prod_{1 \leqslant i \leqslant n} \left(\sum_{i_2,\ldots,i_m=1}^{n} |t_{ii_2\ldots i_m}| \right)^{(m-1)^{n-1}}. \tag{29}$$

Proof. If $\sum_{i_2,\ldots,i_m=1}^{n} |t_{ii_2\ldots i_m}| = 0$ for some $i \in \{1, \ldots, n\}$, then $\mathrm{Det}(\mathcal{T}) = 0$ by Proposition 2.4(i). Consequently, (29) follows trivially.

In the following, suppose that $\sum_{i_2,\ldots,i_m=1}^{n} |t_{ii_2\ldots i_m}| \neq 0$ for all $i \in \{1, \ldots, n\}$. Let tensor $\mathcal{U} \in \mathbb{T}(\mathbb{C}^n, m)$ be defined as

$$u_{ii_2\ldots i_m} := \frac{t_{ii_2\ldots i_m}}{\sum_{i_2,\ldots,i_m=1}^{n} |t_{ii_2\ldots i_m}|}, \quad \forall i, i_2, \ldots, i_m \in \{1, \ldots, n\}. \tag{30}$$

Then, by Lemma 8.1, we have that $\rho(\mathcal{U}) \leqslant 1$. This, together with Corollary 6.5, further implies that

$$\left|\mathrm{Det}(\mathcal{U})\right| \leqslant 1.$$

Moreover, by Proposition 2.4(ii) and (30), it is clear that

$$\left|\mathrm{Det}(\mathcal{U})\right| = \frac{|\mathrm{Det}(\mathcal{T})|}{\prod_{1 \leqslant i \leqslant n} (\sum_{i_2,\ldots,i_m=1}^{n} |t_{ii_2\ldots i_m}|)^{(m-1)^{n-1}}}.$$

Consequently, (29) follows. The proof is complete. □

9. Final remarks

In this paper, we introduced the determinant of a tensor and investigated its various properties. The simple application in Section 5 demonstrates that the determinant theory is applicable and worth further exploring.

Certainly, there are many other issues for further research in the theory of the determinant, which is surely one of the foundations of the eigenvalue theory of tensors. For example:

- More explicit formulae and relations for the higher order traces and the characteristic polynomials of general tensors, like those in Theorems 6.10, 7.5 and 7.7, and Corollary 7.8. This can provide approaches for the computation of the eigenvalues, which are crucial in applications.
- The applications to the spectral hypergraph theory. In Corollary 3.14 in Cooper and Dutle (2012), the traces $\mathrm{Tr}_i(\mathcal{H})$ for $i \in \{1, \ldots, k-1\}$ for the adjacency tensor $\mathcal{H}$ of a k-uniform hypergraph are proved to be zero. This, together with Theorem 6.10, implies that the spectrum of a uniform hypergraph has special structures.

- More properties about the determinants. For example, given a matrix A, Laplace's formula (see p. 7 in Horn and Johnson (1985)) reads:

$$\mathrm{Det}(A)=\sum_{j=1}^{n}(-1)^{i+j}a_{ij}M_{ij} \tag{31}$$

with minor M_{ij} being defined to be the determinant of the $(n-1)\times(n-1)$ matrix that results from A by removing the i-th row and the j-th column. If a generalization of (31) can be derived for the determinant of a tensor, many other useful inequalities, for example, Oppenheim's inequality, can be proved for the determinants.

Acknowledgement

The authors are very grateful to the three referees for their valuable suggestions and constructive comments, which have considerably improved the paper.

References

Balan, V., 2012. Spectra of symmetric tensors and m-root Finsler models. Linear Algebra Appl. 436, 152–162.
Bézout, E., 1779. Théorie Générale des Equations Algébriques. De l'Imprimerie de Ph.-D. Pierres, Paris.
Brachat, J., Comon, P., Mourrain, B., Tsigaridas, E., 2010. Symmetric tensor decomposition. Linear Algebra Appl. 433, 1851–1872.
Cartwright, D., Sturmfels, B., 2011. The number of eigenvalues of a tensor. Linear Algebra Appl., http://dx.doi.org/10.1016/j.laa.2011.05.040.
Cayley, A., 1843. On the theory of determinants. Trans. Cambridge Phil. Soc. VIII, 1–16.
Cayley, A., 1845. On the theory of linear transformations. Cambridge Math. J. 4, 193–209. Reprinted in vol. 1 of his collected mathematical papers, 1845, pp. 80–94.
Chang, K.C., Pearson, K., Zhang, T., 2008. Perron–Frobenius theorem for nonnegative tensors. Commun. Math. Sci. 6, 507–520.
Chang, K.C., Pearson, K., Zhang, T., 2009. On eigenvalue problems of real symmetric tensors. J. Math. Anal. Appl. 350, 416–422.
Chang, K.C., Pearson, K., Zhang, T., 2011. Primitivity, the convergence of the NZQ method, and the largest eigenvalue for non-negative tensors. SIAM J. Matrix Anal. Appl. 32, 806–819.
Comon, P., Golub, G., Lim, L.-H., Mourrain, B., 2008. Symmetric tensors and symmetric tensor rank. SIAM J. Matrix Anal. Appl. 30, 1254–1279.
Cooper, J., Dutle, A., 2012. Spectra of uniform hypergraphs. Linear Algebra Appl. 436, 3268–3292.
Cox, D., Little, J., O'Shea, D., 1998. Using Algebraic Geometry. Springer-Verlag, New York.
Cox, D., Little, J., O'Shea, D., 2006. Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra. Springer-Verlag, New York.
de Silva, V., Lim, L.-H., 2008. Tensor rank and the ill-posedness of the best low-rank approximation problem. SIAM J. Matrix Anal. Appl. 30, 1084–1127.
Dixon, A.L., 1906. On a form of the eliminant of two quantics. Proc. Lond. Math. Soc. 6, 468–478.
Dolotin, V., Morozov, A., 2007. Introduction to Non-Linear Algebra. World Scientific.
Gelfand, I.M., Kapranov, M.M., Zelevinsky, A.V., 1992. Hyperdeterminants. Adv. Math. 96, 226–263.
Gelfand, I.M., Kapranov, M.M., Zelevinsky, A.V., 1994. Discriminants, Resultants and Multidimensional Determinants. Birkhäuser, Boston.
Horn, R.A., Johnson, C.R., 1985. Matrix Analysis. Cambridge University Press, New York.
Horn, R.A., Johnson, C.R., 1994. Topics in Matrix Analysis. Cambridge University Press, New York.
Hu, S., Huang, Z.H., Ni, H.Y., Qi, L., 2012. Positive definiteness of diffusion kurtosis imaging. Inverse Probl. Imag. 6, 57–75.
Hu, S., Qi, L., 2012. Algebraic connectivity of an even uniform hypergraph. J. Comb. Optim. 24, 564–579.
Landsberg, J.M., 2011. Tensors: Geometry and Applications. Grad. Stud. Math., vol. 128. American Mathematical Society, Providence.
Lang, S., 1987. Linear Algebra, revised third edition. Springer-Verlag, New York.
Lang, S., 2002. Algebra, revised third edition. Springer-Verlag, New York.
Li, A.-M., Qi, L., Zhang, B., 2013. E-characteristic polynomials of tensors. Commun. Math. Sci. 11, 33–53.
Lim, L.-H., 2005. Singular values and eigenvalues of tensors: a variational approach. In: Proceedings of the IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, CAMSAP '05, 1, pp. 129–132.
Lim, L.-H., 2007. Foundations of Numerical Multilinear Algebra: decomposition and approximation of tensors. PhD thesis. Stanford University, USA.
Macaulay, F.S., 1902. On some formulas in elimination. Proc. Lond. Math. Soc. 3, 3–27.
Morozov, A., Shakirov, Sh., 2010. New and old results in resultant theory. Theoret. Math. Phys. 163, 587–617.
Morozov, A., Shakirov, Sh., 2011. Analogue of the identity Log Det = Trace Log for resultants. J. Geom. Phys. 61, 708–726.
Ng, M., Qi, L., Zhou, G., 2009. Finding the largest eigenvalue of a non-negative tensor. SIAM J. Matrix Anal. Appl. 31, 1090–1099.
Oeding, L., Ottaviani, G., 2011. Eigenvectors of tensors and algorithms for Waring decomposition. arXiv:1103.0203.
Pearson, K.J., 2010. Essentially positive tensors. Int. J. Algebra 4, 421–427.

Qi, L., 2005. Eigenvalues of a real supersymmetric tensor. J. Symb. Comput. 40, 1302–1324.
Qi, L., 2006. Rank and eigenvalues of a supersymmetric tensor, a multivariate homogeneous polynomial and an algebraic surface defined by them. J. Symb. Comput. 41, 1309–1327.
Qi, L., 2007. Eigenvalues and invariants of tensors. J. Math. Anal. Appl. 325, 1363–1377.
Qi, L., 2011. The best rank-one approximation ratio of a tensor space. SIAM J. Matrix Anal. Appl. 32, 430–442.
Qi, L., Yu, G., Wu, E.X., 2010. Higher order positive semi-definite diffusion tensor imaging. SIAM J. Imag. Sci. 3, 416–433.
Ragnarsson, S., Van Loan, C.F., 2011. Block tensors and symmetric embeddings. Linear Algebra Appl., http://dx.doi.org/10.1016/j.laa.2011.04.014.
Rota Bulò, S., Pelillo, M., 2009. A generalization of the Motzkin–Straus theorem to hypergraphs. Optim. Lett. 3, 187–295.
Strang, G., 1993. Introduction to Linear Algebra. Wellesley Cambridge Press.
Sturmfels, B., 2002. Solving Systems of Polynomial Equations. American Mathematical Society, Providence.
Sylvester, J.J., 1840. A method of determining by mere inspection the derivatives from two equations of any degree. Philos. Mag. 16, 132–135.
Yang, Y., Yang, Q., 2010. Further results for Perron–Frobenius theorem for nonnegative tensors. SIAM J. Matrix Anal. Appl. 31, 2517–2530.
Zhang, X., Qi, L., Ye, Y., 2012. The cubic spherical optimization problems. Math. Comput. 81, 1513–1525.