ADVANCES IN APPLIED MATHEMATICS AND MECHANICS IN CHINA
VOL. 4

G17P2
2

Editor Chien Weizang
Associate Editor Fu Zizhi

INTERNATIONAL ACADEMIC PUBLISHERS

ADVANCES IN APPLIED MATHEMATICS AND MECHANICS IN CHINA

VOL. 4

Editor
Prof. Chien Weizang

Associate Editor
Prof. Fu Zizhi

INTERNATIONAL ACADEMIC PUBLISHERS

(京)新登字 141 号

Published and Distributed by
International Academic Publishers
137 Chaonei Dajie
Beijing 100010, People's Republic of China

First edition 1992

Chien Weizang and Fu Zizhi
Advances in Applied Mathematics & Mechanics in China, Vol. 4

ISBN 7-80003-186-1/O·42

Printed by the Printing House of China Building Industry Press

中国应用数学与力学进展　　第四卷

主编　钱伟长　　副主编　付子智

万国学术出版社出版

(北京朝内大街 137 号, 邮编 100010)

新华书店发行　　建筑工业出版社印刷厂印装

开本 787×1092　1/16　印张 21.25　字数 510 千字

1993 年 4 月第一版　　1993 年 4 月第一次印刷

印数 0001–0300

ISBN 7-80003-186-1/O·42　　定价 96.00 元

Contents

Exterior-Algebraic Method in Tensor Calculus

Guo Zhong-heng (郭仲衡)
(*Dept. of Mathematics, Peking University*)

Abstract

Using the exterior-algebraic technique, this paper provides a systematic formal derivation of expressions for the principal invariants of an endomorphism, an intrinsic proof of Cayley-Hamilton theorem, a direct proof of Newton's formulae, and a direct derivation of the derivatives of the principal invariants. The present approach substantially improves those in existence.

The aim of the present paper is to show that the usage of the exterior algebra changes significantly the feature of tensor calculus. In order to expose the elegance and potential of this technique, after some preliminaries, this paper provides derivations or proofs for 4 problems, which improve substantially the existing ones. The whole development is n-dimensional. Throughout this paper, the range of index is from 1 to n and the summation convention is applied.

I. Preliminaries

Let $\boldsymbol{R}$ be the field of real numbers, $\mathcal{V}$ an n-dimensional (real) vector space and $\mathcal{V}^*$ its dual. A co-vector $\boldsymbol{\beta} \in \mathcal{V}^*$ is a linear functional on $\mathcal{V}$:

$$\boldsymbol{\beta} : \mathcal{V} \to \boldsymbol{R} : \boldsymbol{u} \mapsto \boldsymbol{\beta}(\boldsymbol{u}) =:< \boldsymbol{\beta}, \boldsymbol{u} > . \tag{1.1}$$

The set of all endomorphisms $\boldsymbol{A}, \boldsymbol{B}, \dots$ from $\mathcal{V}$ into itself are denoted by $\mathcal{L}(\mathcal{V})$.

Definition 1.1 For any positive integer r, the dual pairing of exterior products $\boldsymbol{\beta}^1 \wedge \cdots \wedge \boldsymbol{\beta}^r$ and $\boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_r (\boldsymbol{\beta}^1, \dots, \boldsymbol{\beta}^r \in \mathcal{V}^*; \boldsymbol{u}_1, \dots, \boldsymbol{u}_r \in \mathcal{V})$ are defined as

$$\langle \boldsymbol{\beta}^1 \wedge \cdots \wedge \boldsymbol{\beta}^r,\ \boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_r \rangle := \begin{vmatrix} \langle \boldsymbol{\beta}^1,\ \boldsymbol{u}_1 \rangle & \dots & \langle \boldsymbol{\beta}^1,\ \boldsymbol{u}_r \rangle \\ \vdots & \ddots & \vdots \\ \langle \boldsymbol{\beta}^r,\ \boldsymbol{u}_1 \rangle & \dots & \langle \boldsymbol{\beta}^r,\ \boldsymbol{u}_r \rangle \end{vmatrix} . \tag{1.2}$$

Definition 1.2 Throughout this paper, $\{e_i\}$ and $\{\alpha^i\}$ always denote the dual bases of $\mathcal{V}$ and $\mathcal{V}^*$, satisfying

$$\langle \alpha^i, e_j \rangle = \delta^i_j. \tag{1.3}$$

For any positive integer r, the generalized Kronecker delta of r^{th}-order is defined as

$$\delta^{i_1 \dots i_r}_{j_1 \dots j_r} := \langle \alpha^{i_1} \wedge \cdots \wedge \alpha^{i_r}, e_{j_1} \wedge \cdots \wedge e_{j_r} \rangle. \tag{1.4}$$

Corollary 1.3

$$\delta^{1 \dots n}_{1 \dots n} = \langle \alpha^1 \wedge \cdots \wedge \alpha^n,\ e_1 \wedge \cdots \wedge e_n \rangle = 1. \tag{1.5}$$

Definition 1.4 The trace of the endomorphism A is defined as

$$\mathrm{tr} A := \langle \alpha^i, A e_i \rangle. \tag{1.6}$$

It is easy to show:

Lemma 1.5

$$\langle \alpha^j, B e_i \rangle A e_j = A B e_i, \quad \forall A, B \in \mathcal{L}(\mathcal{V}). \tag{1.7}$$

Lemma 1.6 Let $\{u_i\} \subset \mathcal{V}$ and $\{\beta^i\} \subset \mathcal{V}^*$ are sets of n linearly independent vectors and co-vectors, respectively. Then

$$\langle \beta^1 \wedge \cdots \wedge \beta^n,\ u_1 \wedge \cdots \wedge u_n \rangle = \det(\langle \beta^i, u_j \rangle) \neq 0. \tag{1.8}$$

Proof Consider the homogeneous linear system

$$\langle \beta^i, u_j \rangle x^j = 0 \tag{1.9}$$

with $\{x^j\} \subset R$ as unknowns. Any co-vector $\xi \in \mathcal{V}^*$ may be expressed as $\xi = \xi_i \beta^i$. Multiplying (1.9) by ξ_i and summing it with respect to index i, we get

$$\langle \xi, x^j u_j \rangle = 0.$$

In virtue of arbitrariness of ξ and definiteness of dual pairing, we have

$$x^j u_j = 0$$

and the linear independence of $\{u_i\}$ yields $x^j = 0$. This means that (1.8) has only trivial solution and $\det(\langle \beta^i, u_j \rangle) \neq 0$.

II. Principal Invariants of an Endomorphism

The formal derivation of the classical componential expression for the principal invariants of an endomorphism in n-dimensional space is unavailable in normal textbooks[1,2]. We can find this derivation in Ref. [3]. In this section we derive three expressions in an alternative way. The first two (2.10) and (2.17) are in terms of exterior algebra and the third one (2.18) coincides with the classical expression.

Theorem 2.1 Relation

$$\boldsymbol{A}\boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{A}\boldsymbol{u}_n = (\det \boldsymbol{A})\boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_n, \quad \forall \boldsymbol{A} \in \mathcal{L}(\mathcal{V});\ \boldsymbol{u}_i \in \mathcal{V} \tag{2.1}$$

holds, where $\det\boldsymbol{A} \in \boldsymbol{R}$ is called the determinant of $\boldsymbol{A}$. $\det\boldsymbol{A}$ does not depend upon the choice of $\{\boldsymbol{u}_i\}$. If $\{\boldsymbol{u}_i\} \subset \mathcal{V}$ and $\{\boldsymbol{\beta}^i\} \subset \mathcal{V}^*$ are sets of linearly independent vectors and co-vectors, then $\det\boldsymbol{A}$ has the expression

$$\det\boldsymbol{A} = \frac{\det(\langle \boldsymbol{\beta}^i, \boldsymbol{A}\boldsymbol{u}_j\rangle)}{\det(\langle \boldsymbol{\beta}^i, \boldsymbol{u}_j\rangle)}. \tag{2.2}$$

Proof If $\{\boldsymbol{u}_i\}$ is a set of linearly dependent vectors, then $\{\boldsymbol{A}\boldsymbol{u}_i\}$ is also such a set, and (2.1) holds automatically. Thus, it remains for us to prove (2.1) for any independent set $\{\boldsymbol{u}_i\}$. In this case, $\boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_n$ is a non-vanishing n-form and any n-form is its multiple. In particular, we have

$$\boldsymbol{A}\boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{A}\boldsymbol{u}_n = \mu \boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_n \tag{2.3}$$

and

$$\boldsymbol{A}\boldsymbol{u}_{1'} \wedge \cdots \wedge \boldsymbol{A}\boldsymbol{u}_{n'} = \mu' \boldsymbol{u}_{1'} \wedge \cdots \wedge \boldsymbol{u}_{n'} \tag{2.4}$$

for another independent set $\{\boldsymbol{u}_{i'}\}$:

$$\boldsymbol{u}_i = \Phi_i^{i'} \boldsymbol{u}_{i'}; \quad \det(\Phi_i^{i'}) \neq 0. \tag{2.5}$$

Constructing the dual pairing of (2.3) and (2.4) with the exterior product of any linearly independent co-vector set $\{\boldsymbol{\beta}^i\}$, and taking (1.8) into account, we obtain

$$\mu = \frac{\det(\langle \boldsymbol{\beta}^i, \boldsymbol{A}\boldsymbol{u}_j\rangle)}{\det(\langle \boldsymbol{\beta}^i, \boldsymbol{u}_j\rangle)} \tag{2.6}$$

and

$$\mu' = \frac{\det(\langle \boldsymbol{\beta}^i, \boldsymbol{A}\boldsymbol{u}_{j'}\rangle)}{\det(\langle \boldsymbol{\beta}^i, \boldsymbol{u}_{j'}\rangle)}.$$

Condition (2.5) yields that $\mu = \mu'$. From expression (2.6) we can see that μ is also independent of the choice of $\{\boldsymbol{\beta}^i\}$. Thus, the real number μ is just the determinant of $\boldsymbol{A}$ with the property stated in the theorem. □

Now we substitute $\boldsymbol{A} - \lambda \boldsymbol{I}$ ($\forall \lambda \in \boldsymbol{R}$, $\boldsymbol{I}$—identity endomorphism) into (2.1) for $\boldsymbol{A}$, we have

$$(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{u}_1 \wedge \cdots \wedge (\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{u}_n = (\det(\boldsymbol{A} - \lambda\boldsymbol{I}))\boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_n. \tag{2.7}$$

By virtue of linearity of exterior multiplication, the left-hand side of (2.7) may be expanded into a polynomial of λ with n-form "coefficients".

$$\sum_{r=0}^{n}\Big(\sum_{1\le i_1<\cdots<i_r\le n} \boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{A}\boldsymbol{u}_{i_1} \wedge \cdots \wedge \boldsymbol{A}\boldsymbol{u}_{i_r} \wedge \cdots \wedge \boldsymbol{u}_n\Big)(-\lambda)^{n-r} = (\det(\boldsymbol{A} - \lambda\boldsymbol{I}))\boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_n. \tag{2.8}$$

If $\{\boldsymbol{u}_i\}$ is a linearly independent set, then each "coefficient" can be expressed in terms of $\boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_n$:

$$\sum_{1\le i_1<\cdots<i_r\le n} \boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{A}\boldsymbol{u}_{i_1} \wedge \cdots \wedge \boldsymbol{A}\boldsymbol{u}_{i_r} \wedge \cdots \wedge \boldsymbol{u}_n = I_r \boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_n, \quad r = 1, \ldots, n. \tag{2.9}$$

This relation holds even when $\boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_n = 0$ (i.e. $\{\boldsymbol{u}_i\}$ is linearly dependent), because in this case each "coefficient" of $(-\lambda)^{n-r}$ on the left-hand side of (2.8) must vanish. The real numbers $I_1, I_2, \ldots, I_n$ are called the principal invariants of $\boldsymbol{A}$. The next theorem shows that they are also independent of the choice of $\{\boldsymbol{u}_i\}$.

Theorem 2.2 The principal invariants $I_r (r = 1, 2, \ldots, n)$ of an endomorphism $\boldsymbol{A}$ do not depend upon the choice of $\{\boldsymbol{u}_i\}$. If $\{\boldsymbol{u}_i\} \subset \mathcal{V}$ and $\{\boldsymbol{\beta}^i\} \subset \mathcal{V}^*$ are linearly independent sets, then I_r has the expression:

$$I_r = \frac{\langle \boldsymbol{\beta}^1 \wedge \cdots \wedge \boldsymbol{\beta}^n, \sum_{1\le i_1<\cdots<i_r\le n} \boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{A}\boldsymbol{u}_{i_1} \wedge \cdots \wedge \boldsymbol{A}\boldsymbol{u}_{i_r} \wedge \cdots \wedge \boldsymbol{u}_n\rangle}{\langle \boldsymbol{\beta}^1 \wedge \cdots \wedge \boldsymbol{\beta}^n, \boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_n\rangle}. \tag{2.10}$$

Proof Constructing dual pairing of (2.9) with $\boldsymbol{\beta}^1 \wedge \cdots \wedge \boldsymbol{\beta}^n$, we get (2.10). Analogous to the last theorem, the proof of the first part of the theorem reduces to the proof of

$$\boldsymbol{S}(i) = \det(\Phi_i^{i'})\boldsymbol{S}(i'), \tag{2.11}$$

where

$$\boldsymbol{S}(i) \equiv \sum_{1\le i_1<\cdots<i_r\le n} \boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{A}\boldsymbol{u}_{i_1} \wedge \cdots \wedge \boldsymbol{A}\boldsymbol{u}_{i_r} \wedge \cdots \wedge \boldsymbol{u}_n, \tag{2.12}$$

$$\boldsymbol{S}(i') = \sum_{1'\le i'_1<\cdots<i'_r\le n'} \boldsymbol{u}_{1'} \wedge \cdots \wedge \boldsymbol{A}\boldsymbol{u}_{i'_1} \wedge \cdots \wedge \boldsymbol{A}\boldsymbol{u}_{i'_r} \wedge \cdots \wedge \boldsymbol{u}_{n'},$$

and $\{u_{i'}\}$ is an arbitrary independent set satisfying (2.5). Substituting (2.5) into (2.12), we have

$$S(i) = \sum_{1\le i_1<\cdots<i_r\le n} \Phi_1^{i'_1}\ldots\Phi_{i_1}^{i'_p}\ldots\Phi_{i_r}^{i'_q}\ldots\Phi_n^{i'_n} u_{i'_i}\wedge\cdots\wedge Au_{i'_p(i_1)}\wedge$$

$$\cdots\wedge Au_{i'_q(i_r)}\wedge\cdots\wedge u_{i'_n}.$$

The vector with subscript (i_1) is the $i_1{}^{\text{th}}$-element in the exterior product. Notice that the lower indices in the product of n coefficients $\Phi_i^{i'}$ are always $(1,2,\ldots,n)$, independent of $(i_1,\ldots,i_r)$. Hence

$$S(i) = \Phi_1^{i'_1}\ldots\Phi_n^{i'_n}\sum_{1\le i_1<\cdots<i_r\le n} u_{i'_1}\wedge\cdots\wedge Au_{i'_p(i_1)}\wedge\cdots\wedge Au_{i'_q(i_r)}\wedge\cdots\wedge u_{i'_n}. \tag{2.13}$$

It is easy to check that in

$$T' = \sum_{1\le i_1<\cdots<i_r\le n} u_{i'_1}\wedge\cdots\wedge Au_{i'_p(i_1)}\wedge\cdots\wedge Au_{i'_q(i_r)}\wedge\cdots\wedge u_{i'_n} \tag{2.14}$$

an interchange of two indices i' and j' causes the change of sign. Therefore, it suffices to confine the consideration to sums T' with distinct lower indices. If $(i'_1,\ldots,i'_n)$ is a permutation of $(1',\ldots,n')$: $i'_k=\sigma(k'), k'=1',\ldots,n'$, then any T' and (2.13) may be written as

$$\begin{aligned} T' &= \operatorname{sgn}\sigma \sum_{1'\le i'_1<\cdots<i'_r\le n'} u_{1'}\wedge\cdots\wedge Au_{i'_1}\wedge\cdots\wedge Au_{i'_r}\wedge\cdots\wedge u_{n'} \\ &= \operatorname{sgn}\sigma S(i') \end{aligned}$$

and

$$S(i) = \sum_{\sigma\in\mathcal{P}_n} \operatorname{sgn}\sigma\Phi_1^{\sigma(i')}\ldots\Phi_n^{\sigma(n')}S(i') = \det(\Phi_i^{i'})S(i'),$$

where $\mathcal{P}_n$ is an n-element permutation group. This is just the formula (2.11) to be shown. □

With the notation of

$$I_0 = 1 \tag{2.15}$$

and expression (2.9), (2.7) may be written as

$$\Big[\sum_{r=0}^{n} I_r(-\lambda)^{n-r} - \det(A-\lambda I)\Big] u_1\wedge\cdots\wedge u_n = 0. \tag{2.16}$$

If we choose the dual bases $\{\boldsymbol{e}_i\}$ and $\{\boldsymbol{\alpha}^i\}$ for $\{\boldsymbol{u}_i\}$ and $\{\boldsymbol{\beta}^i\}$ in (2.10) respectively, then in virtue of (1.3) and (1.5), expression (2.10) reduces to a simpler form

$$\begin{aligned} I_r &= \sum_{1\le i_1<\cdots<i_r\le n} \langle \boldsymbol{\alpha}^1\wedge\cdots\wedge\boldsymbol{\alpha}^n, \boldsymbol{e}_1\wedge\cdots\wedge \boldsymbol{A}\boldsymbol{e}_{i_1}\wedge\cdots\wedge\boldsymbol{A}\boldsymbol{e}_{i_r}\wedge\cdots\wedge\boldsymbol{e}_n\rangle \\ &= \sum_{1\le i_1<\cdots<i_r\le n} \langle \boldsymbol{\alpha}^{i_1}\wedge\cdots\wedge\boldsymbol{\alpha}^{i_r}, \boldsymbol{A}\boldsymbol{e}_{i_1}\wedge\cdots\wedge\boldsymbol{A}\boldsymbol{e}_{i_r}\rangle \\ &= \frac{1}{r!}\langle \boldsymbol{\alpha}^{i_1}\wedge\cdots\wedge\boldsymbol{\alpha}^{i_r}, \boldsymbol{A}\boldsymbol{e}_{i_1}\wedge\cdots\wedge\boldsymbol{A}\boldsymbol{e}_{i_r}\rangle. \end{aligned} \tag{2.17}$$

We can see that

$$I_1 = \mathrm{tr}\boldsymbol{A}, \quad I_n = \det\boldsymbol{A}.$$

Adopting the decomposition

$$\boldsymbol{A}\boldsymbol{e}_i = A^j{}_i\boldsymbol{e}_j$$

and taking (1.4) into account, from (2.17) we get the classical componential expression

$$\begin{aligned} I_r &= \frac{1}{r!}A^{j_1}{}_{i_1}\ldots A^{j_r}{}_{i_r}\langle \boldsymbol{\alpha}^{i_1}\wedge\cdots\wedge\boldsymbol{\alpha}^{i_r}, \boldsymbol{e}_{j_1}\wedge\cdots\wedge\boldsymbol{e}_{j_r}\rangle \\ &= \frac{1}{r!}\delta^{i_1\ldots i_r}_{j_1\ldots j_r}A^{j_1}{}_{i_1}\ldots A^{j_r}{}_{i_r}. \end{aligned} \tag{2.18}$$

If λ is an eigenvalue of $\boldsymbol{A}$, we can take the corresponding eigenvector to be $\boldsymbol{u}_1$, say, in (2.7), then

$$\det(\boldsymbol{A}-\lambda\boldsymbol{I}) = 0.$$

In other words, on the basis of (2.16), the eigenvalue λ is a root of the characteristic equation of $\boldsymbol{A}$:

$$f(\lambda) := \sum_{r=0}^{n} I_r(-\lambda)^{n-r} = 0. \tag{2.19}$$

III. Cayley-Hamilton Theorem

The existing proofs of n-dimensional Cayley-Hamilton theorem are componental except for Truesdell and Noll's one.[4] Under the assumed invertibility of the endomorphism $\boldsymbol{A}$, Truesdell and Noll gave an intrinsic proof. For the intrinsic proof given here, no restriction is needed to be imposed on $\boldsymbol{A}$.

Theorem 3.1 Any endomorphism $\boldsymbol{A}$ satisfies its characteristic equation:

$$f(\boldsymbol{A}) := \sum_{r=0}^{n} I_r(-\boldsymbol{A})^{n-r} = 0. \tag{3.1}$$

Proof Denoting

$$S_r = \sum_{1\le i_1<\cdots<i_r<n} \boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{A}\boldsymbol{u}_{i_1} \wedge \cdots \wedge \boldsymbol{A}\boldsymbol{u}_{i_r} \wedge \cdots \wedge \boldsymbol{u}_{n-1},$$

$$r = 1, \ldots, n-1$$

and

$$S_0 = \boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_{n-1},$$

and using (2.15), we have

$$I_0 \boldsymbol{u}_1 \wedge \cdots \wedge (-\boldsymbol{A})^{n-0}\boldsymbol{u}_n = \boldsymbol{S}_0 \wedge (-\boldsymbol{A})^{n-0}\boldsymbol{u}_n. \tag{3.2}$$

Replacing $\boldsymbol{u}_n$ in (2.9) by $(-\boldsymbol{A})^{n-r}\boldsymbol{u}_n$, we obtain n equations:

$$I_r \boldsymbol{u}_1 \wedge \cdots \wedge (-\boldsymbol{A})^{n-r}\boldsymbol{u}_n = \boldsymbol{S}_r \wedge (-\boldsymbol{A})^{n-r}\boldsymbol{u}_n - \boldsymbol{S}_{r-1} \wedge (-\boldsymbol{A})^{n-(r-1)}\boldsymbol{u}_n,$$

$$r = 1, \ldots, n-1 \tag{3.3}$$

and

$$I_n \boldsymbol{u}_1 \wedge \cdots \wedge (-\boldsymbol{A})^{n-n}\boldsymbol{u}_n = -\boldsymbol{S}_{n-1} \wedge (-\boldsymbol{A})^{n-(n-1)}\boldsymbol{u}_n. \tag{3.4}$$

Summing the $n+1$ equations (3.2) – (3.4), we get

$$\sum_{r=0}^{n} I_r \boldsymbol{u}_1 \wedge \cdots \wedge (-\boldsymbol{A})^{n-r}\boldsymbol{u}_n = \sum_{r=0}^{n-1} \boldsymbol{S}_r \wedge (-\boldsymbol{A})^{n-r}\boldsymbol{u}_n - \sum_{r=1}^{n} \boldsymbol{S}_{r-1} \wedge (-\boldsymbol{A})^{n-(r-1)}\boldsymbol{u}_n$$

$$= 0$$

or

$$\boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_{n-1} \wedge \left[\sum_{r=0}^{n} I_r (-\boldsymbol{A})^{n-r}\right]\boldsymbol{u}_n = 0. \tag{3.5}$$

In virtue of arbitrariness of $\{\boldsymbol{u}_i\}$, (3.5) yields (3.1).

IV. Key Recurrence Formula

In this section we shall derive a recurrence formula which will be essential for further development.

Lemma 4.1 For any positive integers r and s, satisfying $0 < r - s \le n$, and $\boldsymbol{A}, \boldsymbol{B} \in \mathcal{L}(\mathcal{V})$, recurrence formula

$$I_{r-s}\mathrm{tr}((-\boldsymbol{A})^{s-1}\boldsymbol{B}) = F(\boldsymbol{A}, \boldsymbol{B}; r, s) - F(\boldsymbol{A}, \boldsymbol{B}; r, s+1) \tag{4.1}$$

holds, where

$$F(\boldsymbol{A},\boldsymbol{B};r,s)=\frac{1}{(r-s)!}\langle\boldsymbol{\alpha}^{i_1}\wedge\cdots\wedge\boldsymbol{\alpha}^{i_{r-(s-1)}},$$
$$\boldsymbol{A}\boldsymbol{e}_{i_1}\wedge\cdots\wedge\boldsymbol{A}\boldsymbol{e}_{i_{r-s}}\wedge(-\boldsymbol{A})^{s-1}\boldsymbol{B}\boldsymbol{e}_{i_{r-(s-1)}}\rangle. \tag{4.2}$$

The condition imposed on r and s implies that $r>1$ and $s<r$.

Proof Taking (1.6), (1.7) and (2.17) into account, keeping Definition 1.2 in mind, and expanding the determinants on the right-hand side of (4.2) with respect to the last column, we get

$$\begin{aligned}F(\boldsymbol{A},\boldsymbol{B};r,s)&=\frac{1}{(r-s)!}\langle\boldsymbol{\alpha}^{i_1}\wedge\cdots\wedge\boldsymbol{\alpha}^{i_{r-s}},\boldsymbol{A}\boldsymbol{e}_{i_1}\wedge\cdots\wedge\boldsymbol{A}\boldsymbol{e}_{i_{r-s}}\rangle\\
&\quad\langle\boldsymbol{\alpha}^{i_{r-s+1}},(-\boldsymbol{A})^{s-1}\boldsymbol{B}\boldsymbol{e}_{i_{r-s+1}}\rangle\\
&\quad+\frac{1}{(r-s)!}\sum_{p=1}^{r-s}(-1)^{r-s+p+1}\langle\boldsymbol{\alpha}^{i_p},(-\boldsymbol{A})^{s-1}\boldsymbol{B}\boldsymbol{e}_{i_{r-s+1}}\rangle\\
&\quad\times\langle\boldsymbol{\alpha}^{i_1}\wedge\cdots\wedge\boldsymbol{\alpha}^{i_{p-1}}\wedge\boldsymbol{\alpha}^{i_{p+1}}\wedge\cdots\wedge\boldsymbol{\alpha}^{i_{r-s+1}},\\
&\quad\boldsymbol{A}\boldsymbol{e}_{i_1}\wedge\cdots\wedge\boldsymbol{A}\boldsymbol{e}_{i_p}\wedge\cdots\wedge\boldsymbol{A}\boldsymbol{e}_{i_{r-s}}\rangle\\
&=I_{r-s}\mathrm{tr}((-\boldsymbol{A})^{s-1}\boldsymbol{B})\\
&\quad+\frac{1}{(r-s)!}\sum_{p=1}^{r-s}(-1)^{r-s+p}\langle\boldsymbol{\alpha}^{i_1}\wedge\cdots\wedge\boldsymbol{\alpha}^{i_{p-1}}\wedge\boldsymbol{\alpha}^{i_{p+1}}\wedge\cdots\\
&\quad\wedge\boldsymbol{\alpha}^{i_{r-s+1}},\boldsymbol{A}\boldsymbol{e}_{i_1}\wedge\cdots\wedge(-\boldsymbol{A})^{s}\boldsymbol{B}\boldsymbol{e}_{i_{r-s+1}}\wedge\cdots\wedge\boldsymbol{A}\boldsymbol{e}_{i_{r-s}})\\
&=I_{r-s}\mathrm{tr}((-\boldsymbol{A})^{s-1}\boldsymbol{B})+\frac{1}{(r-(s+1))!}\langle\boldsymbol{\alpha}^{i_1}\wedge\cdots\wedge\boldsymbol{\alpha}^{i_{r-s}},\\
&\quad\boldsymbol{A}\boldsymbol{e}_{i_1}\wedge\cdots\wedge\boldsymbol{A}\boldsymbol{e}_{i_{r-s+1}}\wedge(-\boldsymbol{A})^{s}\boldsymbol{B}\boldsymbol{e}_{i_{r-s}}\rangle.\end{aligned}$$

The last term is just $F(\boldsymbol{A},\boldsymbol{B}:r,s+1)$. □

Making $\boldsymbol{B}=-\boldsymbol{A}$ in Lemma 4.1, we have:

Corollary 4.2 For r,s satisfying the condition in Lemma 4.1 and $\boldsymbol{A}\in\mathcal{L}(\mathcal{V})$, recurrence formula

$$I_{r-s}\mathrm{tr}(-\boldsymbol{A})^{s}=G(\boldsymbol{A};r,s)-G(\boldsymbol{A};r,s+1) \tag{4.3}$$

holds, where

$$\begin{aligned}G(\boldsymbol{A};r,s)&\equiv F(\boldsymbol{A},-\boldsymbol{A};r,s)\\
&=\frac{1}{(r-s)!}\langle\boldsymbol{\alpha}^{i_1}\wedge\cdots\wedge\boldsymbol{\alpha}^{i_{r-s+1}},\boldsymbol{A}\boldsymbol{e}_{i_1}\wedge\cdots\wedge\boldsymbol{A}\boldsymbol{e}_{i_{r-s}}\wedge(-\boldsymbol{A})^{s}\boldsymbol{e}_{i_{r-s+1}}\rangle.\quad\square\end{aligned} \tag{4.4}$$

For $s = r$, expressions (4.2) and (4.4) are still meaningful and in the form:

$$F(\boldsymbol{A}, \boldsymbol{B}; r, r) = \langle \boldsymbol{\alpha}^i, (-\boldsymbol{A})^{r-1}\boldsymbol{B}\boldsymbol{e}_i\rangle = \mathrm{tr}((-\boldsymbol{A})^{r-1}\boldsymbol{B}), \tag{4.5}$$

$$G(\boldsymbol{A}; r, r) = \mathrm{tr}(-\boldsymbol{A})^r. \tag{4.6}$$

V. Newton's Formulae

Newton's formulae are important in the theory of symmetric polynomials. All existing proofs of these formulae are complicated[5-8]. Here, using the key recurrence formula, we offer a simpler proof.

Theorem 5.1 For any endomorphism $\boldsymbol{A}$, Newton's formulae

$$rI_r + \sum_{s=1}^{r} I_{r-s}\mathrm{tr}(-\boldsymbol{A})^s = 0, \quad 1 \le r \le n \tag{5.1}$$

and

$$\sum_{s=r-n}^{r} I_{r-s}\mathrm{tr}(-\boldsymbol{A})^s = 0, \quad r > n \tag{5.2}$$

hold.

Proof Obviously, (5.1) holds for $r = 1$. For the remaining cases with $1 < r \le n$, Corollary 4.2 can be used. To this end, using (4.4) to rewrite expression (2.17) for I_r, we have

$$-rI_r = G(\boldsymbol{A}; r, 1). \tag{5.3}$$

Summing (4.3) for s from 1 to $r-1$, and adding (4.6) to the resulting expression, we get

$$\begin{aligned}\sum_{s=1}^{r} I_{r-s}\mathrm{tr}(-\boldsymbol{A})^s &= \sum_{s=1}^{r} G(\boldsymbol{A}; r, s) - \sum_{s=1}^{r-1} G(\boldsymbol{A}; r, s+1)\\ &= \sum_{s=1}^{r} G(\boldsymbol{A}; r, s) - \sum_{s=2}^{r} G(\boldsymbol{A}; r, s)\\ &= G(\boldsymbol{A}; r, 1).\end{aligned}$$

Inserting this result into (5.3) yields (5.1). The proof of the first Newton's formula is completed. Similarly, summing (4.3) for s from $r-n$ to $r-1$, and adding (4.6) to the result, we obtain

$$\begin{aligned}\sum_{s=r-n}^{r} I_{r-s}\mathrm{tr}(-\boldsymbol{A})^s &= \sum_{s=r-n}^{r} G(\boldsymbol{A};r,s) - \sum_{s=r-n}^{r-1} G(\boldsymbol{A};r,s+1)\\ &= G(\boldsymbol{A};r,r-n)\\ &= \frac{1}{n!}\langle \alpha^{i_1}\wedge\cdots\wedge\alpha^{i_{n+1}}, \boldsymbol{A}e_{i_1}\wedge\cdots\wedge \boldsymbol{A}e_{i_n}\wedge(-\boldsymbol{A})^{r-n}e_{i_{n+1}}\rangle\\ &= 0.\end{aligned}$$

This is just the second Newton's formula (5.2). The last equality is the consequence of the fact that any $(n+1)$-form is zero.

VI. Derivatives of Principal Invariants

Recently Carlson and Hoger[9] derived in an indirect way the expression for the derivatives of the principal invariants of an endomorphism $\boldsymbol{A} \in \mathcal{L}(\mathcal{V})$:

$$DI_r(\boldsymbol{A}) \equiv \frac{\partial I_r}{\partial \boldsymbol{A}} = \Big[\sum_{\beta=0}^{r-1} I_\beta(-\boldsymbol{A})^{r-\beta-1}\Big]^T, \quad r = 1,2,\ldots,n, \tag{6.1}$$

generalizing the derivation of Refs. [1], [4] and [10]. Ref. [1] assumes $\boldsymbol{A}$ symmetric, and Refs. [4] and [10] assume $\boldsymbol{A}$ invertible. Using exterior algebra, we derive expression (6.1) in a direct way without any restriction.

Assume $I_r(\boldsymbol{A})$ $(r = 1,2,\ldots,n)$ Fréchet differentiable. Then the directional derivative of $I_r(\boldsymbol{A})$ along $\boldsymbol{B} \in \mathcal{L}(\mathcal{V})$ is

$$DI_r(\boldsymbol{A})[\boldsymbol{B}] = \frac{d}{dt} I_r(\boldsymbol{A}+t\boldsymbol{B})\Big|_{t=0}. \tag{6.2}$$

On the other hand, we have

$$DI_r(\boldsymbol{A})[\boldsymbol{B}] = \mathrm{tr}[(DI_r(\boldsymbol{A}))^T\boldsymbol{B}]. \tag{6.3}$$

If we can transform the right-hand side of (6.2) into the form $\mathrm{tr}(\boldsymbol{C}^T\boldsymbol{B})$ $(\boldsymbol{C} \in \mathcal{L}(\mathcal{V}))$, then the derivative of I_r at $\boldsymbol{A}$ is readily found to be $DI_r(\boldsymbol{A}) = \boldsymbol{C}$.

Theorem 6.1 The derivative of the r^{th} principal invariant I_r of an endomorphism $\boldsymbol{A}$ is given by

$$DI_r(\boldsymbol{A}) = \sum_{s=1}^{r} I_{r-s}(-\boldsymbol{A}^T)^{s-1}. \quad \square \tag{6.4}$$

Expression (6.4) coincides with (6.1), if the summation parameter $\beta = r - s$ is used instead of s.

Proof We calculate the right-hand side of (6.2):

$$\frac{d}{dt}I_r(\boldsymbol{A}+t\boldsymbol{B})\Big|_{t=0} = \frac{d}{dt}\Big[\frac{1}{r!}\langle\alpha^{i_1}\wedge\cdots\wedge\alpha^{i_p},(\boldsymbol{A}+t\boldsymbol{B})e_{i_1}$$
$$\wedge\cdots\wedge(\boldsymbol{A}+t\boldsymbol{B})e_{i_r}\rangle\Big]\Big|_{t=0}$$
$$= \frac{1}{r!}\sum_{p=1}^{r}\langle\alpha^{i_1}\wedge\cdots\wedge\alpha^{i_p}\wedge\cdots\wedge\alpha^{i_r},$$
$$\boldsymbol{A}e_{i_1}\wedge\cdots\wedge\boldsymbol{B}e_{i_p}\wedge\cdots\wedge\boldsymbol{A}e_{i_r}\rangle$$
$$= \frac{1}{(r-1)!}\langle\alpha^{i_1}\wedge\cdots\wedge\alpha^{i_r},\boldsymbol{A}e_{i_1}\wedge\cdots\wedge\boldsymbol{A}e_{i_{r-1}}\wedge\boldsymbol{B}e_{i_r}\rangle$$
$$= F(\boldsymbol{A},\boldsymbol{B};r,1). \tag{6.5}$$

For $r = 1$, making use of (4.5), (6.5) assumes the form

$$\frac{d}{dt}I_1(\boldsymbol{A}+t\boldsymbol{B})\Big|_{t=0} = F(\boldsymbol{A},\boldsymbol{B};1,1) = \mathrm{tr}(\boldsymbol{I}\boldsymbol{B}).$$

Thus

$$DI_1(\boldsymbol{A}) = \boldsymbol{I} = I_0(-\boldsymbol{A}^T)^{1-1}. \tag{6.6}$$

Hence, (6.4) holds for $r = 1$. For the remaining cases $1 < r \leq n$, Lemma 4.1 can be used. To this end, summing (4.1) for s from 1 to $r-1$, and adding (4.5) to the resulting expression, we get

$$\sum_{s=1}^{r} I_{r-s}\mathrm{tr}((-\boldsymbol{A})^{s-1}\boldsymbol{B}) = \sum_{s=1}^{r} F(\boldsymbol{A},\boldsymbol{B};r,s) - \sum_{s=1}^{r-1} F(\boldsymbol{A},\boldsymbol{B};r,s+1)$$
$$= F(\boldsymbol{A},\boldsymbol{B};r,1).$$

Comparing this result with (6.5), we have

$$\frac{d}{dt}I_r(\boldsymbol{A}+t\boldsymbol{B})\Big|_{t=0} = \mathrm{tr}\Big[\sum_{s=1}^{r} I_{r-s}(-\boldsymbol{A})^{s-1}\boldsymbol{B}\Big].$$

According to (6.2) and (6.3), this expression gives the transpose of (6.4):

$$(DI_r(\boldsymbol{A}))^T = \sum_{s=1}^{r} I_{r-s}(-\boldsymbol{A})^{s-1}.$$

Together with (6.6), it completes the proof of the theorem.

References

[1] Ericksen, J. L., *Tensor Field, Handbuch der Physik*, **3**, 1, Springer, New York, 1960.

[2] Kotchin, N. E., *Vector Calculus and Elements of Tensor Calculus* (in Russian), Press of Soviet Acad. of Sciences, Moscow, 1961.

[3] Bowen, R. M. and Wang, C. C., *Introduction to Vectors and Tensors*, Plenum Press, New York, 1976.

[4] Truesdell, C. and Noll, W., *The Non-linear Field Theories of Mechanics, Handbuch der Physik*, **3**, 3, Springer, Berlin, 1965.

[5] Gantmacher, F. R., *The Theory of Matrices*, **1**, Chelsea Publishing, New York, 1959.

[6] Chrystal, G., *Algebra*, **1**, Chelsea Publishing, New York, 1964.

[7] Algebra Group of Dept. of Math.-Mech. of Peking University, *Lecture Notes on High Algebra* (in Chinese), High Education Publisher, Beijing, 1965.

[8] Grov, L. C., *Algebra*, Acad. Press, New York, 1983.

[9] Carlson, D. E. and Hoger, A., On the derivatives of the principal invariants of a second-order tensor, *J. Elasticity*, **16** (1986), 221–224.

[10] Gurtin, M. E., *An Introduction to Continuum Mechanics*, Acad. Press, New York, 1981.