Variation of induced linear operators[☆]

Rajendra Bhatia [a,*], J.A. Dias da Silva [b]

[a] *Indian Statistical Institute, 7 S.J.S. Sansanwal Marg, New Delhi 110016, India*
[b] *Departamento de Matemática, da Universidade de Lisboa-CELC, Av Gama Pinto 2, 1699 Lisboa Codex, Portugal*

Submitted by M.J. Tsatsomeros

Dedicated to T. Ando

Abstract

Let V be an n-dimensional inner product space. Let λ be an irreducible character of the symmetric group S_m, and let V_λ be the symmetry class of tensors associated with it. Let A be a linear operator on V and let $K_\lambda(A)$ be the operator it induces on V_λ. We obtain an explicit expression for the norm of the derivative of the map $A \to K_\lambda(A)$ in terms of the singular values of A. Two special cases of this problem—antisymmetric and symmetric tensor products—have been studied earlier, and our results reduce to the earlier ones in these cases.

Keywords: Symmetry class of tensors; Induced linear operator; Derivative; Norm; Positive linear operator

1. Introduction

Let $\mathscr{L}(V)$ be the space of bounded linear operators on a Hilbert space V. The norm of an element A of $\mathscr{L}(V)$ is defined as

$$\|A\| = \sup\left\{\|Av\| : v \in V, \|v\| = 1\right\}.$$

In this paper V is finite-dimensional. Then $\|A\|$ is the largest singular value of A.

☆ This work was supported by grants from the National Board for Higher Mathematics, India, and Fundação para a Ciência e Tecnologia, Portugal.

* Corresponding author.
E-mail address: rbh@isid.ac.in (R. Bhatia).

Functions $f : \mathscr{L}(V) \to \mathscr{L}(W)$ are studied often in different contexts. Sometimes f is defined on an open subset of $\mathscr{L}(V)$ such as the set of invertible operators. In perturbation theory, numerical analysis, and physics, one often wants to know the effect of changes in A on $f(A)$. When the map f is differentiable, it is helpful to have estimates of the norm of its derivative. The derivative of f at A is a linear map $Df(A)$ from $\mathscr{L}(V)$ into $\mathscr{L}(W)$ and its norm is defined as

$$\|Df(A)\| = \sup\left\{\|Df(A)(B)\| : B \in \mathscr{L}(V), \|B\| = 1\right\}. \tag{1}$$

Estimates of this lead to first-order perturbation bounds for f. See the discussion in [1, Chapter X] and the papers [4,6,15,17] for different perspectives on this question.

Recall that

$$Df(A)(B) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f(A + tB). \tag{2}$$

Since A and B do not always commute several difficulties arise in estimating $\|Df(A)\|$. Finding *exact* values of $\|Df(A)\|$ is even more difficult, and very few such results are known. Some of them have led to intriguing questions [5,7].

In this paper we obtain exact formulas for $\|Df(A)\|$ when $f(A)$ is any of the operators induced by A on a symmetry class of tensors corresponding to the (full) symmetric group. Two special cases have been studied earlier [2,3]. To put our results in perspective we first recall these results. We need some basic facts, notations, and terminology of multilinear algebra. Further details may be found in [12] or [13].

Let $\dim V = n$, and for $A \in \mathscr{L}(V)$ let

$$\nu_1 \geqslant \nu_2 \geqslant \cdots \geqslant \nu_n \geqslant 0$$

be the singular values of A. Let $\otimes^m V = V \otimes V \otimes \cdots \otimes V$ be the m-fold tensor power of V and let $\otimes^m A$ be the corresponding tensor power of A. It is easy to see that [2]

$$\|D \otimes^m (A)\| = m\|A\|^{m-1}. \tag{3}$$

Now let $1 \leqslant m \leqslant n$, let $\wedge^m V$ be the subspace of $\otimes^m V$ consisting of antisymmetric tensors, and let $\wedge^m A$ be the restriction of $\otimes^m A$ to this subspace. This is sometimes called the exterior power of A or the Grassmann power of A. In [2] it was shown that

$$\|D \wedge^m (A)\| = s_{m-1}(\nu_1, \nu_2, \ldots, \nu_m), \tag{4}$$

where s_{m-1} is the $(m-1)$th elementary symmetric polynomial in $\nu_1, \ldots, \nu_m$; i.e.,

$$s_{m-1}(\nu_1, \ldots, \nu_m) = \sum_{j=1}^{m} \prod_{\substack{i=1 \\ i \neq j}}^{m} \nu_i. \tag{5}$$

The corresponding problem for the symmetric tensor power $\vee^m A$ (obtained by restricting $\otimes^m A$ to the space $\vee^m V$ of symmetric tensors) was studied in [3], where it was shown that

$$\|D \vee^m (A)\| = m\|A\|^{m-1} = m\nu_1^{m-1}, \tag{6}$$

and a speculation was made about a general result that would subsume (4) and (6). The precise formulation and proof of such a result is the principal outcome of this paper.

Let S_m be the symmetric group of degree m. Each element σ of S_m gives rise to a linear operator $P(\sigma)$ on $\otimes^m V$. This is defined as

$$P(\sigma)(v_1 \otimes v_2 \otimes \cdots \otimes v_m) = v_{\sigma^{-1}(1)} \otimes v_{\sigma^{-1}(2)} \otimes \cdots \otimes v_{\sigma^{-1}(m)} \tag{7}$$

on decomposable tensors and then extended linearly to all of $\otimes^m V$.

The map $\sigma \to P(\sigma)$ is a unitary representation of S_m in $\otimes^m V$. In other words, $P(\sigma_1)P(\sigma_2) = P(\sigma_1\sigma_2)$ and $P(\sigma)^{-1} = P(\sigma^{-1}) = P(\sigma)^*$.

Let G be a subgroup of S_m, and let λ be an irreducible character of G. Let

$$T(G, \lambda) = \frac{\lambda\,(\mathrm{id})}{|G|} \sum_{\sigma \in G} \lambda(\sigma) P(\sigma), \tag{8}$$

where id stands for the identity element and $|G|$ for the order of the group G. This linear operator on $\otimes^m V$ is an orthoprojector and is called a *symmetriser map*. Its range is called the *symmetry class of tensors* associated with λ and G.

We will study symmetry classes associated with the full symmetric group $G = S_m$. Then the alternating character $\lambda(\sigma) = \varepsilon_\sigma$ (the signature of the permutation σ) leads to the symmetry class $\wedge^m V$; whereas the principal character $\lambda(\sigma) \equiv 1$ leads to the symmetry class $\vee^m V$.

There is a standard canonical correspondence between irreducible characters of S_m and partitions of the integer m [10]. We use the same symbol λ to denote an irreducible character and the corresponding partition. Recall that a partition π of m is a k-tuple of positive integers $\pi = (\pi_1, \ldots, \pi_k)$ such that $\pi_1 \geqslant \cdots \geqslant \pi_k$ and $\pi_1 + \cdots + \pi_k = m$. For convenience we think of a partition of m also as an m-tuple with nonnegative integer entries by putting some zeros at the end if necessary. We adopt a similar convention for decreasing sequences of nonnegative real numbers. If $\lambda = (1, \ldots, 1)$, then $V_\lambda(S_m) = \wedge^m V$; and if $\lambda = (m, 0, \ldots, 0)$, then $V_\lambda(S_m) = \vee^m V$.

Let $\ell(\lambda)$ be the length of the partition λ—this is the number of nonzero entries in λ. For each $1 \leqslant t \leqslant m$ we denote by $\lambda_{(t)}$ the m-tuple defined as

$$\lambda_{(t)} = \begin{cases} (\lambda_1, \ldots, \lambda_{t-1}, \lambda_t - 1, \lambda_{t+1}, \ldots, \lambda_m) & \text{if } t \leqslant \ell(\lambda), \\ (\lambda_1, \ldots, \lambda_{t-1}, -\infty, \lambda_{t+1}, \ldots, \lambda_m) & \text{if } \ell(\lambda) < t. \end{cases} \tag{9}$$

Given any n-tuple of nonnegative real numbers $(\nu_1, \nu_2, \ldots, \nu_n)$, and a k-tuple $(\gamma_1, \ldots, \gamma_k)$ whose entries are either nonnegative integers or $-\infty$, we define ν^γ as

$$\nu^\gamma = \nu_1^{\gamma_1} \nu_2^{\gamma_2} \cdots \nu_n^{\gamma_n}$$

with the convention that $a^0 = 1$ and $a^{-\infty} = 0$ for every nonnegative a.

Now let λ be a partition of m and let $\ell(\lambda) \leqslant n$. Put

$$S_{\lambda,\nu} = \lambda_1 \nu^{\lambda(1)} + \lambda_2 \nu^{\lambda(2)} + \cdots + \lambda_m \nu^{\lambda(m)}. \tag{10}$$

Note that if $\lambda = (1, 1, \ldots, 1)$, then

$$S_{\lambda,\nu} = \sum_{j=1}^{m} \prod_{\substack{i=1 \\ i \neq j}}^{m} \nu_i = s_{m-1}(\nu_1, \ldots, \nu_m).$$

If $\lambda = (m, 0, \ldots, 0)$, then

$$S_{\lambda,\nu} = m\nu_1^{m-1}.$$

Now return to symmetry classes of tensors. It is well known that $V_\lambda(S_m) \neq \{0\}$ if and only if $\ell(\lambda) \leqslant n$; see [14]. Given any $A \in \mathscr{L}(V)$ we denote by $K_\lambda(A)$ the restriction of the operator $\otimes^m A$ to the subspace $V_\lambda(S_m)$. This is called the operator *induced* by A on the symmetry class $V_\lambda(S_m)$. Our principal result is the following theorem.

Theorem 1. *Let V be an n-dimensional Hilbert space. Let m be a positive integer. Let λ be a partition of m such that $\ell(\lambda) \leqslant n$. Let $A \to K_\lambda(A)$ be the map that associates to each element A of $\mathscr{L}(V)$ the induced operator $K_\lambda(A)$ on the symmetry class $V_\lambda(S_m)$. Then the norm of the derivative of this map at A is given by the formula*

$$\|DK_\lambda(A)\| = S_{\lambda,\nu}, \tag{11}$$

where $\nu_1 \geqslant \nu_2 \geqslant \cdots \geqslant \nu_n$ are the singular values of A, and $S_{\lambda,\nu}$ is the polynomial defined by (10).

Note that Theorem 1 includes as very special cases the results (4) and (6) obtained in [2,3].

To guide the reader through the proof we highlight its salient features. Let A have the singular value decomposition $A = U_1 P U_2$. Using the unitary invariance of the norm and of the singular values one sees that $\|DK_\lambda(A)\| = \|DK_\lambda(P)\|$. So, one may replace A by the positive diagonal matrix P. Then one observes that $DK_\lambda(P)$ is a positive linear map between two matrix algebras. By a general theorem of Russo and Dye, such a map between any two unital C^*-algebras attains its norm at the identity I. This simplifies our calculations immensely because we do not have to consider arbitrary A and B in expression (2) for derivatives. Even after this simplification some difficulties remain. While in the special examples $\wedge^m V$ and $\vee^m V$ good orthonormal bases corresponding to the standard basis in V can be found immediately, this is not the case in other symmetry classes. We explain how a suitable basis may be chosen for our purposes. This choice leads to a partition of m; and finally we have to study the relation between this partition and λ, and the corresponding functions $S_{\lambda,\nu}$. Here we prove a majorisation theorem that is of interest in its own right.

The idea of replacing A by P in calculating $\|D \wedge^k (A)\|$ occurs in [2]. It is also shown there that $\|D \wedge^k (P)\| = \|D \wedge^k (P)(I)\|$. The idea of proving the same result using completely positive maps is due to Sunder [16].

2. Preliminaries

Given a symmetriser map $T(G, \lambda)$ let

$$v_1 * v_2 * \cdots * v_m = T(G, \lambda)(v_1 \otimes v_2 \otimes \cdots \otimes v_m).$$

These vectors belong to $V_\lambda(G)$ and are called decomposable symmetrised tensors.

Let $\Gamma_{m,n}$ be the set of all maps from the set $\{1, \ldots, m\}$ into the set $\{1, \ldots, n\}$. This set can be identified with the collection of all multiindices $\{(i_1, \ldots, i_m) : 1 \leqslant i, j \leqslant n\}$. If $\alpha \in \Gamma_{m,n}$, this correspondence associates the index $(\alpha(1), \ldots, \alpha(m))$ with it. We order $\Gamma_{m,n}$ by the lexicographic order.

Every subgroup G of S_m acts on $\Gamma_{m,n}$ by the action $(\sigma, \alpha) \to \alpha\sigma^{-1}$, $\sigma \in G, \alpha \in \Gamma_{m,n}$. The subgroup G_α of G defined as

$$G_\alpha = \left\{\sigma \in G : \alpha\sigma = \alpha\right\}$$

is called the *stabiliser* of α.

Let $\{e_1, \ldots, e_n\}$ be a basis of V. Then $\{e_\alpha^\otimes := e_{\alpha(1)} \otimes \cdots \otimes e_{\alpha(m)} : \alpha \in \Gamma_{m,n}\}$ is a basis for $\otimes^m V$. Hence the set

$$\left\{e_\alpha^* := T(\lambda, G)e_\alpha^\otimes : \alpha \in \Gamma_{m,n}\right\}$$

spans the space $V_\lambda(G)$. However, the elements of this set need not be linearly independent. Some of them may even be zero. Let

$$\Omega = \Omega_\lambda = \left\{\alpha \in \Gamma_{m,n} : \sum_{\sigma \in G_\alpha} \lambda(\sigma) \neq 0\right\}. \tag{12}$$

It is easy to see that

$$\|e_\alpha\|^2 = \frac{\lambda\,(\mathrm{id})}{|G|} \sum_{\sigma \in G_\alpha} \lambda(\sigma).$$

So the set $\{e_\alpha^* : \alpha \in \Omega\}$ consists of the nonzero elements of $\{e_\alpha^* : \alpha \in \Gamma_{m,n}\}$.

Let Δ be the system of distinct representatives for the set $\Gamma_{m,n}/G$, constructed by choosing the smallest element (in the lexicographic order) from each orbit. Let

$$\overline{\Delta} = \overline{\Delta}_\lambda = \Delta \cap \Omega_\lambda.$$

It can be proved that $\{e_\alpha^* : \alpha \in \overline{\Delta}\}$ is a linearly independent set. Since the set $\{e_\alpha^* : \alpha \in \Omega\}$ spans $V_\lambda(G)$ there exists a set $\widehat{\Delta}$ such that $\overline{\Delta} \subseteq \widehat{\Delta} \subseteq \Omega$ and

$$\left\{e_\alpha^* : \alpha \in \widehat{\Delta}\right\} \tag{13}$$

is a basis for $V_\lambda(G)$, not necessarily orthonormal. See [13] for details.

Each element α of $\Gamma_{m,n}$ gives rise to a partition of m in the following way. Let range $\alpha = \{i_1, \ldots, i_\ell\}$, where $i_1, \ldots, i_\ell$ are labelled in such a way that

$$|\alpha^{-1}(i_1)| \geqslant |\alpha^{-1}(i_2)| \geqslant \cdots \geqslant |\alpha^{-1}(i_\ell)|.$$

Then

$$\mu^{(\alpha)} := \left(|\alpha^{-1}(i_1)|, |\alpha^{-1}(i_2)|, \ldots, |\alpha^{-1}(i_\ell)|\right) \tag{14}$$

is a partition of m of length ℓ.

On the set of partitions of m, we define a partial order $\prec$ as follows: we say that $\mu \prec \lambda$ if for all $1 \leqslant k \leqslant m$

$$\sum_{j=1}^{k} \mu_j \leqslant \sum_{j=1}^{k} \lambda_j.$$

(This is the usual *majorisation* order between m-tuples [1] when we identify partitions with m-tuples.) We will need the following theorem of Merris [14].

Theorem 2 (Merris). *Let λ be a partition of m and α an element of $\Gamma_{m,n}$. Let Ω_λ and $\mu^{(\alpha)}$ be as defined in* (12) *and* (14). *Then $\alpha \in \Omega_\lambda$ if and only if $\mu^{(\alpha)} \prec \lambda$.*

Let λ, μ be two partitions of m. We say that $\mu \lhd \lambda$, if there exist indices $i, j \in \{1, \ldots, m\}$ such that

(i) $i < j$;

(ii) $\mu_i = \lambda_i - 1, \mu_j = \lambda_j + 1$, and $\lambda_k = \mu_k$ for $k \neq i, j$;

(iii) either $i = j - 1$ or $\mu_i = \mu_j$.

We will need the following result [10, p. 24].

Proposition 3. *If $\mu \prec \lambda$, then there exists a sequence of partitions $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}$ such that*

$$\mu = \lambda^{(1)} \lhd \lambda^{(2)} \lhd \cdots \lhd \lambda^{(k)} = \lambda.$$

For brevity we say that $A \in \mathscr{L}(V)$ is *positive* if it is positive semidefinite. A linear map $\Phi : \mathscr{L}(V) \to \mathscr{L}(W)$ is called *positive* if it maps positive elements of $\mathscr{L}(V)$ into positive elements of $\mathscr{L}(W)$. We say that Φ is unital if $\Phi(I) = I$.

Positive linear maps Φ enjoy a very special property: $\|\Phi\| = \|\Phi(I)\|$. This is a consequence of the well-known Russo–Dye Theorem [11] valid in C^*-algebras.

3. Proofs

Let $V_\lambda = V_\lambda(S_m)$ be the symmetry class of tensors associated with λ and let $K_\lambda : \mathscr{L}(V) \to \mathscr{L}(V_\lambda)$ be the induced map. For brevity let $D_\lambda(A, B) = DK_\lambda(A)(B)$, the image of B under the derivative $DK_\lambda(A)$. Then $D_\lambda(A, B)$ is the restriction to V_λ of the operator on $\otimes^m V$ defined as

$$\begin{aligned} D(A, B) := {} & B \otimes A \otimes A \otimes \cdots \otimes A + A \otimes B \otimes A \otimes \cdots \otimes A \\ & + \cdots + A \otimes \cdots \otimes A \otimes B. \end{aligned}$$

Note that if A and B are positive, then so is $D(A, B)$.

Let $A = U_1 P U_2$ be the singular value decomposition of A. Using unitary invariance of the norm and the fact that $K_\lambda(U)$ is unitary if U is unitary, we see that

$$\|DK_\lambda(A)\| = \|DK_\lambda(P)\|.$$

From the description above it is clear that $DK_\lambda(P)$ is a positive linear map. Hence by the Russo–Dye Theorem

$$\|DK_\lambda(A)\| = \|D_\lambda(P, I)\|. \tag{15}$$

So we have to calculate the maximum eigenvalue of $D_\lambda(P, I)$. We will do this by finding a basis for V_λ in which $D_\lambda(P, I)$ is diagonal. Then the diagonal entries of this matrix are the eigenvalues of $D_\lambda(P, I)$; our basis need not be orthonormal for this.

Let $\alpha \in \Gamma_{m,n}$ and let $\mu^{(\alpha)}$ be the partition of length ℓ associated with α as in (14). Let

$$\nu_\alpha = (\nu_{i_1}, \ldots, \nu_{i_\ell}) \tag{16}$$

be the largest (in the lexicographic order) sequence such that $(i_1, \ldots, i_\ell)$ satisfies (14). (For example, if $\ell = 4$ and $|\alpha^{-1}(6)| = |\alpha^{-1}(7)| > |\alpha^{-1}(4)| = |\alpha^{-1}(3)|$, then $\nu_\alpha = (\nu_6, \nu_7, \nu_3, \nu_4)$.)

Given any partition λ, let ω_λ be the element of $\Gamma_{m,n}$ defined as

$$\omega_\lambda = \big(\underbrace{1, \ldots, 1}_{\lambda_1 \text{ times}}, \underbrace{2, \ldots, 2}_{\lambda_2 \text{ times}}, \ldots, \underbrace{\ell(\lambda), \ldots, \ell(\lambda)}_{\lambda_{\ell(\lambda)} \text{ times}}\big). \tag{17}$$

Then clearly

$$\mu^{(\omega_\lambda)} = (\lambda_1, \ldots, \lambda_{\ell(\lambda)}) = \lambda, \quad \nu_{\omega_\lambda} = (\nu_1, \ldots, \nu_{\ell(\lambda)}). \tag{18}$$

Proposition 4. *Let P be a positive linear operator on V, and suppose $E = \{e_1, \ldots, e_n\}$ is an orthonormal basis for V in which the matrix of P is diagonal with diagonal entries $\nu_1 \geqslant \cdots \geqslant \nu_n$. Let $\{e^*_\alpha : \alpha \in \widehat{\Delta}\}$ be a basis for V_λ as in* (13). *Then in this basis $D_\lambda(P, I)$ is diagonal and its (α, α) entry is*

$$D_\lambda(P, I)_{\alpha,\alpha} = \sum_{j=1}^{m} \prod_{\substack{i=1 \\ i \neq j}}^{m} \nu_{\alpha(i)} = S_{\mu^{(\alpha)}, \nu_\alpha}, \quad \alpha \in \widehat{\Delta}. \tag{19}$$

Proof. Recall that for any $\alpha \in \Gamma_{m,n}$

$$e^*_\alpha = \frac{\lambda(\mathrm{id})}{m!} \sum_{\sigma \in S_m} \lambda(\sigma) e^{\otimes}_{\alpha\sigma}.$$

Note that

$$D_\lambda(P, I) \left(\sum_\sigma \lambda(\sigma) e^{\otimes}_{\alpha\sigma} \right)$$

$$
\begin{aligned}
&= D(P,I)\left(\sum_\sigma \lambda(\sigma) e^{\otimes}_{\alpha\sigma}\right)\\
&= (I\otimes P\otimes P\otimes\cdots\otimes P)\left(\sum_\sigma \lambda(\sigma) e^{\otimes}_{\alpha\sigma}\right)\\
&\quad + \cdots + (P\otimes P\otimes\cdots\otimes I)\left(\sum_\sigma \lambda(\sigma) e^{\otimes}_{\alpha\sigma}\right)\\
&= \left(\sum_\sigma \lambda(\sigma)\prod_{i=2}^{m} \nu_{\alpha\sigma(i)} e^{\otimes}_{\alpha\sigma}\right) + \cdots + \left(\sum_\sigma \lambda(\sigma)\prod_{i=1}^{m-1} \nu_{\alpha\sigma(i)} e^{\otimes}_{\alpha\sigma}\right).
\end{aligned}
$$

For each $1 \leqslant k \leqslant m$

$$
\prod_{\substack{i=1\\ i\neq j}}^{m} \nu_{\alpha\sigma(i)} = \prod_{\substack{i=1\\ i\neq\sigma(j)}}^{m} \nu_{\alpha(i)}.
$$

This shows that

$$
D_\lambda(P,I)e^*_\alpha = \left(\sum_{j=1}^{m}\prod_{\substack{i=1\\ i\neq j}}^{m} \nu_{\alpha(i)}\right) e^*_\alpha.
$$

Thus the matrix of $D_\lambda(P,I)$ in the basis $\{e^*_\alpha : \alpha \in \widehat{\Delta}\}$ is diagonal with entries given in (19).

By definitions (14) and (19)

$$
\begin{aligned}
\sum_{j=1}^{m}\prod_{\substack{i=1\\ i\neq j}}^{m} \nu_{\alpha(i)} &= \mu_1^{(\alpha)} \nu_{i_1}^{\mu_1^{(\alpha)}-1} \nu_{i_2}^{\mu_2^{(\alpha)}} \cdots \nu_{i_\ell}^{\mu_\ell^{(\alpha)}} + \mu_2^{(\alpha)} \nu_{i_1}^{\mu_1^{(\alpha)}} \nu_{i_2}^{\mu_2^{(\alpha)}-1} \cdots \nu_{i_\ell}^{\mu_\ell^{(\alpha)}}\\
&\quad + \cdots + \mu_\ell^{(\alpha)} \nu_{i_1}^{\mu_1^{(\alpha)}} \nu_{i_2}^{\mu_2^{(\alpha)}} \cdots \nu_{i_\ell}^{\mu_\ell^{(\alpha)}-1}\\
&= S_{\mu^{(\alpha)}\nu_\alpha}. \qquad \square
\end{aligned}
$$

Proposition 5. *Let λ and μ be partitions of m, and let $\nu_1 \geqslant \cdots \geqslant \nu_m \geqslant 0$ be any decreasing sequence of nonnegative numbers. If $\mu \prec \lambda$, then $S_{\mu,\nu} \leqslant S_{\lambda,\nu}$.*

Proof. By Proposition 3, it is enough to prove this when $\mu \lhd \lambda$. Assume $\nu_m > 0$; the general case follows from this by continuity.

Use the notations as in definition of $\mu \lhd \lambda$, before Proposition 3. Then for $k \neq i, j$

$$
\lambda_k \nu^{\lambda(k)} = \lambda_k \nu_k^{\lambda_k - 1} \prod_{r\neq k} \nu_r^{\lambda_r}
$$

$$
\begin{aligned}
&= \mu_k \nu_k^{\mu_k - 1} \prod_{r \neq k} \nu_r^{\lambda_r} \\
&\geqslant \mu_k \nu_k^{\mu_k - 1} \left[\prod_{r \neq i,j,k} \nu_r^{\lambda_r} \right] \nu_i^{\lambda_i - 1} \nu_j^{\lambda_j + 1} \\
&= \mu_k \nu_k^{\mu_k - 1} \prod_{r \neq k} \nu_r^{\mu_r} \\
&= \mu_k \nu^{\mu_{(k)}}.
\end{aligned}
$$

Next note that

$$
\begin{aligned}
&\left[\lambda_i \nu^{\lambda_{(i)}} + \lambda_j \nu^{\lambda_{(j)}}\right] - \left[\mu_i \nu^{\mu_{(i)}} + \mu_j \nu^{\mu_{(j)}}\right] \\
&\quad = \frac{\nu^\lambda}{\nu_i^2 \nu_j} \left[\lambda_i \nu_i \nu_j + \lambda_j \nu_i^2 - (\lambda_i - 1)\nu_j^2 - (\lambda_j + 1)\nu_i \nu_j \right] \\
&\quad = \frac{\nu^\lambda}{\nu_i^2 \nu_j} \left[(\lambda_i - 1)\nu_i \nu_j - (\lambda_i - 1)\nu_j^2 + \lambda_j (\nu_i^2 - \nu_i \nu_j) \right] \\
&\quad \geqslant \frac{\nu^\lambda}{\nu_i^2 \nu_j} \left[(\lambda_i - 1)(\nu_i \nu_j - \nu_j^2) \right] \quad (\text{since } \nu_i^2 \geqslant \nu_i \nu_j) \\
&\quad \geqslant 0 \quad (\text{since } \lambda_i = \mu_i + 1 \geqslant 1).
\end{aligned}
$$

Taken together, the two inequalities we have obtained, prove the proposition. □

Let $\lambda = (\lambda_1, \ldots, \lambda_\ell, 0, \ldots, 0)$ be a partition of m. Then we denote by λ^* the partition of $m - \ell$ given as

$$\lambda^* = (\lambda_1^*, \ldots, \lambda_{m-\ell}^*),$$

where $\lambda_i^* = \lambda_i - 1$ if $i \leqslant \ell$ and $\lambda_i^* = 0$ if $i > \ell$.

Given an m-tuple $(\theta_1, \ldots, \theta_m)$ of real numbers we denote by $\theta^\downarrow$ its decreasing rearrangement; i.e., $\theta^\downarrow = (\theta_1^\downarrow, \ldots, \theta_m^\downarrow)$, where $\theta_1^\downarrow \geqslant \cdots \geqslant \theta_m^\downarrow$ are the numbers $\theta_1, \ldots, \theta_m$ rearranged. We use the notation $\nu \geqslant \theta$ to mean $\nu_j \geqslant \theta_j$ for all j.

Proposition 6. *Let ν, θ be m-tuples of nonnegative real numbers such that ν is decreasing and $\nu \geqslant \theta^\downarrow$. Then for every partition λ of m we have $S_{\lambda,\nu} \geqslant S_{\lambda,\theta}$.*

Proof. Note first that

$$\nu^\lambda \geqslant (\theta^\downarrow)^\lambda \geqslant \theta^\lambda. \tag{20}$$

For any m-tuple $\rho = (\rho_1, \ldots, \rho_m)$ of nonnegative real numbers let

$$T_{\lambda,\rho} = \sum_{i=1}^{m} \rho^{\lambda_{(i)}}.$$

Then, bearing in mind that $\lambda_{(i)}(i) = -\infty$ if $i > \ell$ we have

$$\begin{aligned} T_{\lambda,\rho} &= \sum_{i=1}^{\ell} \rho^{\lambda_{(i)}} \\ &= \rho_1^{\lambda_1-1}\rho_2^{\lambda_2-1}\cdots\rho_\ell^{\lambda_\ell-1}(\rho_2\cdots\rho_\ell + \rho_1\rho_3\cdots\rho_\ell + \cdots + \rho_1\cdots\rho_{\ell-1}) \\ &= \rho^{\lambda^*} s_{\ell-1}(\rho_1,\ldots,\rho_\ell), \end{aligned}$$

where $s_{\ell-1}$ is the $(\ell-1)$th elementary symmetric polynomial in ℓ variables. So from (20) and using the symmetry of $s_{\ell-1}$ we have

$$T_{\lambda,\nu} \geqslant T_{\lambda,\theta}. \tag{21}$$

Next note that

$$\begin{aligned} S_{\lambda,\rho} &= T_{\lambda,\rho} + (\lambda_1-1)\rho^{\lambda_{(1)}} + (\lambda_2-1)\rho^{\lambda_{(2)}} + \cdots + (\lambda_\ell-1)\rho^{\lambda_{(\ell)}} \\ &= T_{\lambda,\rho} + \rho_1\cdots\rho_\ell\left(\lambda_1^*\rho^{\lambda^*_{(1)}} + \cdots + \lambda^*_{m-\ell}\rho^{\lambda^*_{(m-\ell)}}\right) \\ &= T_{\lambda,\rho} + \rho_1\cdots\rho_\ell S_{\lambda^*,\rho}. \end{aligned} \tag{22}$$

We prove the assertion $S_{\lambda,\nu} \geqslant S_{\lambda,\theta}$ by induction on the integer λ_1. If $\lambda_1 = 1$, then

$$S_{\lambda,\nu} = s_{m-1}(\nu_1,\ldots,\nu_m) \geqslant s_{m-1}(\theta_1,\ldots,\theta_m) = S_{\lambda,\theta}.$$

If $\lambda_1 > 1$, use (22) to write

$$S_{\lambda,\nu} = T_{\lambda,\nu} + \nu_1\cdots\nu_\ell S_{\lambda^*,\nu}.$$

Then use (21), the inequalities $\nu \geqslant \theta^{\downarrow}$, and the induction hypothesis to conclude that $S_{\lambda,\nu} \geqslant S_{\lambda,\theta}$. □

Combining Propositions 5 and 6 we have:

Proposition 7. *Let ν, θ be m-tuples of nonnegative real numbers such that ν is decreasing and $\theta^{\downarrow} \leqslant \nu$. Let λ, μ be partitions of m such that $\mu \prec \lambda$. Then*

$$S_{\mu,\theta} \leqslant S_{\lambda,\nu}.$$

Proof of Theorem 1. By Proposition 4, the matrix of $D_\lambda(P,I)$ is diagonal in the basis $\{e^*_\alpha : \alpha \in \widehat{\Delta}\}$, and the diagonal elements are given by $S_{\mu^{(\alpha)},\nu_\alpha}$.

Let ω_λ be the element of $\Gamma_{m,n}$ associated with λ by (17). Then $\omega_\lambda \in \overline{\Delta} \subseteq \widehat{\Delta}$. By the proof of Proposition 4, we also have $D_\lambda(P,I)_{\omega_\lambda,\omega_\lambda} = S_{\lambda,\nu}$ (see the relations (18)). So $\|D_\lambda(P,I)\| \geqslant S_{\lambda,\nu}$.

By Theorem 2, $\mu^{(\alpha)} \prec \lambda$. It is obvious that $\nu_\alpha^{\downarrow} \leqslant \nu$. Hence $S_{\mu^{(\alpha)},\nu_\alpha} \leqslant S_{\lambda,\nu}$ by Proposition 7.

Since $S_{\mu^{(\alpha)},\nu_\alpha}$, $\alpha \in \widehat{\Delta}$, is an enumeration of all the eigenvalues of the positive operator $D_\lambda(P,I)$, this implies $\|D_\lambda(P,I)\| \leqslant S_{\lambda,\nu}$. Thus $\|D_\lambda(P,I)\| = S_{\lambda,\nu}$. Use (15) to complete the proof. □

4. Remarks

1. Using standard results of Calculus [1, Chapter X] we can obtain from Theorem 1 perturbation bounds for K_λ. Thus we have for B close to A the first-order perturbation bound

$$\|K_\lambda(A) - K_\lambda(B)\| \leqslant S_{\lambda,\nu}\|A - B\| + \mathrm{O}(\|A - B\|^2). \tag{23}$$

2. Given an irreducible character λ of S_m, let

$$d_\lambda(A) = \frac{1}{\lambda(\mathrm{id})} \sum_{\sigma \in S_m} \lambda(\sigma) \prod_{j=1}^{m} a_{j\sigma(j)}.$$

This is called an *immanant* of A. These functions are important in representation theory and combinatorics.
Let $m = n$. When $\lambda(\sigma) = \varepsilon(\sigma)$ the function d_λ is the determinant, and when $\lambda(\sigma) \equiv 1$, it is the permanent. It is well known that we can choose an orthonormal basis for $V_\lambda(S_n)$ such that $d_\lambda(A)$ is one of the diagonal entries of $K_\lambda(A)$ in this basis. So from (23) we obtain

$$|d_\lambda(A) - d_\lambda(B)| \leqslant S_{\lambda,\nu}\|A - B\| + \mathrm{O}(\|A - B\|^2). \tag{24}$$

3. For simplicity we have restricted our discussion to symmetry classes associated with the full symmetric group. Similar results can be obtained for general symmetry classes. Let G be a subgroup of S_m and let λ be a complex irreducible character of G. Denote by π_λ the multilinearity partition of λ [8]. Using arguments similar to those that have been used to prove Theorem 1 and the results in [9], we can see that

$$\|DK_\lambda(A)\| \leqslant S_{\pi_\lambda,\nu}.$$

Furthermore if the inner product $(\pi_\lambda, \lambda)_G$ is different from zero, then it can be proved that

$$\|DK_\lambda(A)\| = S_{\pi_\lambda,\nu}.$$

Acknowledgement

This work was done while the second author had a visiting position at the Indian Statistical Institute, New Delhi. He thanks the Institute for its wonderful hospitality.

References

[1] R. Bhatia, Matrix Analysis, Springer, New York, 1997.
[2] R. Bhatia, S. Friedland, Variation of Grassmann powers and spectra, Linear Algebra Appl. 40 (1981) 11–18.

[3] R. Bhatia, Variation of symmetric tensor powers and permanents, Linear Algebra Appl. 62 (1984) 269–276.

[4] R. Bhatia, Matrix factorisations and their perturbations, Linear Algebra Appl. 197/198 (1994) 245–276.

[5] R. Bhatia, K.B. Sinha, Variation of real powers of positive operators, Indiana Univ. Math. J. 43 (1994) 913–925.

[6] R. Bhatia, K.B. Sinha, Derivations, derivatives and chain rules, Linear Algebra Appl. 302/303 (1999) 231–234.

[7] R. Bhatia, J.A. Holbrook, Fréchet derivatives of the power function, Indiana Univ. Math. J. 49 (2000) 1155–1173.

[8] J.A. Dias da Silva, A. Fonseca, On the multilinearity partition of an irreducible character, Linear and Multilinear Algebra 20 (1987) 203–218.

[9] J.A. Dias da Silva, A. Fonseca, Nonzero star products, Linear and Multilinear Algebra 27 (1990) 49–55.

[10] G. James, A. Kerber, The Representation Theory of the Symmetric Group, Addison-Wesley, New York, 1981.

[11] R.V. Kadison, J.R. Ringrose, Fundamentals of the Theory of Operator Algebras, vol. 4, Birkhäuser, Boston, MA, 1992.

[12] M. Marcus, Finite Dimensional Multilinear Algebra—Part I, Marcel Dekker, New York, 1973.

[13] R. Merris, Multilinear Algebra, Gordon and Breach, Amsterdam, 1997.

[14] R. Merris, Nonzero decomposable symmetrized tensors, Linear Algebra and Appl. 17 (1977) 287–292.

[15] G.K. Pedersen, Operator differentiable functions, Publ. Res. Inst. Math. Sci. Kyoto 36 (2000) 139–157.

[16] V.S. Sunder, A noncommutative analogue of $D|X|^k = k|X|^{k-1}$, Linear Algebra Appl. 44 (1982) 87–95.

[17] M. Suzuki, Quantum analysis—non-commutative differential calculi, Commun. Math. Phys. 183 (1997) 339–363.