

THE NORM OF THE K -TH DERIVATIVE OF THE χ -SYMMETRIC POWER OF AN OPERATOR*

SÓNIA CARVALHO[†] AND PEDRO J. FREITAS[‡]

Abstract. In this paper, the exact value for the norm of directional derivatives, of all orders, for symmetric tensor powers of operators on finite dimensional vector spaces is presented. Using this result, an upper bound for the norm of all directional derivatives of immanants is obtained.

This work is inspired in results by R. Bhatia, J. Dias da Silva, P. Grover, and T. Jain.

Key words. Norms, Derivatives, Compound matrix, Induced power of a matrix, Immanant.

AMS subject classifications. 15A69, 15A60, 15A15.

1. Introduction. Let V and U be finite-dimensional complex vector spaces, and denote by $\mathcal{L}(V)$ the vector space of linear operators from V to itself. The k -th derivative of a map $f : V \rightarrow U$ is a multilinear map $D^k f(T)$ from $(\mathcal{L}(V))^k$ to $\mathcal{L}(U)$ defined as

$$D^k f(T)(X^1, \dots, X^k) = \frac{\partial^k}{\partial t_1 \dots \partial t_k} \Big|_{t_1 = \dots = t_k = 0} f(T + t_1 X^1 + \dots + t_k X^k).$$

The norm of a multilinear operator $\Phi : (\mathcal{L}(V))^k \rightarrow \mathcal{L}(U)$ is given by

$$\|\Phi\| = \sup_{\|X^1\| = \dots = \|X^k\| = 1} \|\Phi(X^1, \dots, X^k)\|.$$

Let $T \in \mathcal{L}(V)$, with $n = \dim V$. We will obtain exact values for the norm of k -th derivative of the operator $f(T) = K_\chi(T)$, where $K_\chi(T)$ represents the χ -symmetric tensor power of the operator T , that is, the restriction of the operator $\otimes^m T$ to the subspace of χ -symmetric tensors, which we will denote by V_χ .

Let $\nu_1 \geq \dots \geq \nu_n$ be the singular values of the operator T and denote by p_t the symmetric polynomial of degree t (the number of variables to be specified). In the

*Received by the editors on April 8, 2013. Accepted for publication on October 6, 2013. Handling Editor: João Filipe Queirós.

[†]Centro de Estruturas Lineares e Combinatória da Universidade de Lisboa, Av Prof Gama Pinto 2, P-1649-003 Lisboa, and Departamento de Matemática do ISEL, Rua Conselheiro Emídio Navarro 1, 1959-007 Lisbon, Portugal (scarvalho@adm.isel.pt).

[‡]Centro de Estruturas Lineares e Combinatória, Av Prof Gama Pinto 2, P-1649-003 Lisboa, and Departamento de Matemática da Faculdade de Ciências, Campo Grande, Edifício C6, piso 2, P-1749-016 Lisboa, Universidade de Lisboa (pjfreitas@fc.ul.pt).

papers [3], [6] and [7], the following values were obtained:

$$\|D^k \otimes^m T\| = \|D^k \vee^m T\| = \frac{m!}{(m-k)!} \|T\|^{m-k} = \frac{m!}{(m-k)!} \nu_1^{m-k},$$

$$\|D^k \wedge^m T\| = k! p_{m-k}(\nu_1, \dots, \nu_m).$$

On the other hand, R. Bhatia and J. Dias da Silva established in [1] a formula for the norm of the first derivative of the χ -symmetric tensor power of an operator. It is again $k!$ times of the value of the elementary symmetric polynomial of degree $m-1$ taken on m singular values of the operator T , where k is the order of the derivative. In all cases, we note that the norm is the value of the elementary symmetric polynomial of degree $m-k$ applied to a certain family of m singular values of T (eventually with repetitions), multiplied by $k!$.

In this paper, we will present a result that generalizes all these cases. We use techniques developed in [1] and [3], and a result from [5].

2. Results on χ -symmetric tensor powers. We now present some classic facts and notation about χ -symmetric powers that can be found in [8, Chapter 6]. Let S_m be the symmetric group of degree m , χ an irreducible character of S_m and define $K_\chi \in \mathcal{L}(\otimes^m V)$ as

$$K_\chi = \frac{\chi(\text{id})}{m!} \sum_{\sigma \in S_m} \chi(\sigma) P(\sigma),$$

where id stands for the identity element of S_m and $P(\sigma)(v_1 \otimes \dots \otimes v_m) = v_{\sigma^{-1}(1)} \otimes \dots \otimes v_{\sigma^{-1}(m)}$. The range of K_χ is called the *symmetry class of tensors* associated with the irreducible character χ and it is represented by $V_\chi = K_\chi(\otimes^m V)$. We denote

$$v_1 * v_2 * \dots * v_m = K_\chi(v_1 \otimes v_2 \otimes \dots \otimes v_m).$$

These vectors belong to V_χ and are called *decomposable symmetrised tensors*.

Given $T \in \mathcal{L}(V)$, it is known that V_χ is an invariant subspace for $\otimes^m T$. We define the χ -symmetric tensor power of T as the restriction of $\otimes^m T$ to V_χ , and denote it by $K_\chi(T)$.

Let $\Gamma_{m,n}$ be the set of all maps from the set $\{1, \dots, m\}$ into the set $\{1, \dots, n\}$. This set can also be identified with the collection of multi-indices $\{(i_1, \dots, i_m) : i_j \leq n\}$. If $\alpha \in \Gamma_{m,n}$, this correspondence associates to α , the m -tuple $(\alpha(1), \dots, \alpha(m))$. We will consider the lexicographic order in the set $\Gamma_{m,n}$. The group S_m acts on $\Gamma_{m,n}$ by the action $(\sigma, \alpha) \longrightarrow \alpha\sigma^{-1}$ where $\sigma \in S_m$ and $\alpha \in \Gamma_{m,n}$. The set

$$\{\alpha\sigma : \sigma \in S_m\} \subseteq \Gamma_{m,n}$$

is then the orbit of α and the stabilizer of α is the following subgroup of S_m :

$$G_\alpha = \{\sigma \in S_m : \alpha\sigma = \alpha\}.$$

Now suppose that V is a Hilbert space and let $\{e_1, \dots, e_n\}$ be an orthonormal basis of V . Then

$$\{e_\alpha^\otimes = e_{\alpha(1)} \otimes e_{\alpha(2)} \otimes \dots \otimes e_{\alpha(m)} : \alpha \in \Gamma_{m,n}\}$$

is a basis of the m -th tensor power of V . So, by the definition of the space V_χ , the set

$$\{e_\alpha^* = K_\chi(e_\alpha^\otimes) : \alpha \in \Gamma_{m,n}\}$$

spans V_χ . However, this set need not be a basis of V_χ , because its elements might not be linearly independent, some of them may even be zero. Let

$$(2.1) \quad \Omega = \Omega_\chi = \{\alpha \in \Gamma_{m,n} : \sum_{\sigma \in G_\alpha} \chi(\sigma) \neq 0\}.$$

With simple calculations, we can conclude that

$$(2.2) \quad \|e_\alpha^*\|^2 = \frac{\chi(\text{id})}{m!} \sum_{\sigma \in G_\alpha} \chi(\sigma).$$

So the nonzero decomposable symmetrised tensors are $\{e_\alpha^* : \alpha \in \Omega\}$. Now, let Δ be the system of distinct representatives for the quotient set $\Gamma_{m,n}/S_m$, constructed by choosing the first element in each orbit, for the lexicographic order of indices. It is easy to check that $\Delta \subseteq G_{m,n}$, where $G_{m,n}$ is the set of all increasing sequences of $\Gamma_{m,n}$. Let

$$\overline{\Delta} = \Delta \cap \Omega.$$

It can be proved that the set $\{e_\alpha^* : \alpha \in \overline{\Delta}\}$ is linearly independent. We have already seen that the set $\{e_\alpha^* : \alpha \in \Omega\}$, spans V_χ , so there is a set $\widehat{\Delta}$, such that $\overline{\Delta} \subseteq \widehat{\Delta} \subseteq \Omega$ and

$$\{e_\alpha^* : \alpha \in \widehat{\Delta}\}$$

is a basis for V_χ . It is also known that this basis is orthogonal if χ is a linear character, i.e., if $\chi \equiv 1$ or $\chi = \text{sgn}$. If χ is the principal character, $\widehat{\Delta} = \overline{\Delta} = G_{m,n}$ and if χ is the alternating character, $\widehat{\Delta} = \overline{\Delta} = Q_{m,n}$. In both cases the corresponding bases are orthogonal for the induced inner product in $\otimes^m V$.

A partition π of m is an r -tuple of positive integers $\pi = (\pi_1, \dots, \pi_r)$, such that

$$\bullet \quad \pi_1 \geq \dots \geq \pi_r,$$

$$\bullet \pi_1 + \cdots + \pi_r = m.$$

Sometimes it is useful to consider a partition of m with exactly m entries, so we complete the list with zeros. The number of nonzero entries in the partition π is called the *length of π* and is represented by $l(\pi)$.

Given an n -tuple of real numbers $x = (x_1, \dots, x_n)$ and $\alpha \in \Gamma_{m,n}$, we define the m -tuple

$$x_\alpha := (x_{\alpha(1)}, x_{\alpha(2)}, \dots, x_{\alpha(m)}).$$

It is known from representation theory that there is a canonical correspondence between the irreducible characters of S_m and the partitions of m , it is usual to use the same notation to represent both of them. Recall that if $\chi = (1, \dots, 1)$ then $\chi = \text{sgn}$ and $V_\chi = \wedge^m V$ is the Grassmann space, and if $\chi = (m, 0, \dots, 0)$, then $\chi \equiv 1$ and $V_\chi = \vee^m V$.

For every partition $\pi = (\pi_1, \pi_2, \dots, \pi_{l(\pi)})$ of m , we define $\omega(\pi)$ as

$$\omega(\pi) := (\underbrace{1, \dots, 1}_{\pi_1 \text{ times}}, \underbrace{2, \dots, 2}_{\pi_2 \text{ times}}, \dots, \underbrace{l(\pi), \dots, l(\pi)}_{\pi_{l(\pi)} \text{ times}}) \in G_{m,n} \subseteq \Gamma_{m,n}.$$

For each $\alpha \in \Gamma_{m,n}$, let $\text{Im } \alpha = \{i_1, \dots, i_l\}$, suppose that $|\alpha^{-1}(i_1)| \geq \cdots \geq |\alpha^{-1}(i_l)|$. The partition of m

$$(2.3) \quad \mu(\alpha) := (|\alpha^{-1}(i_1)|, \dots, |\alpha^{-1}(i_l)|)$$

is called the *multiplicity partition* of α .

REMARK 1. The multiplicity partition of $\omega(\pi)$ is equal to the partition π :

$$\mu(\omega(\pi)) = \pi.$$

We have that $\text{Im } \omega(\pi) = \{1, 2, \dots, l(\pi)\}$ and that $|\alpha^{-1}(i)| = \pi_i$, for every $i = 1, 2, \dots, l(\pi)$. So

$$\mu(\omega(\pi)) = (|\alpha^{-1}(1)|, |\alpha^{-1}(2)|, \dots, |\alpha^{-1}(l(\pi))|) = (\pi_1, \pi_2, \dots, \pi_{l(\pi)}) = \pi.$$

We recall a well known order defined on the set of partitions of m . A partition μ precedes λ , written $\mu \preceq \lambda$, if for all $1 \leq s \leq m$,

$$\sum_{j=1}^s \mu_j \leq \sum_{j=1}^s \lambda_j.$$

We will also need the following classical result that characterizes the set Ω_χ and is [8, Theorem 6.3.7].

THEOREM 2.1. *Let χ be a partition of m and $\alpha \in \Gamma_{m,n}$. Let Ω_χ and $\mu(\alpha)$ be as defined in (2.1) and (2.3). Then $\alpha \in \Omega_\chi$ if and only if χ majorizes $\mu(\alpha)$.*

3. Formulas for k -th derivatives. We now present a formula for higher order derivatives that generalizes formulas in [2] and [3]. It is known that, given $X^1, \dots, X^m \in \mathcal{L}(V)$, the space V_χ is invariant for the map defined as

$$X^1 \tilde{\otimes} X^2 \tilde{\otimes} \dots \tilde{\otimes} X^m := \frac{1}{m!} \sum_{\sigma \in S_m} X^{\sigma(1)} \otimes \dots \otimes X^{\sigma(m)}.$$

See for instance [8, p. 184]. We will denote the restriction of this map to V_χ by $X^1 * \dots * X^m$ and call it the *symmetrized χ -symmetric tensor product* of the operators $X^1, \dots, X^m$. We remark that this notation does not convey the fact that the product depends on the character χ . In [3], the following formula is deduced:

$$(3.1) \quad D^k(\otimes^m T)(X^1, \dots, X^k) = \frac{m!}{(m-k)!} \underbrace{T \tilde{\otimes} \dots \tilde{\otimes} T}_{m-k \text{ copies}} \tilde{\otimes} X^1 \tilde{\otimes} \dots \tilde{\otimes} X^k.$$

If $k > m$ all derivatives are zero. From this we can deduce a formula for the derivative $D^k K_\chi(T)(X^1, \dots, X^k)$, using the same techniques.

If L is a linear map, it is known that the derivative of L is L at each point. Then, applying the chain rule, we have that

$$D(L \circ f)(a)(x) = L \circ D(f(a))(x).$$

If the map f is k times differentiable then

$$(3.2) \quad D^k(L \circ f)(a)(x^1, \dots, x^k) = L \circ D^k f(a)(x^1, \dots, x^k).$$

THEOREM 3.1. *Using the notation we have established, we have*

$$D^k K_\chi(T)(X^1, \dots, X^k) = \frac{m!}{(m-k)!} T * \dots * T * X^1 * \dots * X^k.$$

If $m = k$ this formula does not depend on T , and if $k > m$ all derivatives are zero.

Proof. Let Q be the inclusion map defined as $Q : V_\chi \longrightarrow \otimes^m V$, so its adjoint operator Q^* is the projection of $\otimes^m V$ onto V_χ . We have

$$T_1 * \dots * T_m = Q^*(T_1 \tilde{\otimes} \dots \tilde{\otimes} T_m)Q.$$

Both maps $L \mapsto Q^*L$ and $L \mapsto LQ$ are linear, so we can apply formulas (3.1) and (3.2) and get

$$\begin{aligned} D^k K_\chi(T)(X^1, \dots, X^n) &= D^k(Q^*(\otimes^m T)Q)(X^1, \dots, X^k) \\ &= Q^*D^k(\otimes^m T)(X^1, \dots, X^k)Q \\ &= \frac{m!}{(m-k)!} Q^* \underbrace{(T \tilde{\otimes} \dots \tilde{\otimes} T)_{m-k \text{ times}}}_{m-k \text{ times}} X^1 \tilde{\otimes} \dots \tilde{\otimes} X^k Q \\ &= \frac{m!}{(m-k)!} T * \dots * T * X^1 * \dots * X^k. \end{aligned}$$

This concludes the proof. $\square$

4. Norm of the k -th derivative of $K_\chi(T)$. We recall that the norm of a multilinear operator $\Phi : (\mathcal{L}(V))^k \rightarrow \mathcal{L}(U)$ is given by

$$\|\Phi\| = \sup_{\|X^1\|=\dots=\|X^k\|=1} \|\Phi(X^1, \dots, X^k)\|.$$

The main result of this section is the following theorem.

THEOREM 4.1. *Let V be an n -dimensional Hilbert space. Let m and k be positive integers such that $1 \leq k \leq m \leq n$, and let χ be a partition of m . Let $T \rightarrow K_\chi(T)$ be the map that associates to each element of $\mathcal{L}(V)$ the induced operator $K_\chi(T)$ on the symmetry class V_χ . Then the norm of the derivative of order k of this map is given by the formula*

$$\|D^k K_\chi(T)\| = k! p_{m-k}(\nu_{\omega(\chi)}),$$

where p_{m-k} is the symmetric polynomial of degree $m-k$ in m variables and $\nu_1 \geq \dots \geq \nu_n$ are the singular values of T .

The proof of our main result is inspired in the techniques used in [1]. We will now highlight the most important features of the proof of our main theorem.

First we will use the polar decomposition of the operator T , in the following form: $P = TW$, with P positive semidefinite and W unitary. We will see that $\|D^k K_\chi(T)\| = \|D^k K_\chi(P)\|$. This allows us to replace T by P . After that we observe that the multilinear map $D^k K_\chi(P)$ is positive between the two algebras in question, so it is possible to use a multilinear version of the famous Russo-Dye theorem that states that the norm for a positive multilinear map is attained in $(I, I, \dots, I)$, where I is the identity operator. This result considerably simplifies the calculations needed to obtain the expression stated in our theorem.

The second part of our proof consists in finding the largest singular value of $D^k K_\chi(P)(I, I, \dots, I)$, which coincides with the value of the norm of $D^k K_\chi(T)$.

First we will need some properties of the operator K_χ . They all follow from the definitions.

PROPOSITION 4.2. *Let χ be an irreducible character of S_m and suppose that S and T are in $\mathcal{L}(V)$ and $v_1, \dots, v_m \in V$. Then*

1. $K_\chi(ST) = K_\chi(S)K_\chi(T)$,
2. $K_\chi(T)(v_1 * \dots * v_m) = T(v_1) * \dots * T(v_m)$,
3. $K_\chi(T)^* = K_\chi(T^*)$, where T^* is the adjoint operator of T ,
4. $K_\chi(T)$ is invertible for all invertible T and in this case $K_\chi(T)^{-1} = K_\chi(T^{-1})$.

By the polar decomposition, we know that for every $T \in \mathcal{L}(V)$ there are a positive semidefinite operator P and a unitary operator W such that $T = PW$. Moreover, the eigenvalues of P are the singular values of T .

PROPOSITION 4.3. *With the above notation, we have*

$$\|D^k K_\chi(T)\| = \|D^k K_\chi(P)\|.$$

Proof. Let $T = PW$, with W unitary. Then $K_\chi(W)$ is also unitary, because

$$[K_\chi(W)]^{-1} = K_\chi(W^{-1}) = K_\chi(W^*) = [K_\chi(W)]^*.$$

So, we have

$$\begin{aligned} \|D^k K_\chi(T)(X^1, \dots, X^k)\| &= \\ &= \|D^k K_\chi(T)(X^1, \dots, X^k)K_\chi(W)\| \\ &= \left\| \left(\frac{\partial^k}{\partial t_1 \dots \partial t_k} \Big|_{t_1=\dots=t_k=0} K_\chi(T + t_1 X^1 + \dots + t_k X^k) \right) K_\chi(W) \right\| \\ &= \left\| \frac{\partial^k}{\partial t_1 \dots \partial t_k} \Big|_{t_1=\dots=t_k=0} K_\chi(T + t_1 X^1 + \dots + t_k X^k) K_\chi(W) \right\| \\ &= \left\| \frac{\partial^k}{\partial t_1 \dots \partial t_k} \Big|_{t_1=\dots=t_k=0} K_\chi(P + t_1 X^1 W + \dots + t_k X^k W) \right\|. \\ &= \|D^k K_\chi(P)(X^1 W, \dots, X^k W)\| \end{aligned}$$

We have $\|X^i W\| = \|X^i\|$, and moreover, $\{XW : \|X\| = 1\}$ is the set of all operators of norm 1, so

$$\begin{aligned} \|D^k K_\chi(T)\| &= \sup_{\|X^1\|=\dots=\|X^k\|=1} \|D^k K_\chi(T)(X^1, \dots, X^k)\| \\ &= \sup_{\|X^1\|=\dots=\|X^k\|=1} \|D^k K_\chi(P)(X^1 W, \dots, X^k W)\| \\ &= \|D^k K_\chi(P)\|. \end{aligned}$$

This concludes the proof. $\square$

Now we need to estimate the norm of the operator $D^k K_\chi(P)$. For this, we use a result from [3], a multilinear version of the Russo-Dye theorem, which we quote here. A multilinear operator Φ is said to be *positive* if $\Phi(X^1, \dots, X^k)$ is a positive semidefinite operator whenever $X^1, \dots, X^k$ are so.

THEOREM 4.4 (Russo-Dye multilinear version). *Let $\Phi : \mathcal{L}(V)^k \rightarrow \mathcal{L}(U)$ be a positive multilinear operator. Then*

$$\|\Phi\| = \|\Phi(I, I, \dots, I)\|.$$

We have that $D^k K_\chi(P)$ is a positive multilinear operator, since if $X^1, \dots, X^k$ are positive semidefinite, then by the formula in Theorem 3.1, $D^k(P)(X^1, \dots, X^k)$ is the restriction of a positive semidefinite operator to an invariant subspace, and thus is positive semidefinite.

Therefore,

$$\|D^k K_\chi(T)\| = \|D^k K_\chi(P)\| = \|D^k K_\chi(P)(I, I, \dots, I)\|.$$

Now we have to find the maximum eigenvalue of $D^k K_\chi(P)(I, I, \dots, I)$. This will be done by finding a basis of V_χ formed by eigenvectors for $D^k K_\chi(P)(I, I, \dots, I)$. If $E = \{e_1, \dots, e_n\}$ is an orthonormal basis of eigenvectors for P , then $\{e_\alpha^* : \alpha \in \widehat{\Delta}\}$ will be a basis of eigenvectors for $D^k K_\chi(P)(I, I, \dots, I)$ (in general, it will not be orthonormal).

For $\beta \in Q_{m-k,m}$, define $\otimes_\beta^m P$ as the tensor $X^1 \otimes \dots \otimes X^m$, in which $X^i = P$ if $i \in \text{Im } \beta$ and $X^i = I$ otherwise.

LEMMA 4.5. *We refer to the notation we have established so far.*

1. *We have*

$$\underbrace{P \tilde{\otimes} \dots \tilde{\otimes} P}_{m-k \text{ times}} \tilde{\otimes} I \tilde{\otimes} \dots \tilde{\otimes} I = \frac{k!(m-k)!}{m!} \sum_{\beta \in Q_{m-k,m}} \otimes_\beta^m P.$$

2. *Let $v_1, \dots, v_m$ be eigenvectors for P with eigenvalues $\lambda_1, \dots, \lambda_m$. Then*

$$\sum_{\beta \in Q_{m-k,m}} \otimes_\beta^m P(v_1 \otimes \dots \otimes v_m) = p_{m-k}(\lambda_1, \dots, \lambda_m) v_1 \otimes \dots \otimes v_m.$$

Proof.

1. It is a matter of carrying out the computations. The factors $k!$ and $(m-k)!$ account for the permutations of I and P , respectively, in each summand. We note

that the final factor is exactly the inverse of the number of summands, as in the definition of $P \tilde{\otimes} \cdots \tilde{\otimes} P \tilde{\otimes} I \tilde{\otimes} \cdots \tilde{\otimes} I$.

2. For each $\beta \in Q_{m-k,m}$ we have that

$$\otimes_{\beta}^m P(v_1 \otimes \cdots \otimes v_m) = \prod_{i=1}^{m-k} \lambda_{\beta(i)}(v_1 \otimes \cdots \otimes v_m).$$

So,

$$\begin{aligned} \sum_{\beta \in Q_{m-k,m}} \otimes_{\beta}^m P(v_1 \otimes \cdots \otimes v_m) &= \sum_{\beta \in Q_{m-k,m}} \prod_{i=1}^{m-k} \lambda_{\beta(i)}(v_1 \otimes \cdots \otimes v_m) \\ &= p_{m-k}(\lambda_1, \dots, \lambda_m) v_1 \otimes \cdots \otimes v_m. \end{aligned}$$

This concludes the proof. $\square$

The following proposition gives the expression for the eigenvalues of the derivative $D^k K_{\chi}(P)(I, I, \dots, I)$.

PROPOSITION 4.6. *Let $\alpha \in \hat{\Delta}$ and define*

$$\lambda(\alpha) := k! p_{m-k}(\nu_{\alpha}).$$

Then $\lambda(\alpha)$ is the eigenvalue of $D^k K_{\chi}(P)(I, I, \dots, I)$ associated with the eigenvector e_{α}^ .*

Proof. Recall that $E = \{e_1, \dots, e_n\}$ is an orthonormal basis of eigenvectors for P , with eigenvalues $\nu_1, \dots, \nu_n$. For every $\alpha \in \hat{\Delta}$ we have

$$e_{\alpha}^* = \frac{\chi(\text{id})}{m!} \sum_{\sigma \in S_m} \chi(\sigma) e_{\alpha\sigma}^{\otimes}.$$

Then

$$\begin{aligned} D^k K_{\chi}(P)(I, \dots, I)(e_{\alpha}^*) &= \frac{m!}{(m-k)!} \underbrace{(P \tilde{\otimes} \cdots \tilde{\otimes} P \tilde{\otimes} I \tilde{\otimes} \cdots \tilde{\otimes} I)}_{m-k \text{ times}}(e_{\alpha}^*) \\ &= \frac{m!}{(m-k)!} \frac{k!(m-k)!}{m!} \sum_{\beta \in Q_{m-k,m}} \otimes_{\beta}^m P(e_{\alpha}^*) \\ &= k! \sum_{\sigma \in S_m} \chi(\sigma) \sum_{\beta \in Q_{m-k,m}} \otimes_{\beta}^m P(e_{\alpha\sigma}^{\otimes}) \\ &= k! \sum_{\sigma \in S_m} \chi(\sigma) p_{m-k}(\nu_{\alpha\sigma}) e_{\alpha\sigma}^{\otimes} \\ &= k! \sum_{\sigma \in S_m} \chi(\sigma) p_{m-k}(\nu_{\alpha}) e_{\alpha\sigma}^{\otimes} \\ &= k! p_{m-k}(\nu_{\alpha}) e_{\alpha}^*. \end{aligned}$$

In the last equations, we used the previous lemma and the symmetry of the polynomial p_{m-k} . So, the eigenvalue associated with e_α^* is $\lambda(\alpha)$. $\square$

We have obtained the expression for all the eigenvalues of $D^k K_\chi(P)(I, \dots, I)$, now we have to find the largest one.

LEMMA 4.7. *If $\alpha, \beta \in \widehat{\Delta}$ are in the same orbit, then $\lambda(\alpha) = \lambda(\beta)$.*

Proof. If α and β are in the same orbit, then there is $\sigma \in S_m$ such that $\alpha\sigma = \beta$. So by the definition of the symmetric elementary polynomials, we have

$$p_{m-k}(\nu_\beta) = p_{m-k}(\nu_{\alpha\sigma}) = p_{m-k}(\nu_\alpha).$$

This concludes the proof. $\square$

According to the results in section 2, every orbit has a representative in $G_{m,n}$, and this is the first element in each orbit (for the lexicographic order). Therefore, the norm of the k -th derivative of $K_\chi(T)$ is attained at some $\lambda(\alpha)$ with $\alpha \in \overline{\Delta} \subseteq G_{m,n}$. We now compare eigenvalues coming from different elements of $\overline{\Delta}$.

LEMMA 4.8. *Let α, β be elements of $\overline{\Delta} \subseteq G_{m,n}$. Then $\lambda(\alpha) \geq \lambda(\beta)$ if and only if α precedes β in the lexicographic order.*

Proof. The result follows directly from the expression of the eigenvalues of $D^k K_\chi(P)(I, \dots, I)$ given in Proposition 4.6. $\square$

We are now ready to complete the proof of the main theorem. From now on we will also write χ to represent also the partition of m associated with the irreducible character χ .

Proof of Theorem 4.1. We have that $\omega(\chi) \in \overline{\Delta}$, so we must have

$$\|D^k K_\chi(P)(I, \dots, I)\| \geq \lambda(\omega(\chi)).$$

Now let $\alpha \in \overline{\Delta}$. Using the results from Theorem 2.1 and Remark 1, we have that $\chi = \mu(\omega(\chi))$ majorizes $\mu(\alpha)$. By the definition of multiplicity partition, we have that $\omega(\chi)$ precedes α in the lexicographic order. By Lemma 4.8, we then have $\lambda(\omega(\chi)) \geq \lambda(\alpha)$ and

$$\|D^k K_\chi(T)\| = \|D^k K_\chi(P)(I, \dots, I)\| = \lambda(\omega(\chi)) = k! p_{m-k}(\nu_{\omega(\chi)}).$$

This concludes the proof of the theorem. $\square$

Now we will see that the formulas obtained by Jain [7] and Grover [6] are particular cases of this last formula.

If $\chi = (m, 0, \dots, 0)$ then $K_\chi(T) = \vee^m T$. In this case $\nu_{\omega(\chi)} = (\nu_1, \dots, \nu_1)$. So, we

have

$$\begin{aligned}\|D^k \vee^m T\| &= k! p_{m-k}(\nu_{\omega(\chi)}) \\ &= k! p_{m-k}(\nu_1, \nu_1, \dots, \nu_1) \\ &= k! \binom{m}{k} \nu_1^{m-k} = \frac{m!}{(m-k)!} \nu_1^{m-k} \\ &= \frac{m!}{(m-k)!} \|T\|^{m-k}.\end{aligned}$$

Also, if $\chi = (1, 1, \dots, 1)$, then $K_\chi(T) = \wedge^m T$ and $\nu_{\omega(\chi)} = (\nu_1, \nu_2, \dots, \nu_m)$. In this case we have that

$$\|D^k \wedge^m T\| = k! p_{m-k}(\nu_1, \nu_2, \dots, \nu_m),$$

where $p_{m-k}(\nu_1, \nu_2, \dots, \nu_m)$ is the symmetric elementary polynomial of degree $m-k$ calculated on the top m singular values of T .

Our main formula also generalizes the result for the norm of the first derivative of $K_\chi(T)$ obtained by R. Bhatia and J. Dias da Silva in [1]. Just notice that if $k=1$, we have that $Q_{1,m} = \{1, 2, \dots, m\}$, so

$$\begin{aligned}\|DK_\chi(T)\| &= p_{m-1}(\nu_{\omega(\chi)}) \\ &= \nu_{\omega(\chi)(2)} \nu_{\omega(\chi)(3)} \cdots \nu_{\omega(\chi)(m)} + \nu_{\omega(\chi)(1)} \nu_{\omega(\chi)(3)} \cdots \nu_{\omega(\chi)(m)} + \cdots \\ &\quad \cdots + \nu_{\omega(\chi)(1)} \nu_{\omega(\chi)(2)} \cdots \nu_{\omega(\chi)(m-1)} \\ &= \sum_{j=1}^m \prod_{\substack{i=1 \\ i \neq j}}^m \nu_{\omega(\chi)(i)}.\end{aligned}$$

5. Norm of the k -th derivative of the immanant. We now wish to establish an upper bound for the k -th derivative of the immanant, which is defined as

$$d_\chi(A) = \sum_{\sigma \in S_n} \chi(\sigma) \prod_{i=1}^n a_{i\sigma(i)},$$

where A is a complex $n \times n$ matrix and χ is an irreducible character of S_n .

The *mixed immanant* of $X^1, \dots, X^n \in M_n(\mathbb{C})$ is defined as

$$\Delta_\chi(X^1, \dots, X^n) := \frac{1}{n!} \sum_{\sigma \in S_n} d_\chi(X_{[1]}^{\sigma(1)}, \dots, X_{[n]}^{\sigma(n)}),$$

where $X_{[j]}^i$ denotes the j -th column of the matrix X^i . If $X^1 = \dots = X^t = A$, for some $t \leq n$ and $A \in M_n(\mathbb{C})$, we denote the mixed immanant by $\Delta_\chi(A; X^{t+1}, \dots, X^n)$.

Now, we recall the definition of $K_\chi(A)$, the m -th χ -symmetric tensor power of the matrix A — from now on, we take χ to be a character of S_m . We find this definition in [5], along with other results we will now use.

A natural way to define this matrix is to fix an orthonormal basis E in V , and consider the linear endomorphism T such that $A = M(T, E)$. Define $\mathcal{E}' = (e_\alpha^* : \alpha \in \widehat{\Delta})$. Let $\mathcal{E} = (v_\alpha : \alpha \in \widehat{\Delta})$ be the orthonormal basis of the m -th χ -symmetric tensor power of the vector space V obtained by applying the Gram-Schmidt orthonormalization procedure to $\mathcal{E}'$. We define

$$K_\chi(A) := M(K_\chi(T), \mathcal{E}).$$

The matrix $K_\chi(A)$ has rows and columns indexed in $\widehat{\Delta}$, with $Q_{m,n} \subseteq \widehat{\Delta}$. This definition admits, as special cases, the m -th compound and the m -th induced power of a matrix, as defined in [8, p. 236].

Since the basis chosen in V_χ is orthonormal, the result for the norm of the operator applies to this matrix:

$$\|K_\chi(A)\| \leq k! p_{m-k}(\nu_{\omega(\chi)}),$$

where $\nu_1 \geq \dots \geq \nu_n$ are the singular values of A . This upper bound is what we will need for the main result in this section.

Before that, we present an explicit formula for $K_\chi(A)$, from [5], and show a similarity between the formulas for $D^k K_\chi(T)$, obtained in Theorem 3.1, and for $D^k K_\chi(A)$, obtained in [5].

Denote by $\text{imm}_\chi(A)$ the square matrix with rows and columns indexed by $\widehat{\Delta}$, whose (γ, δ) entry is $d_\chi(A[\gamma|\delta])$, where $A[\gamma|\delta]$ is the matrix obtained from A , picking rows $\gamma(1), \dots, \gamma(m)$ and columns $\delta(1), \dots, \delta(m)$. Let $B = (b_{\alpha\beta})$ be the change of basis matrix from $\mathcal{E}$ to $\mathcal{E}'$. This means that for each $\alpha \in \widehat{\Delta}$,

$$v_\alpha = \sum_{\gamma \in \widehat{\Delta}} b_{\gamma\alpha} e_\gamma^*.$$

This matrix B does not depend on the choice of the basis E as long as it is orthonormal (it encodes the Gram-Schmidt procedure applied to $\mathcal{E}'$).

With these matrices, we can write

$$(5.1) \quad K_\chi(A) = \frac{\chi(\text{id})}{m!} B^* \text{imm}_\chi(A) B.$$

For $X^1, \dots, X^m \in M_n(\mathbb{C})$, denote by $\text{miximm}_\chi(X^1, \dots, X^m)$ the square matrix with rows and columns indexed by $\widehat{\Delta}$, whose (γ, δ) entry is $\Delta_\chi(X^1[\gamma|\delta], \dots, X^m[\gamma|\delta])$,

so that

$$\text{miximm}_\chi(A, \dots, A) = \text{imm}_\chi(A).$$

We use the same shorthand as with the mixed immanant: For $k \leq m$,

$$\text{miximm}_\chi(A; X^1, \dots, X^k) := \text{miximm}_\chi(A, \dots, A, X^1, \dots, X^k).$$

We have

$$D^k K_\chi(A)(X^1, \dots, X^k) = \frac{\chi(\text{id})}{(m-k)!} B^* \text{miximm}_\chi(A; X^1, \dots, X^k) B.$$

Notice the similarity with the formula in Theorem 3.1. In fact, one can check that if X^i is the matrix of the operator $S^i \in \mathcal{L}(V)$ with respect to the orthonormal basis $\{e_1, \dots, e_n\}$, then the previous matrix is the matrix of $D^k K_\chi(T)(S^1, \dots, S^k)$ with respect to the basis $\mathcal{E}$.

We now use the results on the norm in order to get an upper bound for the norm of the k -th derivative of the immanant.

THEOREM 5.1. *Keeping with the notation established, we have that, for $k \leq n$,*

$$\|D^k d_\chi(A)\| \leq k! p_{n-k}(\nu_{\omega(\chi)}).$$

Proof. We always have $Q_{m,n} \subseteq \widehat{\Delta}$. We now take $m = n$ and denote $\gamma := (1, 2, \dots, n) \in Q_{n,n} \subseteq \widehat{\Delta}$ (this is the only element in $Q_{n,n}$). By definition, $d_\chi(A)$ is the (γ, γ) entry of $\text{imm}_\chi(A)$, and, according to formula (5.1), we have

$$\text{imm}_\chi(A) = \frac{n!}{\chi(\text{id})} (B^*)^{-1} K_\chi(A) B^{-1}.$$

Since multiplication by a constant matrix is a linear map, we have

$$D^k ((B^*)^{-1} K_\chi(A) B^{-1})(X^1, \dots, X^k) = (B^*)^{-1} D^k K_\chi(A)(X^1, \dots, X^k) B^{-1}.$$

We denote by C the column γ of the matrix B^{-1} :

$$C = (B^{-1})_{[\gamma]} = (b'_{\alpha\gamma}), \quad \alpha \in \widehat{\Delta}.$$

Then

$$D^k d_\chi(A)(X^1, \dots, X^k) = \frac{n!}{\chi(\text{id})} C^* D^k K_\chi(A)(X^1, \dots, X^k) C.$$

By formula (2.2), we have that

$$\|e_\gamma^*\|^2 = \frac{\chi(\text{id})}{n!}.$$

By definition of the matrix B , we have

$$e_\gamma^* = \sum_{\beta \in \widehat{\Delta}} b'_{\beta\gamma} v_\beta$$

with $C = [b'_{\beta\gamma} : \beta \in \widehat{\Delta}]$. Since the basis $(v_\alpha : \alpha \in \widehat{\Delta})$ is orthonormal, we have

$$\|C\|^2 = \|C\|_2^2 = \|e_\gamma^*\|^2 = \frac{\chi(\text{id})}{n!},$$

where $\|C\|_2$ is the Euclidean norm of C . Therefore,

$$\begin{aligned} \|D^k d_\chi(A)\| &= \frac{n!}{\chi(\text{id})} \|CD^k K_\chi(A)C^*\| \\ &\leq \frac{n!}{\chi(\text{id})} \|C\|^2 \|D^k K_\chi(A)\| \\ &= k! p_{n-k}(\nu_{\omega(\chi)}). \end{aligned}$$

This concludes the proof. $\square$

In [4], it is proved that this upper bound coincides with the norm of the derivative of the determinant. In [3], when $d_\chi = \text{per}$, the upper bound presented in formula (52) is, using our notation, $(n!/(n-k)!)\|A\|^{n-k}$. Using our formula, we get the same value: For $\omega(\chi) = (1, 1, \dots, 1)$,

$$k! p_{n-k}(\nu_{\omega(\chi)}) = k! \binom{n}{n-k} \nu_1^{n-k} = \frac{n!}{(n-k)!} \|A\|^{n-k}.$$

It is also shown that for

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

we have strict inequality, for $d_\chi = \text{per}$.

One of the purposes of having upper bounds for norms is the possibility of estimating the magnitude of perturbations. Taylor's formula states that if f is a p times differentiable function between two normed spaces, then

$$f(a+x) - f(a) = \sum_{k=1}^p \frac{1}{k!} D^k f(a)(x, \dots, x) + O(\|x\|^{p+1}).$$

Therefore,

$$\|f(a+x) - f(a)\| \leq \sum_{k=1}^p \frac{1}{k!} \|D^k f(a)\| \|x\|^k.$$

Using our formulas, we get the following result.

COROLLARY 5.2. *According to our notation, we have, for $T, X \in \mathcal{L}(V)$ and $A, Y \in M_n(\mathbb{C})$:*

$$\|K_\chi(T) - K_\chi(T+X)\| \leq \sum_{k=1}^m p_{m-k}(\nu_{\omega(\chi)}) \|X\|^k,$$

$$|d_\chi(A) - d_\chi(A+Y)| \leq \sum_{k=1}^n p_{n-k}(\nu_{\omega(\chi)}) \|Y\|^k.$$

Acknowledgements. We would like to thank Professors Rajendra Bhatia and José Dias da Silva for all the discussions about this subject, both about classical theory and the topics relating to this particular problem.

REFERENCES

- [1] R. Bhatia and J.A. Dias da Silva. Variation of induced linear operators. *Linear Algebra and its Applications*, 341:391–402, 2002.
- [2] R. Bhatia and S. Friedland. Variation of Grassman powers and spectra. *Linear Algebra and its Applications*, 40:1–18, 1981.
- [3] R. Bhatia, P. Grover, and T. Jain. Derivatives of tensor powers and their norms. *Electronic Journal of Linear Algebra*, 26:604–619, 2013.
- [4] R. Bhatia and T. Jain. Higher order derivatives and perturbation bounds for determinants. *Linear Algebra and its Applications*, 431:2102–2108, 2009.
- [5] S. Carvalho and P.J. Freitas. The k -th derivatives of the immanant and the χ -symmetric power of an operator. *Electronic Journal of Linear Algebra*, 27:284–301, 2014.
- [6] P. Grover. Derivatives of multilinear functions of matrices. *Matrix Information Geometry*, Springer, Heidelberg, 93–109, 2013.
- [7] T. Jain. Derivatives for antisymmetric tensor powers and perturbation bounds, *Linear Algebra and its Applications*, 435:1111–1121, 2011.
- [8] R. Merris. *Multilinear Algebra*. Gordon and Breach Science Publishers, Amsterdam, 1997.