UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE MATEMÁTICA

# Higher Order Derivatives and Norms of Certain Matrix Functions

**Sónia Raquel Ferreira Carvalho**

Doutoramento em Matemática
Especialidade de Álgebra, Lógica e Fundamentos

2014

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE MATEMÁTICA

# Higher Order Derivatives and Norms of Certain Matrix Functions

**Sónia Raquel Ferreira Carvalho**

Tese orientada pelo Professor Pedro J. Freitas e pelo Professor Rajendra Bhatia especialmente elaborada para a obtenção do grau de doutor em Matemática especialidade de Álgebra, Lógica e Fundamentos

2014

*To my parents, Gorete and António*

# Contents

# Acknowledgements

Behind the work of a person there are the advice, suggestions, contributions, corrections and support of so many others. I could not be happier to acknowledge all the persons that have made part of this work in one way or another.

I would like to express my deep gratitude to my research supervisors, Professor Pedro Freitas and Professor Rajendra Bhatia. I would like to thank Professor Bhatia for presenting me the idea for this research work and for his advice and precious suggestions throughout the whole process, especially the comments and remarks made at the last versions of the thesis. They have greatly contributed to improve the final result. Having the opportunity to work with someone with such a broad knowledge has been a rewarding experience.

I would like to thank Professor Freitas for his infinite patience, guidance, enthusiastic encouragement, useful critiques of this research work and also for his friendship. It has been quite a long journey. Since its beginning I have always had the support and dedication of Professor Freitas, starting with the PhD programme, qualification exams, PhD project, seminars until the writing of this thesis.

I would also like to thank Professor J. A. Dias da Silva, for his advice and assistance and for his contribution in finding and calculating some heavy technical results that I would not achieve without his help. I also would like to thank him for being a reference as a teacher and researcher, throughout all my academic career.

I would also like to extend my thanks to my colleagues and my friends at

Instituto Superior de Engenharia de Lisboa that have given me the possibility to work on my PhD research in full time by giving me leave of absence during four semesters. In particular, I have to thank Isabel Melo for taking care of all the issues related with tuition with professionalism and dedication.

Last, but not least, my lovely family. I wish to thank my parents Gorete and António and my sister Cláudia for their unconditional support and full encouragement throughout my study. They have always been there. I would like to dedicate this thesis to my mother that has been my strength, my role model, the reason why I have never given up.

I would also like to thank my sweet daughters Madalena and Margarida for understanding my, not so good, mood and for the peace and quiet that they, occasionally, gave me when I was working at home. Finally, I extend my gratitude to Carlos, for his shoulder in my weakest moments and for always believing.

Thank you all for being a part of this long, hard but rewarding journey.

> *It is good to have an end to journey toward; but it is the journey that matters, in the end.*
>
> Ernest Hemingway.

# Resumo

A ideia de obter fórmulas para as derivadas e para as respectivas normas de certas c funções de matrizes foi sugerida pelo Professor Rajendra Bhatia, no sentido de generalizar resultados obtidos anteriormente.

O pioneiro neste tipo de problemas foi Carl Gustav Jacob Jacobi (século XIX) que calculou a primeira derivada da função determinante. Seja det : $M_n(\mathbb{C}) \longrightarrow \mathbb{C}$ a aplicação que a cada matriz quadrada de ordem $n$ faz corresponder o seu determinante, onde $M_n(\mathbb{C})$ representa o espaço vectorial das matrizes quadadradas de ordem $n$ sobre o corpo dos números complexos. Pela definição de derivada direccional no ponto $A \in M_n(\mathbb{C})$, para cada $X \in M_n(\mathbb{C})$,

$$D \det(A)(X) = \frac{d}{dt}\Big|_{t=0} \det(A + tX).$$

A famosa *Fórmula de Jacobi* é

$$D \det(A)(X) = \operatorname{tr}(\operatorname{adj}(A)X),$$

onde $\operatorname{adj}(A)$ representa a matriz adjunta de $A$ e tr o traço da matriz.

A questão que se coloca é como pode este resultado ser generalizado. Por um lado, podemos tentar generalizá-lo para derivadas de ordem superior, por outro podemos considerar a primeira derivada de outras funções matriciais, das quais o determinante é um caso particular. O determinante pode ser encarado como um caso particular de uma função generalizada de matrizes, no entanto uma vez que $\det(A) = \wedge^n A$, este também pode ser visto como uma das potências de Grassmman de $A$, i.e. $\wedge^m A$, no caso em que $m = n$.

A noção de $k$-ésima derivada de uma função $\phi$ é dada do seguinte modo: considerando $A, X^1, \ldots X^k \in M_n(\mathbb{C})$,a $k$-ésima derivada no ponto $A$ e nas direcções de $(X^1, \ldots, X^k)$ é uma função multilinear definida por

$$D^k\phi(A)(X^1, \ldots, X^k) := \frac{\partial^k}{\partial t_1 \ldots \partial t_k}\Big|_{t_1=\ldots=t_k=0} \phi(A + t_1X^1 + \ldots + t_kX^k).$$

Dado $\chi$ um carácter irredutível do grupo de permutações de ordem $n$, $S_n$ sabe-se que o imanente de $A$, associado a $\chi$ é dado pela expressão:

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

e tem como casos particulares o determinante, que corresponde ao carácter alternante, e o permanente que corresponde ao carácter principal. Considerando $\chi$ um carácter irredutível de $S_m$, $\wedge^m A$ e $\vee^m A$ são também casos particulares da $m$-ésima potência $\chi$-simétrica de $A$, $K_\chi(A)$.

Em [9], R. Bhatia e T. Jain calcularam várias expressões para as derivadas de ordem superior do determinante. Em seguida T. Jain [17] obteve fórmulas para as derivadas de ordem superior das restantes potências de Grassmman de $A$.

Recentemente, R. Bhatia, T. Jain e P. Grover [9], [8] deram um passo noutro sentido da generalização da fórmula de Jacobi e calcularam diversas expressões para a $k$-ésima derivada de um outro imanente, o permanente. Também obtiveram fórmulas para as $k$-ésimas derivadas de outra potência $\chi$-simétrica de $A$, a $m$-ésima potência induzida da matriz $A$, representada por $\vee^m A$, onde $1 \leq k \leq m \leq n$.

A outra questão que estudámos nesta dissertação prende-se com as normas das derivadas de ordem superior do imanente e da potência $\chi$- simétrica, em todos os problemas a norma considerada é a norma espectral.

Em 1981 R. Bhatia e S. Friedland, em [7] deram o primeiro passo neste tipo de problemas quando demonstraram que

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

onde $\nu_1 \geq \nu_2 \geq \ldots \geq \nu_n$ são os valores singulares da matriz $A$ e $p_{m-1}$ representa o polinómo simétrico elementar de grau $m-1$, neste caso em $m$ variáveis.

Pouco tempo depois, R. Bhatia, em [5], obteve uma expressão para a norma da derivada da potência simétrica,

$$\|D \vee^m A\| = m\|A\|^{m-1} = m\nu_1^{m-1}. \tag{2}$$

Em 2002 R. Bhatia and J. Dias da Silva demonstraram em [6], um resultado para a norma da primeira derivada de qualquer classe simétrica, generalizando deste modo as expressões (1) e (2).

Recentemente, R. Bhatia, P. Grover e T. Jain estenderam as mesmas expressões (1) e (2) noutra direcção na medida em que estabeleceram fórmulas para as normas das derivadas de ordem superior das potências de Grassmman e simétricas, respectivamente.

Em [17] e [8], os autores provaram que

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

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

Em todos os casos estudados, notamos que a norma é dada por

$$k!\, p_{m-k}(\nu_{i_1}, \ldots, \nu_{i_m}),$$

em que $p_{m-k}$ é um polinómio simétrico elementar em $m$ variáveis e $\nu_{i_1}, \ldots, \nu_{i_m}$ um conjunto de $m$ valores singulares do operador $T$ (eventualmente com repetições).

Apesar de se tratarem de resultados que abordam assuntos similares, convem notar que as técnicas e áreas envolvidas nos processos de demonstração de cada um deles são bastante diferentes.
Nesta dissertação de doutoramento generalizamos todos os resultados anteriores, ou seja, são calculadas fórmulas para as $k$-ésimas derivadas do imanente e da $m$-ésima potência $\chi$-simétrica de $A$. Também são analisadas as normas de algumas destas derivadas.

A ideia que levou ao estudo destes problemas pode ser inserida na área da análise matricial, uma vez que queremos calcular expressões para derivadas de funções de matrizes. No entanto, ao observarmos as funções estudadas, este problema insere-se claramente no contexto da álgebra multilinear, que por sua vez, está fortemente ligada à combinatória e à teoria da representação. Portanto, como acontece na grande maioria das vezes, estamos perante um problema multidisciplinar. Os fundamentos teóricos que foram necessários à compreensão e resolução destes problemas encontram-se nos livros [3], [12] e [25], que são verdadeiras referências para qualquer tipo de trabalho que se insira nestas áreas do saber.
É também muito pertinente salientar a importância que a notação tem neste tipo de áreas da matemática. Por um lado utilizamos notação clássica da álgebra multilinear e combinatória, por outro introduzimos muita notação

nova, com o objectivo de simplificar conceitos bastante profundos e de alguma forma manter-nos compatíveis com a notação existente.

O primeiro problema surgiu com a função imanente. As únicas propriedades dos imanentes que são herdadas do determinante são aquelas que decorrem do facto de qualquer imanente ser também uma função multilinear nas colunas (ou linhas) da matriz A. Assim, foi possível generalizar a expressão da primeira derivada do imanente e, recorrendo ao argumento de multilinearidade e alguma notação bastante complexa, generalizar a primeira expressão para a sua $k$-ésima derivada. Para obter a segunda expressão para a $k$-ésima derivada foi necessário provar a expansão de Laplace para imanentes. Para isso tivemos que contornar o facto de o imanente de uma matriz de ordem $n$ não se calcular através de imanentes de submatrizes. O que fizemos para ultrapassar esta questão foi através da generalização da soma directa usual de matrizes.

Quando passamos ao cálculo da $k$-ésima derivada da $m$-ésima potência $\chi$-simétrica de $A$, apresentam-se algumas questões. A primeira é precisamente construção da matriz $K_\chi(A)$ que generaliza $\vee^m A$ e $\wedge^m A$. Outra questão prende-se com o facto de que os únicos caracteres lineares de $S_m$ serem o carácter principal e o carácter alternante. Entre outras coisas isto significa que apenas o espaço de Grassmann e o espaço dos tensores completamente simétricos têm bases ortogonais conhecidas, para o produto interno induzido, formadas por tensores decomponíveis.
No caso geral $V_\chi$ para $\chi$ não linear isto não acontece. Como consequência é necessário recorrer ao processo de ortogonalização de Gram-Schmidt na base induzida. Desta forma os cálculos são significativamente mais complicados e as expressões obtidas para o caso geral são muito mais complexas. Para além disso, tal como acontece com o imanente, não é possível efectuar cálculos com submatrizes e temos que, novamente, utilizar uma generalização de soma directa de matrizes

O estudo das normas das derivadas de ordem superior da potência $\chi$-simétrica é feito para o caso da norma espectral. Primeiramente observamos que, dado um operador $T$ e considerando a decomposição polar $T = PW$, que as normas de $D^k K_\chi(T)$ e $D^k K_\chi(P)$ coincidem, uma vez que $K_\chi(W)$ é unitário e a norma considerada é unitariamente invariante. Ainda assim se quiséssemos explicitar a norma de $D^k K_\chi(P)$ os cálculos seriam complicadíssimos.

Porém, T. Jain [8] conseguiu generalizar o corolário do famoso teorema de Russo-Dye para aplicações multilineares, o que nos permitiu concluir que a norma da $k$-ésima derivada de $K_\chi(P)$ é atingida se todas as direções forem iguais à matriz identidade, o que simplifica consideravelmente os cálculos a efectuar. Por fim, pela definição da norma espectral temos que determinar o maior valor próprio de $D^k K_\chi(P)(I, I, \ldots, I)$. Tal foi possível através de um resultado clássico da álgebra multilinear. Usando a expressão obtida para a norma de $D^k K_\chi(T)$ conseguimos obter também uma expressão para as normas das derivadas de ordem superior de $K_\chi(A)$, onde $A$ é uma matriz quadrada de ordem $n$ que representa $T$ em relação a uma base ortonormada. A partir desta expressão obtemos também majorantes para a norma das derivadas de ordem superior qualquer imanente.

Finalmente, as ideias para trabalho futuro nesta área passam por várias questões distintas. A primeira questão que surge é se os resultados continuam válidos se considerarmos que $V$ é um espaço vectorial com dimensão infinita. No caso das desigualdades obtidas para as normas, uma das questões que se pode colocar é para que matrizes é que a igualdade é válida.
No que diz respeito às expressões das derivadas de ordem superior demonstradas, que melhoramentos podemos obter se considerarmos casos particulares de caracteres irredutíveis de $S_m$ (ou partições de $m$), nomeadamente, se considerarmos a família de partições com propriedades já conhecidas e estudadas, como é o caso das *hook partitions*. Numa perspectiva mais generalista podemos pensar se é possível provar para qualquer imanente ou para qualquer potência $\chi$-simétrica outro tipo de resultados conhecidos para o determinante ou para o permanente e para as potências simétrica e anti-simétrica.

**Palavras Chave:** Carácter, imanente, derivada direccional, norma matricial, composta de uma matriz, potência induzida de uma matriz.

# Abstract

In this thesis we obtain formulas for higher order directional derivatives for the immanant map, which is a generalization of the determinant and the permanent maps. We also obtain formulas for the $k$-th derivative of the $m$-th $\chi$-symmetric tensor power of an operator or a matrix, where $\chi$ is an irreducible character of the permutation group $S_m$. Moreover, we calculate the operator norm of these derivatives.

We start by presenting some general concepts of multilinear algebra, representation theory and matrix analysis, in particular some results about characters of $S_m$ and differential calculus applied to matrix functions, which will be useful throughout this work. We also present some well known results about the immanant map, as well as other results such as the generalized Laplace expansion for immanants.

The starting point of this kind of problems is the famous Jacobi formula obtained in the 19th century by Carl Jacobi. This formula gives us the first order derivative of the determinant function. In recent work, R. Bhatia, T. Jain and P. Grover [9], [8] presented us formulas for higher order derivatives of the determinant and the permanent maps and also the expressions for the derivatives of the symmetric and antisymmetric tensor powers. These maps are all particular cases of the immanant and the $\chi$-symmetric tensor power, when $\chi$ is a linear character of $S_m$, namely the principal and alternating characters. The general case has much more complicated features and needs some new concepts and notations, which play a very important role throughout our work.

We also study the norm of these higher order derivatives. This problem was first addressed by R. Bhatia and S. Friedland in [7] where they proved

a formula for $D \wedge^m (A)$. This result has been extended in two different directions. In [9], R. Bhatia and T. Jain study the case for higher order derivatives and they obtain a formula for the norm of $D^k \wedge^m (A)$, whereas in [6], R. Bhatia and J. A. Dias da Silva demonstrate a formula for the norm of the first derivative for all symmetry classes. In our work we obtain a result that subsumes all the previous expressions.

**Keywords:** Immanant, Fréchet derivative, matrix norm, compound matrix, induced power of a matrix.

# List of Symbols

| Symbol | Description | Page |
|---|---|---|
| $\langle , \rangle$ | inner product | 6 |
| $\equiv$ | equivalence relation | 7 |
| $\mathrm{adj}(A)$ | adjugate of $A$ | 48 |
| $\mathrm{adj}_\chi(A)$ | immanantal adjoint of $A$ | 27 |
| $A \bigoplus B$ | direct sum of matrices | 16 |
| $A \bigoplus\limits_{\alpha\|\beta} B$ | $\alpha$, $\beta$ direct sum of matrices | 34 |
| $A(i\|j)$ | | 26 |
| $A(j;X)$ | replace column $j$ by the $j$-th column of $X$ | 48 |
| $A\{\alpha,\beta\}$ | | 32 |
| $A(\alpha;X^1,\ldots,X^k)$ | | 56 |
| $A_{[i]}$ | $i$-th column of the matrix $A$ | 57 |
| $\otimes^m A$ | $m$-th tensor power of $A$ | 9 |
| $A_1 \otimes \ldots \otimes A_m$ | Kronecker product | 9 |
| $\vee^m A$ | permanental coumpound of $A$ | 59 |
| $\wedge^m A$ | determinantal compound of $A$ | 59 |
| $\|\alpha\|$ | $\alpha(1)+\ldots+\alpha(k)$ | 29 |
| $\|\alpha^{-1}(i)\|$ | | 51 |
| $C(\sigma)$ | conjugate class of $\sigma$ | 12 |
| $\chi$ | irreducible character | 11 |
| $\overline{\chi}$ | conjugate character | 12 |
| $d_\chi^G$ | generalized matrix function | 24 |
| $d_\chi(A)$ | immanant of $A$ | 24 |
| $\det(A)$ | determinant of $A$ | 24 |
| $D\phi(A)$ | Fréchet derivative | 46 |
| $D\phi(A)(X)$ | first derivative of $\phi(A)$ in direction $X$ | 46 |
| $D^k\phi(A)(X^1,\ldots,X^k)$ | $k$-th derivative of $\phi(A)$ in directions $(X^1,\ldots,X^k)$ | 47 |

| | | |
|---|---|---|
| $\delta_{\alpha\beta}$ | Kronecker delta | 7 |
| $\Delta$ | | 18 |
| $\overline{\Delta}$ | | 19 |
| $\widehat{\Delta}$ | | 19 |
| $\Delta(B_1, \ldots, B_n)$ | mixed discriminant | 67 |
| $\Delta_\chi(X^1, \ldots, X^n)$ | mixed immanant | 58 |
| $e_\alpha^\otimes$ | decomposable tensor associated to $\alpha$ | 6 |
| $e_\alpha^*$ | decomposable symmetrized tensor associated to $\alpha$ | 16 |
| $\mathcal{E}'$ | induced basis of $V_\chi$ | 19 |
| $\mathcal{E}$ | orthonormal basis of $V_\chi$ | 23 |
| $\\|f\\|$ | norm of a linear operator | 73 |
| $\\|\phi\\|$ | norm of a multilinear operator | 74 |
| $GL(n, \mathbb{C})$ | $n \times n$ invertible complex matrices | 10 |
| $G_\alpha$ | stabilizer subgroup | 17 |
| $G_{m,n}$ | increasing functions in $\Gamma_{m,n}$ | 18 |
| $\Gamma_{m,n}$ | functions $\alpha : \{1, \ldots, m\} \longmapsto \{1, \ldots, n\}$ | 6 |
| $\Gamma(n_1, \ldots, n_m)$ | | 50 |
| $\Gamma_{n,k}^0$ | | 51 |
| $\overline{\Gamma}$ | | 56 |
| id | identity of $S_m$ | 12 |
| $\operatorname{Im} \alpha$ | image of $\alpha$ | 31 |
| $\operatorname{imm}_\chi(A)$ | matrix with $(\alpha, \beta)$-entry $d_\chi(A[\alpha\|\beta])$ | 63 |
| $I$ | identity operator | 8 |
| $I(S_m)$ | set of irreducible characters of $S_m$ | 11 |
| $K_\chi$ | symmetrizer map | 13 |
| $K_\chi(A)$ | $m$-th $\chi$-symmetric tensor power of $A$ | 62 |
| $K_\chi(T)$ | induced transformation | 22 |
| $l(\pi)$ | length of the partition $\pi$ | 69 |
| $\mathcal{L}(V)$ | linear transformations from $V$ to $V$ | 7 |
| $\mathcal{L}(X, Y)$ | bounded linear operators from $X$ to $Y$ | 46 |
| $\lambda \preceq \mu$ | $\lambda$ precedes $\mu$ | 77 |
| $\lambda_\alpha$ | | 53 |
| $\lambda(\alpha)$ | | 83 |
| $\Lambda$ | | 51 |
| $\Lambda_\beta$ | equivalence class of $\beta$ | 52 |
| $\operatorname{miximm}_\chi(A)$ | | 63 |
| $M_n(\mathbb{C})$ | $n \times n$ complex matrices | 10 |

# Introduction

> *It is not uncommon to find a special richness and vitality at the boundary between mathematical disciplines. With roots in linear algebra, group representation theory, and combinatorics, multilinear algebra is an important example. Serious expeditions into any of these fertile areas require substantial preparation, and multilinear algebra is no exception.*
>
> Russell Merris, in Multilinear Algebra.

When we start to read this thesis, the first thing we come across is its title, which is clearly related to the field of matrix analysis. Although, after we look at the contents, we see the strong presence of multilinear algebra and combinatorics.

The idea, the motivation, the main goals and results of this work are, in fact, in the area of matrix analysis, but the tools, the techniques, the lemmas are certainly in the areas of multilinear algebra and combinatorics. That is why we began by quoting the first paragraph of one of the most used references in this area.

This dissertation is divided into three parts. Chapter two is devoted to general concepts of three areas of mathematics, multilinear algebra, combinatorics and representation theory, that we will use. In the other two chapters, we state and prove some new results on higher order derivatives of special matrix functions and we also derive expressions and upper bounds for the operator norm of these derivatives.

*Symmetry, as wide or narrow as you may define its meaning, is one idea by which man through the ages has tried to comprehend and create order, beauty, and perfection.*

Hermann Weyl

In chapter one we give an overview of tensor spaces and symmetry classes of tensors, with a subsection devoted to characters of the permutation group. In the last part of this chapter, we have gathered most of the results that are known about the immmanant of a square matrix, but we also prove some new results that extend the formulas which are already known for determinants and for permanents, such as the *Laplace Expansion for Immanants* and the *Generalized Laplace Expansion for Immanants.* The last formula evaluates the immanant of a square matrix of order $n$ using immanants of special type of a direct sum of submatrices of order $k$ and order $n-k$. In order to do this we introduce some new notation that will be used in the following chapters.

In the next chapter we obtain formulas for the higher order derivatives of the immanant and the $m$-th $\chi$-symmetric tensor power of an operator, where $\chi$ is an irreducible character of the permutation group $S_m$.

It is known that the determinant map and the permanent map are special cases of a more generalized map, which is the immanant, and the compound and the induced power of a matrix are also generalized by other symmetric powers, related to symmetry classes of tensors. These will be our objects of study. We present various expressions for the $k$-th derivatives that extend the formulas previously established for the two special cases.

In the last chapter we obtain exact values for the norm of the $k$-th derivative of the operator $f(T) = K_\chi(T)$, where $K_\chi(T)$ represents the $\chi$-symmetric tensor power of the operator $T$, that is, the restriction of the operator $\otimes^m T$ to the subspace of $\chi$-symmetric tensors, which we will denote by $V_\chi$. This kind of problem was first addressed in [7] by R. Bhatia and S. Friedland where they found the norm of the first derivative of the Grassmann power of a matrix, which led to a striking formula:

$$\|D \wedge^m (A)\| = p_{m-1}(\nu_1, \ldots, \nu_m) \tag{3}$$

where $p_{m-1}$ is the symmetric polynomial of degree $m-1$ in $m$ variables and $\nu_1 \geq \nu_2 \geq \ldots \geq \nu_n$ represent the singular values of the matrix $A$. Later R.

Bhatia and J. A. Dias da Silva have proved a formula that gives an explicit expression for the norm of the first derivative for other symmetry classes. Again, this expression is given by the symmetric polynomial of degree $m-1$ in $m$ variables calculated on the "top $m$" singular values of $A$, these "top $m$" singular values are chosen according to each symmetry class.

Recently, T. Jain have generalized formula (3) in another way. In [17] it is proved that

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

Similar formulas have been obtained by R. Bhatia, P. Grover and T. Jain [8] for the norm of the higher order derivatives of the permanent and for $\vee^m T$. In this last chapter we demonstrate a formula that subsumes all the previous cases, which was published in [11]. Our proof is inspired in the techniques used by R. Bhatia and J. A. Dias da Silva in [6].

Finally, we present a result for the norm of the $k$-th derivative of $K_\chi(A)$, the $m$-th $\chi$-symmetric tensor power of the matrix $A$, a matrix we have defined in the previous chapter. Using this result we are able to obtain an upper bound for the $k$-th derivative of the immanant $d_\chi(A)$ and some inequalities which are consequences of Taylor's formula.

There are some questions that are still unanswered, such as: "Do the formulas hold in infinite dimension?", "Are the inequalities sharp?". We end this thesis by summarizing some problems and questions for future work on this subject.

# Chapter 1

# General Concepts

*The truth is rarely pure and never simple.*

Oscar Wilde

We present some general results of multilinear algebra and some classical definitions that will be used throughout the next chapters. General references for this topic are [12] and [25].

## 1.1 Tensor Power

Let $V$ be a vector space of dimension $n$ over the field of complex numbers $\mathbb{C}$ and let $m$ be an integer such that $1 \leq m \leq n$.

**Definition 1.1.1.** The $m$-th *tensor power* of $V$, denoted by $\otimes^m V$, is the tensor product of $m$ copies of $V$. In particular, $\otimes^1 V = V$ and $\otimes^0 V = \mathbb{C}$.

If $E = \{e_1, e_2, \ldots e_n\}$ is a basis of $V$, then

$$\{e_{i_1} \otimes e_{i_2} \otimes \ldots \otimes e_{i_m} : 1 \leq i_j \leq n, 1 \leq j \leq m\}$$

is a basis of $\otimes^m V$ induced by $E$. In order to simplify this notation we will introduce some concepts.

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 also be identified with the collection of multiindices $\{(i_1,\ldots,i_m) : i_j \leq n, j = 1,\ldots,m\}$. If $\alpha \in \Gamma_{m,n}$, this correspondence associates to $\alpha$ the $m$-tuple

$$\alpha := (\alpha(1),\ldots,\alpha(m)).$$

In the set $\Gamma_{m,n}$ we will consider the lexicographic order.

**Definition 1.1.2.** Let $\alpha$ be an element of $\Gamma_{m,n}$. A *decomposable tensor associated to the element* $\alpha$, $u_\alpha^\otimes$, is an element of $\otimes^m V$ such that

$$u_\alpha^\otimes := u_{\alpha(1)} \otimes u_{\alpha(2)} \otimes \ldots \otimes u_{\alpha(m)}.$$

In particular, if $E = \{e_1, e_2, \ldots e_n\}$ is a basis of $V$ then the set

$$\{e_\alpha^\otimes : \alpha \in \Gamma_{m,n}\}$$

is the basis of $\otimes^m V$, induced by $E$.

We have that $\dim \otimes^m V = n^m$.

Notice that $\otimes^m V \neq \{v_1 \otimes v_2 \otimes \ldots \otimes v_m : v_i \in V, 1 \leq i \leq m\}$, meaning that there are tensors in $\otimes^m V$ that are not decomposable.

**Definition 1.1.3.** A *Hilbert space* $V$ is a vector space $V$ over the field $\mathbb{C}$ together with an inner product, i.e., with a map

$$\langle\, ,\, \rangle : V \times V \longrightarrow \mathbb{C},$$

that satisfies the following three axioms for all vectors $u, v, w \in V$ and all scalars $a, b \in \mathbb{C}$

1. $\langle v, w\rangle = \overline{\langle w, v\rangle}$,
2. $\langle au + bv, w\rangle = a\langle u, w\rangle + b\langle v, w\rangle$,
3. If $u \neq 0$ then $\langle u, u\rangle > 0$.

If $V$ is a Hilbert space, then the $m$-th tensor power of $V$ is also a Hilbert space.

**Proposition 1.1.4.** *Let $V$ be an $n$-dimensional Hilbert space and suppose that $E = \{e_1, e_2, \dots, e_n\}$ is an orthonormal basis of $V$. Let $z, w \in \otimes^m V$ such that*

$$z = \sum_{\alpha \in \Gamma_{m,n}} b_\alpha e_\alpha^\otimes, \qquad w = \sum_{\alpha \in \Gamma_{m,n}} c_\alpha e_\alpha^\otimes.$$

*Then,*

$$\langle z, w \rangle = \sum_{\alpha \in \Gamma_{m,n}} b_\alpha \overline{c_\alpha},$$

*defines an inner product on the space $\otimes^m V$. In particular, if $x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_m \in V$ then*

$$\langle x_1 \otimes x_2 \otimes \dots \otimes x_m, y_1 \otimes y_2 \otimes \dots \otimes y_m \rangle = \prod_{i=1}^{m} \langle x_i, y_i \rangle. \tag{1.1}$$

This is called the *induced inner product* in $\otimes^m V$ and this is the unique inner product that satisfies (1.1). For simplicity we use the same notation to represent the inner product in $V$ and the induced inner product in $\otimes^m V$.
If $E = \{e_1, e_2, \dots, e_n\}$ is an orthonormal basis of $V$, then the induced basis of $\otimes^m V$ is also an orthonormal basis, because for every $\alpha, \beta \in \Gamma_{m,n}$ we have

$$\langle e_\alpha^\otimes, e_\beta^\otimes \rangle = \prod_{i=1}^{m} \langle e_{\alpha(i)}, e_{\beta(i)} \rangle = \prod_{i=1}^{m} \delta_{\alpha(i)\beta(i)} = \delta_{\alpha\beta}.$$

We will denote by $\mathcal{L}(V)$ the vector space consisting of linear operators $T : V \longrightarrow V$. Recall that the adjoint operator of $T$, $T^*$ is the unique operator that satisfies

$$\langle Tu, v \rangle = \langle u, T^* v \rangle,$$

for every $u, v \in V$.

**Definition 1.1.5.** Let $T \in \mathcal{L}(V)$. We say that

i) $T$ is *normal* if $TT^* = T^*T$.

ii) $T$ is *unitary* if $TT^* = T^*T = I$.

iii) $T$ is *hermitian* if $T^* = T$,

where $I$ stands for the identity operator.

Let us recall that every normal operator has an orthonormal basis of eigenvectors, meaning that every matrix representation of $T$ is unitarily similar to a diagonal matrix. Moreover if $T$ is unitary then its eigenvalues are in the unit circle and if $T$ is hermitian then its eigenvalues are all real.

If the operators $T_i \in \mathcal{L}(V)$, $i = 1, \ldots, m$,then there is a unique operator $\mathbf{L} \in \mathcal{L}(\otimes^m V)$ such that

$$\mathbf{L}(v_1 \otimes v_2 \otimes \ldots \otimes v_m) = T_1(v_1) \otimes T_2(v_2) \otimes \ldots \otimes T_m(v_m),$$

for every $v_i \in V$. This operator $\mathbf{L}$ is called the *tensor product of the operators* $T_1, T_2, \ldots T_m$ and it is denoted by

$$\mathbf{L} := T_1 \otimes T_2 \otimes \ldots \otimes T_m.$$

In particular, if $T_1 = T_2 = \ldots = T_m = T$, we will write this as $\otimes^m T$, the $m$-th tensor power of the operator $T$.

Now, we list some properties of the tensor product of linear operators.

**Proposition 1.1.6.** *Let $T_i, S_i \in \mathcal{L}(V)$, $i = 1, \ldots, m$. Then*

1. $(S_1 \otimes S_2 \otimes \ldots \otimes S_m) \circ (T_1 \otimes T_2 \otimes \ldots \otimes T_m) = (S_1 \circ T_1) \otimes (S_2 \circ T_2) \otimes \ldots \otimes (S_m \circ T_m)$.

2. $\otimes^m I_V = I_{\otimes^m V}$.

3. $(T_1 \otimes T_2 \otimes \ldots \otimes T_m)(\otimes^m V) = T_1(V) \otimes T_2(V) \otimes \ldots \otimes T_m(V)$.

4. *If $T_1, T_2, \ldots, T_m$ are injective (respectively: bijective, normal, hermitian or unitary) then $T_1 \otimes T_2 \otimes \ldots \otimes T_m$ is injective (respectively: bijective, normal, hermitian or unitary).*

5. *$T_1 \otimes T_2 \otimes \ldots \otimes T_m$ is invertible if and only if $T_1, T_2, \ldots T_m$ are invertible.*

Since the tensor product of linear operators in $V$ is a linear operator in $\otimes^m V$, it can be represented by a matrix of order $n^m$.

**Definition 1.1.7.** Let $A_p = (a^p_{ij})$ be a $k_p \times n_p$ matrix, $1 \le p \le m$. The *Kronecker product* of the matrices $A_1, A_2, \ldots, A_m$, represented by $A_1 \otimes A_2 \otimes \ldots \otimes A_m$ is a $\prod k_p \times \prod n_p$ matrix whose rows are indexed by the set $\{(i_1, i_2, \ldots, i_m) : 1 \le i_p \le k_p\}$ and whose columns are indexed by the set $\{(j_1, j_2, \ldots, j_m) : 1 \le j_p \le n_p\}$, both ordered lexicographically. The $((i_1, i_2, \ldots, i_m), (j_1, j_2, \ldots, j_m))$- entry of this matrix is

$$\prod_{p=1}^{m} a^p_{i_p j_p}.$$

**Example 1.1.8.** Let $A_1 = (a^1_{ij}), A_2 = (a^2_{kl})$ be $n_1 \times n_1$and $n_2 \times n_2$ matrices. Then $A_1 \otimes A_2$ is a $n_1 n_2 \times n_1 n_2$ matrix, and

$$A_1 \otimes A_2 = \begin{pmatrix} a^1_{11}A_2 & a^1_{12}A_2 & \ldots & a^1_{1n_1}A_2 \\ a^1_{21}A_2 & a^1_{22}A_2 & \ldots & a^1_{2n_1}A_2 \\ \ldots & \ldots & \ldots & \ldots \\ a^1_{n_1 1}A_2 & a^1_{n_1 2}A_2 & \ldots & a^1_{n_1 n_1}A_2 \end{pmatrix}.$$

**Theorem 1.1.9.** *Let $E = \{e_1, e_2, \ldots, e_n\}$ be a basis of the Hilbert space $V$ and $T_i \in \mathcal{L}(V)$, $i = 1, 2, \ldots, m$ such that $A_i$ is the matrix representation of $T_i$ with respect to $E$, i.e. $A_i = M(T_i, E)$.*

*Then the matrix representation of $T_1 \otimes T_2 \otimes \ldots \otimes T_m$ with respect to the basis $\{e^{\otimes}_{\alpha} : \alpha \in \Gamma_{m,n}\}$ is $A_1 \otimes A_2 \otimes \ldots \otimes A_m$.*

In particular, if $A$ is a $n \times n$ complex matrix we write

$$\otimes^m A := A \otimes A \otimes \ldots \otimes A,$$

the Kronecker product of $m$ copies of $A$ which is usually called the $m$-fold tensor power of the matrix $A$. We have that $\otimes^m A$ is an $n^m \times n^m$ matrix.

We list below some properties of these tensor powers. For $A$ and $B$ $n \times n$ complex matrices, we have

1. $(\otimes^m A)(\otimes^m B) = \otimes^m (AB)$.
2. $(\otimes^m A)^{-1} = \otimes^m A^{-1}$ when $A$ is invertible.
3. $(\otimes^m A)^* = \otimes^m A^*$.

4. If $A$ is hermitian, unitary, normal or positive, then so is $\otimes^m A$.

5. If $\lambda_1, \lambda_2, \ldots, \lambda_m$ are $m$ eigenvalues of $A$, not necessarily distinct, with eigenvectors $u_1, u_2, \ldots, u_m$, respectively, then $\lambda_1\lambda_2 \ldots \lambda_m$ is an eigenvalue of the matrix $\otimes^m A$ associated to the eigenvector $u_1 \otimes u_2 \otimes \ldots \otimes u_m$.

All of these properties are true if we replace the matrix $A$ by a linear operator $T$ on the Hilbert space $V$.

## 1.2 Symmetry Classes of Tensors

### Characters of $S_m$

Denote by $M_l(\mathbb{C})$ the set of $l \times l$ complex matrices and by $GL(l, \mathbb{C})$ the multiplicative group of all $l \times l$ invertible matrices.

**Definition 1.2.1.** Let $G$ be a group. A *representation* of the group $G$ of *degree* $l$ is a homomorphism $A : \sigma \longrightarrow GL(l, \mathbb{C})$. If the homomorphism is one-to-one, then the representation is *faithful*.

In this thesis we only study the particular case when $G = S_m$, where $S_m$ denotes the permutation group of order $m$, although in some parts we will also state some interesting properties for a subgroup $G$ of $S_m$.

**Example 1.2.2.** 1. For every $\sigma \in S_m$, define

$$A(\sigma) = (\delta_{i\sigma(j)})_{ij},$$

the $m \times m$ complex matrix whose $(i, j)$-entry is equal to 1 if $\sigma(j) = i$ and 0 otherwise. There are $m!$ different matrices of this kind and they are called *permutation matrices* of order $m$. Notice that if $\sigma_1, \sigma_2 \in S_m$ then:

$$A(\sigma_1\sigma_2) = A(\sigma_1)A(\sigma_2).$$

Then $A$ is representation of $S_m$, of degree $m$. It is easy to see that $A$ is faithful.

2. Using the previous example, for each $\sigma \in S_m$ define

$$B(\sigma) = \det(A(\sigma)).$$

With simple calculations we can see that $B$ is a representation of $S_m$ of degree 1. It is easy to check that $B$ is not faithful.

**Definition 1.2.3.** Let $N$ be a nonempty set and suppose $S = \{A(\nu) : \nu \in N\}$ is a set of $m \times m$ complex matrices indexed by $N$. Then the set $S$ is *reducible* if there is an invertible matrix $Q$ of order $m$ and an integer $p$ with $1 < p < m$, such that for all $\nu \in N$,

$$Q^{-1}A(\nu)Q = \begin{pmatrix} B(\nu) & 0 \\ C(\nu) & D(\nu) \end{pmatrix},$$

where $B(\nu)$ is a $p \times p$ complex matrix. $S$ is *irreducible* if it is not reducible. The representation $A$ of $S_m$ is called reducible or irreducible if the set of matrices $\{A(\sigma) : \sigma \in S_m\}$ has the corresponding property.

**Definition 1.2.4.** Let $A$ be a representation of $S_m$. Let $\chi : S_m \longrightarrow \mathbb{C}$ be defined as

$$\chi(\sigma) = \operatorname{tr} A(\sigma),$$

where $\operatorname{tr} A(\sigma)$ stands for the trace of $A(\sigma)$. The function $\chi$ is called the *character* of the group $S_m$ afforded by the representation $A$.

If $A$ is an irreducible representation, then $\chi$ is an *irreducible character.* We will denote by $I(S_m)$ the set of all irreducible characters of $S_m$.

From now on, $\chi$ will always denote an irreducible character of $S_m$. If $A$ is a representation of degree 1, then $\chi$ is a homomorphism of groups, i.e. for every $\sigma, \tau \in S_m$,

$$\chi(\sigma\tau) = \chi(\sigma)\chi(\tau).$$

In this case $\chi$ is said to be a *linear character.*
The following proposition is well-known.

**Proposition 1.2.5.** *There are only two irreducible linear characters of $S_m$, for every natural number $m$. These linear characters are*

- $\chi \equiv 1$ *the* principal character *of $S_m$.*
- $\chi(\sigma) = \operatorname{sgn}(\sigma)$ *the* alternating character *of $S_m$, where* $\operatorname{sgn}(\sigma)$ *stands for the sign of the permutation $\sigma$.*

Now we list some properties of characters of the permutation group.

**Proposition 1.2.6.** *[25] Let $\chi$ be an irreducible character of the group $S_m$ with degree $l$. Suppose $\sigma, \tau \in S_m$ and* id *is the identity element of $S_m$. Then:*

1. $\chi(\mathrm{id}) = \mathrm{tr}(I) = l$.
2. $|\chi(\sigma)| \leq l$, *for every* $\sigma \in S_m$.
3. $\chi(\sigma\tau) = \chi(\tau\sigma)$.
4. $\chi(\tau^{-1}\sigma\tau) = \chi(\sigma)$, *that is,* $\chi$ *is constant on the conjugacy classes of* $S_m$.
5. $\chi(\sigma)$ *is an integer for every* $\sigma \in S_m$.

**Remark 1.2.7.** Notice that:

1. The elements $\sigma, \tau \in S_m$ are in the same conjugacy class if and only if $\sigma$ and $\tau$ have the same cycle type. We will represent by $C(\sigma)$ the conjugacy class of the element $\sigma$.

2. The last property of the previous proposition is only true if we consider the characters of the whole permutation group. In this case $\sigma$ and $\sigma^{-1}$ are in the same conjugacy class, and as a consequence $\chi$ has to be real. Using some results of Galois theory, it can be proved that $\mathrm{Im}\,\chi$ has to be an integer. However, if we consider a more general case, that is $G$ a subgroup of $S_m$ then the values of $\chi$ need not be real and we only have that $\overline{\chi}$ is an irreducible character and $\overline{\chi}(\sigma) = \overline{\chi(\sigma)} = \chi(\sigma^{-1})$.

Next we state two technical results, which we will use later The first one establishes relations between two irreducible characters of $S_m$. This lemma is an extension of a well known result called the *Orthogonality Relations of the First Kind.*

**Lemma 1.2.8.** *[25] Let $\chi, \xi \in I(S_m)$, then*

$$\sum_{\sigma \in S_m} \chi(\sigma^{-1})\xi(\sigma\tau) = \begin{cases} \dfrac{m!\chi(\tau)}{\chi(\mathrm{id})}, & \textit{if } \chi = \xi; \\ 0, & \textit{otherwise.} \end{cases}$$

In order to state the second lemma we need the following notation. We represent by $o(G)$ the number of elements of the subgroup $G$. This lemma describes the *Orthogonality Relations of the Second Kind.*

**Lemma 1.2.9.** *[25] Let $\sigma$ and $\tau$ be elements of the permutation group $S_m$. Then*

$$\sum_{\sigma \in S_m} \chi(\sigma^{-1})\chi(\tau) = \begin{cases} \dfrac{m!}{o(C(\sigma))}, & \text{if } \tau \in C(\sigma); \\ 0, & \text{otherwise.} \end{cases}$$

We are interested in studying some important subspaces of the space $\otimes^m V$, which are associated to an irreducible character of $S_m$.

**Definition 1.2.10.** For each $\sigma \in S_m$, we define the linear operator $P(\sigma) \in \mathcal{L}(\otimes^m V)$ as

$$P(\sigma)(v_1 \otimes v_2 \otimes \ldots \otimes v_m) = v_{\sigma^{-1}(1)} \otimes v_{\sigma^{-1}(2)} \otimes \ldots \otimes v_{\sigma^{-1}(m)}.$$

We also consider the map $P$ defined as

$$P : S_m \longrightarrow \mathcal{L}(\otimes^m V),$$

$$\sigma \longmapsto P(\sigma).$$

It is easy to check that the map $P$ has the following properties:

- $P(\sigma\tau) = P(\sigma)P(\tau)$, $\sigma, \tau \in S_m$.
- $P(\sigma)$ is invertible and $P(\sigma^{-1}) = P(\sigma)^{-1}$.
- For every $\sigma \in S_m$, $P(\sigma)$ is unitary, i.e. $P(\sigma)^* = P(\sigma)^{-1}$.

In particular, $P$ is said to be a unitary representation of $S_m$, since every matrix $P(\sigma)$ is unitary, for every $\sigma \in S_m$.

**Definition 1.2.11.** Let $\chi$ be an irreducible character of $S_m$ and define the operator

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

where id stands for the identity element of $S_m$. For every $\chi$, $K_\chi$ is a linear operator on $\otimes^m V$. It is called the *symmetrizer map*.

**Theorem 1.2.12.** *Suppose $\chi$ is an irreducible character of $S_m$. Then $K_\chi$ is an orthogonal projection in the space $\otimes^m V$.*

*Proof.* Suppose that $\chi$ is an irreducible character of $S_m$.

1. $K_\chi$ **is an Hermitian operator:**

$$\begin{aligned} K_\chi^* &= \frac{\chi(\mathrm{id})}{m!}\sum_{\sigma\in S_m}\overline{\chi(\sigma)}P(\sigma)^* \\ &= \frac{\chi(\mathrm{id})}{m!}\sum_{\sigma\in S_m}\chi(\sigma^{-1})P(\sigma^{-1}) \\ &= K_\chi. \end{aligned}$$

2. $K_\chi$ **is idempotent.**

   First we notice that $K_\chi$ commutes with $P(\sigma)$ for every $\sigma \in S_m$.

$$\begin{aligned} K_\chi^2 &= \left(\frac{\chi(\mathrm{id})}{m!}\right)^2\left(\sum_{\sigma\in S_m}\chi(\sigma)P(\sigma)\right)\left(\sum_{\tau\in S_m}\chi(\tau)P(\tau)\right) \\ &= \left(\frac{\chi(\mathrm{id})}{m!}\right)^2\sum_{\sigma,\tau\in S_m}\chi(\sigma)\chi(\tau)P(\sigma\tau) \\ &= \frac{\chi(\mathrm{id})}{m!}\sum_{\gamma\in S_m}\frac{\chi(\mathrm{id})}{m!}\left(\sum_{\sigma\in S_m}\chi(\sigma)\chi(\sigma^{-1}\gamma)\right)P(\gamma) \qquad (\gamma=\sigma\tau) \\ &= \frac{\chi(\mathrm{id})}{m!}\sum_{\gamma\in S_m}\chi(\gamma)P(\gamma) \\ &= K_\chi. \end{aligned}$$

   The fourth equality follows from Lemma 1.2.8.

□

**Definition 1.2.13.** Suppose $\chi$ is an irreducible character of $S_m$. The range of $K_\chi$ is called the *symmetry class of tensors* associated with the irreducible character $\chi$ and it is represented by

$$V_\chi = K_\chi(\otimes^m V).$$

There is a relation between the whole tensor product space $\otimes^m V$ and its subspaces $V_\chi$, for an irreducible character $\chi$. This relation is a corollary of the following theorem:

**Theorem 1.2.14.** *Suppose $\chi$ and $\xi$ are irreducible characters of $S_m$. If $\chi \neq \xi$, then $K_\chi K_\xi = 0$. Moreover,*

$$\sum_{\chi \in I(S_m)} K_\chi = I,$$

*where $I$ is the identity operator in $\mathcal{L}(\otimes^m V)$.*

*Proof.* By the definition of the symmetrizer map

$$\begin{aligned}
K_\chi K_\xi &= \frac{\chi(\mathrm{id})\xi(\mathrm{id})}{m!^2}\left(\sum_{\sigma\in S_m}\chi(\sigma)P(\sigma)\right)\left(\sum_{\tau\in S_m}\xi(\tau)P(\tau)\right)\\
&= \frac{\chi(\mathrm{id})\xi(\mathrm{id})}{m!^2}\sum_{\sigma\tau\in S_m}\chi(\sigma)\xi(\tau)P(\sigma\tau)\\
&= \frac{\chi(\mathrm{id})\xi(\mathrm{id})}{m!^2}\sum_{\sigma\in S_m}\left(\sum_{\mu\in S_m}\chi(\sigma)\xi(\sigma^{-1}\mu)\right)P(\mu) \qquad (\mu=\sigma\tau)\\
&= 0,
\end{aligned}$$

by Lemma 1.2.8. On the other hand,

$$\begin{aligned}
\sum_{\chi\in I(S_m)} K_\chi &= \frac{1}{m!}\sum_{\chi\in I(S_m)}\chi(\mathrm{id})\sum_{\sigma\in S_m}\chi(\sigma)P(\sigma)\\
&= \sum_{\sigma\in S_m}\frac{1}{m!}\left(\sum_{\chi\in I(S_m)}\chi(\mathrm{id})\chi(\sigma)\right)P(\sigma)\\
&= P(\mathrm{id})\\
&= I,
\end{aligned}$$

by Lemma 1.2.9. $\square$

**Corollary 1.2.15.** *The space $\otimes^m V$ is the orthogonal direct sum of the symmetry classes $V_\chi$ as $\chi$ ranges over $I(S_m)$. This means, if $I(S_m) = \{\chi_1, \chi_2, \ldots, \chi_l\}$, then*

$$\otimes^m V = V_{\chi_1} \oplus V_{\chi_2} \oplus \ldots \oplus V_{\chi_l}.$$

*Proof.*

$$\begin{aligned}
\otimes^m V &= I(\otimes^m V) \\
&= (K_{\chi_1} + K_{\chi_2} + \ldots + K_{\chi_l})(\otimes^m V) \\
&= K_{\chi_1}(\otimes^m V) \oplus K_{\chi_2}(\otimes^m V) \oplus \ldots \oplus K_{\chi_l}(\otimes^m V) \\
&= V_{\chi_1} \oplus V_{\chi_2} \oplus \ldots \oplus V_{\chi_l}.
\end{aligned}$$

Theorem 2.2.13 assures that the sum is direct. That concludes our proof. □

It is well known that the alternating character $\chi(\sigma) = \mathrm{sgn}(\sigma)$ (sign of the permutation $\sigma$) leads to the symmetry class $\wedge^m V$ which is called the *space of skew-symmetric tensors*, the $m$-th *Grassmann space*, or the *$m$-th exterior power of $V$*. It is also well known that the symmetry class corresponding to the the principal character $\chi(\sigma) \equiv 1$ of the group $S_m$ is represented by $\vee^m V$and is usually called the space of *completely symmetric tensors.*[25]

Given a symmetrizer map $K_\chi$, we denote

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

These vectors belong to $V_\chi$ and are called *decomposable symmetrized tensors.* It is important to observe that the chosen notation does not emphasize the fact that $*$ depends on the irreducible character $\chi$. Again if $\chi(\sigma) = \mathrm{sgn}(\sigma)$ we usually write decomposable symmetrized tensors as

$$v_1 \wedge v_2 \wedge \ldots \wedge v_m$$

and if $\chi$ is the principal character, we write

$$v_1 \vee v_2 \vee \ldots \vee v_m.$$

The space $V_\chi$ contains all of the decomposable symmetrized tensors, but, in general, it is not equal to this set. However, from the previous definitions we can conclude that $V_\chi$ is spanned by the set:

$$\{e^*_\alpha := K_\chi(e^\otimes_\alpha) = K_\chi(e_{\alpha(1)} \otimes e_{\alpha(2)} \otimes \ldots \otimes e_{\alpha(m)}) : \alpha \in \Gamma_{m,n}\}.$$

Now we turn to finding a basis for the space $V_\chi$ . The set $\{e^*_\alpha : \alpha \in \Gamma_{m,n}\}$ is a set of generators. It is easy to check that, in general, this set is not a basis of $V_\chi$ , in fact some of its elements may be equal to zero. In order to find a basis for $V_\chi$ we need some more definitions.

First notice that the group $S_m$ acts on the set $\Gamma_{m,n}$ by the action defined as

$$(\sigma, \alpha) \longrightarrow \alpha\sigma^{-1}$$

where $\sigma \in S_m$ and $\alpha \in \Gamma_{m,n}$.

We recall some standard definitions from group theory.

**Definition 1.2.16.** Let $\alpha \in \Gamma_{m,n}$. The set

$$\{\alpha\sigma : \sigma \in S_m\}$$

is the *orbit* of $\alpha$.

If $\alpha, \beta \in \Gamma_{m,n}$ are in the same orbit then we say that $\alpha$ and $\beta$ are *equivalent* and we write $\alpha \equiv \beta$.

**Example 1.2.17.** Consider $\alpha = (3, 3, 5, 1) \in \Gamma_{4,6}$. Then the orbit of $\alpha$ is:

$$\{(1,3,3,5),(1,3,5,3),(1,5,3,3),(3,1,3,5),(3,1,5,3),(3,3,1,5),$$
$$(3,3,5,1),(3,5,1,3),(3,5,3,1),(5,1,3,3),(5,3,1,3),(5,3,3,1)\}.$$

In particular, $\alpha \equiv (1, 3, 3, 5) \in G_{4,6}$.

**Definition 1.2.18.** Let $\alpha \in \Gamma_{m,n}$. The *stabilizer* of $\alpha$ is the subgroup of $S_m$ defined as

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

**Example 1.2.19.** For $\alpha = (3, 3, 5, 1) \in \Gamma_{4,6}$,

$$G_\alpha = \{\mathrm{id}, (12)\}.$$

Considering the set of generators of $V_\chi$ indexed by $\Gamma_{m,n}$, we want to remove the ones that are equal to zero.

**Lemma 1.2.20.** *Let $\{e_1, \ldots, e_n\}$ be an orthonormal basis of the Hilbert space $V$. For every $\alpha \in \Gamma_{m,n}$ we have*

$$\|e^*_\alpha\|^2 = \frac{\chi(\mathrm{id})}{m!} \sum_{\sigma \in G_\alpha} \chi(\sigma). \tag{1.2}$$

*Proof.* In order to calculate the norm of $e^*_\alpha$ we start by calculating the inner product between two generators of $V_\chi$. Let $\alpha, \beta \in \Gamma_{m,n}$. Since $K_\chi$ is hermitian and idempotent,

$$\begin{aligned} \langle e^*_\alpha, e^*_\beta \rangle &= \langle K_\chi(e^\otimes_\alpha), K_\chi(e^\otimes_\beta) \rangle \\ &= \langle K_\chi(e^\otimes_\alpha), e^\otimes_\beta \rangle. \end{aligned}$$

Using the definition of $K_\chi$,

$$\langle e^*_\alpha, e^*_\beta \rangle = \frac{\chi(\mathrm{id})}{m!} \sum_{\sigma \in S_m} \chi(\sigma) \prod_{t=1}^{m} \langle e_{\alpha(t)}, e_{\beta\sigma(t)} \rangle.$$

Since the basis is orthonormal the only nonzero summands are the ones for which $\alpha$ and $\beta$ are equivalent. So we have that

$$\|e^*_\alpha\|^2 = \frac{\chi(\mathrm{id})}{m!} \sum_{\sigma \in G_\alpha} \chi(\sigma). \tag{1.3}$$

That concludes our proof. $\square$

Now let

$$\Omega = \Omega_\chi = \{\alpha \in \Gamma_{m,n} : \sum_{\sigma \in G_\alpha} \chi(\sigma) \neq 0\}. \tag{1.4}$$

So the nonzero decomposable symmetrized tensors are $\{e^*_\alpha : \alpha \in \Omega\}$, this set, obviously, generates $V_\chi$.

**Definition 1.2.21.** Let $\Delta$ 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.

Recall that $G_{m,n}$ is the set of all increasing sequences of $\Gamma_{m,n}$.

**Lemma 1.2.22.** *With the notation already defined, we have the following*

1. *If $\beta \in \Gamma_{m,n}$, then there are $\alpha \in G_{m,n}$ and $\sigma \in S_m$ such that $\beta = \alpha\sigma$.*

2. *We have $\alpha, \beta \in G_{m,n}$ and $\sigma \in S_m$ with $\beta = \alpha\sigma$ if and only if $\alpha = \beta$.*

Using the previous lemma it is easy to see that $\Delta = G_{m,n}$, that is, in every orbit there is an unique element of $G_{m,n}$ and that element is the smallest in the lexicographic order.
Let

$$\overline{\Delta} = \Delta \cap \Omega \subseteq G_{m,n}.$$

On the one hand it can be proved that the set

$$\{e^*_\alpha : \alpha \in \overline{\Delta}\}$$

is linearly independent. On the other hand, the set $\{e^*_\alpha : \alpha \in \Omega\}$, spans $V_\chi$. So, we can conclude that there is a set $\widehat{\Delta}$, with

$$\overline{\Delta} \subseteq \widehat{\Delta} \subseteq \Omega$$

such that

$$\mathcal{E}' := \{e^*_\alpha : \alpha \in \widehat{\Delta}\}, \tag{1.5}$$

is a basis for $V_\chi$, formed with decomposable symmetrized tensors.
From the classical theory, it is also known that if $\chi$ is one of the two linear characters of $S_m$, then $\overline{\Delta} = \widehat{\Delta}$ and in these two cases, the basis is orthogonal. In particular, if $\chi$ is the alternating character then

$$\overline{\Delta} = Q_{m,n},$$

the subset of $\Gamma_{m,n}$ of all strictly increasing maps, and in this case

$$\dim(\wedge^m V) = \binom{n}{m}.$$

If $\chi$ is the principal character of $S_m$ then

$$\overline{\Delta} = G_{m,n}$$

and

$$\dim(\vee^m V) = \binom{n+m-1}{m}.$$

If $\chi$ is an irreducible character of $S_m$ it is known that $\overline{\chi} = \chi$, because the range of $\chi$ lies in the integers. However if $\chi$ is an irreducible character of a subgroup $G$ of $S_m$ this is not true. In this case we only have that $\overline{\chi}$ is also an irreducible character of $G$. In the next proposition we establish a relation between the symmetry class of tensors associated with $\chi$ and $\overline{\chi}$, which is trivial in the case of characters of $S_m$.

**Proposition 1.2.23.** *Let $G$ be a subgroup of $S_m$. Suppose $\chi$ is an irreducible character of $G$. Then*

$$V_\chi = K_\chi(\otimes^m V) = K_{\overline{\chi}}(\otimes^m V) = V_{\overline{\chi}}.$$

*Proof.* First notice that

$$\overline{\Delta}_\chi = \overline{\Delta}_{\overline{\chi}},$$

it follows from the definition of $\Omega_\chi$ that

$$\sum_{\sigma \in G_\alpha} \chi(\sigma) \neq 0 \,\text{if and only if}\, \sum_{\sigma \in G_\alpha} \overline{\chi}(\sigma) \neq 0$$

with $\alpha \in \Gamma_{m,n}$.

So

$$\widehat{\Delta}_\chi \cap \widehat{\Delta}_{\overline{\chi}} \neq \emptyset$$

and we can choose $\alpha \in \widehat{\Delta}_\chi \cap \widehat{\Delta}_{\overline{\chi}}$.

From 1.2.15 we have that either

$$V_\chi \,\text{and}\, V_{\overline{\chi}} \text{ are ortogonal or } \;V_\chi = V_{\overline{\chi}}.$$

Let $E = \{e_1, e_2, \ldots, e_n\}$ be an orthonormal basis of the Hilbert space $V$. Let $e^*_\alpha \in V_\chi$, $e^{\overline{*}}_\alpha \in V_{\overline{\chi}}$ be elements of the induced bases of the spaces $V_\chi$ and $V_{\overline{\chi}}$, respectively. We compute the inner product of these elements of $\otimes^m V$.

$$\begin{aligned}
\langle e^*_\alpha, e^{\overline{*}}_\alpha \rangle &= \left\langle K_\chi(e^\otimes_\alpha), K_{\overline{\chi}}(e^\otimes_\alpha) \right\rangle \\
&= \left(\frac{\chi(\mathrm{id})}{m!}\right)^2 \left\langle \sum_{\sigma\in G} \chi(\sigma) e^\otimes_{\alpha\sigma}, \sum_{\tau\in S_m} \chi(\tau^{-1}) e^\otimes_{\alpha\tau} \right\rangle \\
&= \left(\frac{\chi(\mathrm{id})}{m!}\right)^2 \sum_{\sigma,\tau\in G_\alpha} \chi(\sigma)\chi(\tau^{-1}) \| e^\otimes_{\alpha\sigma}\|^2 \\
&= \left(\frac{\chi(\mathrm{id})}{m!}\right)^2 \| e^\otimes_{\alpha}\|^2 \sum_{\sigma,\tau\in G_\alpha} \chi(\sigma)\chi(\tau^{-1}) \\
&= \left(\frac{\chi(\mathrm{id})}{m!}\right)^2 \| e^\otimes_{\alpha}\|^2 \sum_{\sigma\in G_\alpha} \chi(\sigma) \sum_{\tau\in G_\alpha} \chi(\tau^{-1})
\end{aligned}$$

Since $\alpha \in \Omega_\chi$, we have that $\sum_{\sigma\in G_\alpha} \chi(\sigma) \neq 0$ and $\sum_{\tau\in G_\alpha} \chi(\tau^{-1}) \neq 0$, by the definition of $\Omega_\chi$ . So

$$\langle e^*_\alpha, e^{\overline{*}}_\alpha \rangle \neq 0.$$

Then we must have

$$V_\chi = V_{\overline{\chi}}.$$

This concludes our proof. □

Given a linear operator $T$ in $\mathcal{L}(V)$ we have previously defined the $m$-th tensor power of $T$ which we have represented by $\otimes^m T$. Now we want to define an induced transformation on the space $V_\chi$. Since $V_\chi \subset \otimes^m V$, in order to define this induced transformation, we consider the restriction of the operator $\otimes^m T$ to $V_\chi$. We will see that the space $V_\chi$ is an invariant subspace for the operator $\otimes^m T$.

**Definition 1.2.24.** Suppose $L \in \mathcal{L}(\otimes^m V)$. Then $L$ is *bisymmetric* if it commutes with $P(\sigma)$, for every $\sigma \in S_m$.

We will need the following lemma which characterizes the bisymmetric operators.

**Lemma 1.2.25.** *[25] Suppose $L \in \mathcal{L}(\otimes^m V)$. Then $L$ is bisymmetric if and only if $L$ belongs to the linear closure of*

$$\{\otimes^m T : T \in \mathcal{L}(V)\}.$$

Using this it is easy to see that the symmetry class of tensors $V_\chi$ is an invariant subspace of $\otimes^m T$. Notice that $P(\sigma)$ is bisymmetric for all $\sigma \in S_m$. Therefore, using the previous lemma, $P(\sigma)$ commutes with $\otimes^m T$ for every $\sigma \in S_m$. Then $\otimes^m T$ commutes with any projection $K_\chi = \dfrac{\chi(\mathrm{id})}{m!} \sum_{\sigma \in S_m} \chi(\sigma) P(\sigma)$.

**Definition 1.2.26.** Let $V$ be a Hilbert space of dimension $n$ and let $1 \leq m \leq n$. Suppose $\chi$ is an irreducible character of $S_m$ and $T \in \mathcal{L}(V)$. The *induced transformation determined by* $\chi$ is the restriction of $\otimes^m T$ to the subspace $V_\chi$. This is represented by the symbol $K_\chi(T)$.

Now we list some properties of the induced transformation determined by the irreducible character $\chi$.

**Proposition 1.2.27.** *Suppose $V$ is a Hilbert space with dimension $n$ and let $1 \leq m \leq n$. Let $\chi$ be an irreducible character of $S_m$ and suppose that $S$ and $T$ are in $\mathcal{L}(V)$ and $v_1, \ldots v_m \in V$. Then*

1. $K_\chi(ST) = K_\chi(S) K_\chi(T)$,
2. $K_\chi(T)(v_1 * \cdots * v_m) = T(v_1) * \cdots * T(v_m)$,
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* $K_\chi(T)^{-1} = K_\chi(T^{-1})$.
5. *If $T$ is a unitary operator then $K_\chi(T)$ is unitary.*

We have already seen that if $E = \{e_1, e_2, \ldots, e_n\}$ is an orthonormal basis of the Hilbert space $V$ then the set

$$\{e_\alpha^\otimes : \alpha \in \Gamma_{m,n}\}$$

is an orthogonal basis of $\otimes^m V$, formed by decomposable tensors of $\otimes^m V$. On the other hand, if $\chi$ is a linear character then

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

is an orthogonal basis of $V_\chi$. Among other things, this is consequence of the fact that for every $\sigma \in S_m$, $e_{\alpha\sigma}^* = \chi(\sigma) e_\alpha^*$, when $\chi$ is linear. [25, p. 165]

In general, if $\chi$ does not have degree one, there are no known orthonormal bases of $V_\chi$ formed by decomposable symmetrized tensors, this means that the induced basis

$$\mathcal{E}' = \{e^*_\alpha : \alpha \in \widehat{\Delta}\},$$

is not, in general, an orthogonal basis of the space $V_\chi$, with the induced inner product in $V_\chi$.

However, if we apply the Gram-Schmidt orthonormalization procedure to the basis $\mathcal{E}'$, we obtain an orthonormal basis of the $m$-th $\chi$-symmetric tensor power of the vector space $V$, which we denote by

$$\mathcal{E} = \{v_\alpha : \alpha \in \widehat{\Delta}\}.$$

Let $T$ be an operator in $\mathcal{L}(V)$ and let $A$ be a complex $n \times n$ complex matrix. Suppose that $A$ represents the operator $T$ in the orthonormal basis $E$, i.e.

$$A = M(T; E).$$

Let $t = |\widehat{\Delta}|$. By the previous definitions, we have already seen that $K_\chi(T)$ is in $\mathcal{L}(V_\chi)$. Suppose $\mathcal{A}$ is a $t \times t$ complex matrix that represents the operator $K_\chi(T)$ in the orthonormal basis $\mathcal{E}$, i.e.

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

Our goal is to find a relation between the entries of $A$ and $\mathcal{A}$, which will lead us to the definition of the $m$-th $\chi$-symmetric tensor power of a matrix $A$. In order to construct this matrix, we study the immanant of a square matrix, in the next section.

## 1.3 Immanant

In this section, we present some properties of a multilinear function called the immanant. The immanant function can be looked at in two different, but equivalent ways. On the one hand it may be studied as a special case of a map called *generalized matrix function* which was first introduced by I. Schur, on the other hand it can be studied as a generalization of the determinant.

In this section we also define the $\chi$-symmetric tensor power of an $n \times n$ complex matrix, where $\chi$ is an irreducible character of $S_m$. Finally we establish a relation between these two concepts.

**Definition 1.3.1.** Let $A$ be an $n \times n$ complex matrix. Suppose $G$ is a subgroup of $S_n$ and $\chi$ a character of $G$. The *generalized matrix function* $d_\chi^G : M_n(\mathbb{C}) \longrightarrow \mathbb{C}$ is defined by

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

In the special case $G = S_n$ and $\chi$ irreducible, we have:

**Definition 1.3.2.** Let $A \in M_n(\mathbb{C})$ and $\chi$ be an irreducible character of $S_n$. We define the *immanant* of $A$ as:

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

The map $d_\chi : M_n(\mathbb{C}) \longrightarrow \mathbb{C}$ is a multilinear map on the columns (and also on the rows) of $A$.

**Example 1.3.3.** The determinant of $A$,

$$\det(A) = d_{\mathrm{sgn}}(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} \mathrm{sgn}(\sigma) a_{i\sigma(i)}$$

is the particular case of an immanant when $\chi$ is the alternating character.

If $\chi$ is the principal character then

$$\mathrm{per}(A) := d_1(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i\sigma(i)},$$

which is the *permanent* of the matrix $A$.

It is important to notice that these two immanants are the only immanants afforded by linear characters of $S_n$, which means, among other things, that in these cases $\chi$ is a homomorphism.
Since the immanant is a multilinear function on the columns (or rows) of the matrix $A$, all the properties of the determinant that are consequences of its multilinearity are still true for every immanant $d_\chi$.

**Proposition 1.3.4.** *Let $\chi$ be an irreducible character of $S_n$ and $A \in M_n(\mathbb{C})$. Then*

1. $d_\chi(A^T) = d_\chi(A)$;
2. $d_\chi(A^*) = \overline{d_\chi(A)}$.

*Proof.* Suppose $A \in M_n(\mathbb{C})$, we want to prove that

$$d_\chi(A^T) = d_\chi(A).$$

$$\begin{aligned}
d_\chi(A^T) &= \sum_{\sigma\in S_n} \chi(\sigma) \prod_{i=1}^{n} (A^T)_{i\sigma(i)} \\
&= \sum_{\sigma\in S_n} \chi(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i} \qquad (i = \sigma^{-1}(j)) \\
&= \sum_{\sigma\in S_n} \chi(\sigma) \prod_{j=1}^{n} a_{j\sigma^{-1}(j)} \qquad (\sigma = \tau^{-1}) \\
&= \sum_{\tau\in S_n} \chi(\tau^{-1}) \prod_{j=1}^{m} a_{j\tau(j)} \\
&= \sum_{\tau\in S_n} \chi(\tau) \prod_{j=1}^{n} a_{j\tau(j)} \\
&= d_\chi(A).
\end{aligned}$$

Recall that $\chi(\tau) = \chi(\tau^{-1})$, because $\chi$ is an irreducible character of the permutation group.

Now we will prove that $d_\chi(A^*) = \overline{d_\chi(A)}$.

$$\begin{aligned}
d_\chi(A^*) &= \sum_{\sigma\in S_n} \chi(\sigma) \prod_{i=1}^{n} (A^*)_{i\sigma(i)} \\
&= \sum_{\sigma\in S_n} \chi(\sigma) \prod_{i=1}^{n} \overline{a_{\sigma(i)i}} \\
&= \sum_{\sigma\in S_n} \chi(\sigma) \prod_{j=1}^{n} \overline{a_{j\sigma^{-1}(j)}} \qquad (\sigma(i) = j) \\
&= \sum_{\tau\in S_n} \chi(\tau^{-1}) \prod_{j=1}^{n} \overline{a_{j\tau(j)}} \qquad (\sigma = \tau^{-1}) \\
&= \sum_{\tau\in S_n} \overline{\chi(\tau)} \prod_{j=1}^{n} \overline{a_{j\tau(j)}} \\
&= \overline{\sum_{\tau\in S_n} \chi(\tau) \prod_{j=1}^{n} a_{j\tau(j)}} \\
&= \overline{d_\chi(A)}.
\end{aligned}$$

This concludes our proof. □

Using the last proposition, we can see that for the permanent function:

1. $\mathrm{per}(A^T) = \mathrm{per}(A)$,
2. $\mathrm{per}(A^*) = \overline{\mathrm{per}(A)}$.

Now, if we consider the determinant function, the Laplace Expansion gives us a formula to calculate the determinant of a matrix of order $n$ using determinants of submatrices of order $n-1$. In general, there is no natural way to associate a character of $S_n$ with a character of $S_{n-1}$, so there is no natural way to relate the immanant of an $n \times n$ matrix with immanants of its submatrices. We intend to generalize the Laplace Expansion for every immanant. In order to do that we need to relate the immanant of the matrix $A$ with the immanant of matrices of the same size.

First we need to introduce some notation and definitions.
We denote by $A(i|j)$ the $n \times n$ matrix that is obtained from $A$ by replacing

the $i$-th row and $j$-th column with zero entries, except entry $(i, j)$ which we set to 1. For example, suppose

$$A = \begin{pmatrix} 2 & -1 & 0 & 3 \\ 1 & 1 & 1 & 0 \\ 0 & -2 & -1 & 3 \\ -2 & -2 & 6 & 1 \end{pmatrix}$$

then

$$A(3|2) = \begin{pmatrix} 2 & 0 & 0 & 3 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & 0 & 6 & 1 \end{pmatrix}.$$

**Definition 1.3.5.** Let $A \in M_n(\mathbb{C})$ and let $\chi$ be an irreducible character of $S_n$. The *immanantal adjoint* of $A$, $\mathrm{adj}_\chi(A)$ is the $n \times n$ matrix in which the entry $(i, j)$ is $d_\chi(A(i|j))$.

This definition agrees with the definition of permanental adjoint in [25], but not with the usual adjugate matrix. In this case we would need to consider the transpose matrix. This is only a matter of convention.

Using the fact that the immanant is a multilinear function, we get the following result.

**Proposition 1.3.6.** *(Laplace Expansion for Immanants) Let $A \in M_n(\mathbb{C})$ and $\chi$ an irreducible character of $S_n$.*
*For every $1 \le j \le n$,*

$$d_\chi(A) = \sum_{i=1}^{n} a_{ij} d_\chi(A(i|j)).$$

*For every $1 \le i \le n$,*

$$d_\chi(A) = \sum_{j=1}^{n} a_{ij} d_\chi(A(i|j)).$$

*Proof.* First notice that

$$a_{ij} d_\chi(A(i|j)) = d_\chi \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \dots & \dots & \dots & & \dots & \dots\dots \\ 0 & 0 & \dots & a_{ij} & \dots & 0 \\ \dots & \dots & \dots & & \dots & \dots\dots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{nn} \end{pmatrix}.$$

Since $d_\chi$ is multilinear we have

$$d_\chi(A) = d_\chi \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \dots & \dots & \dots & & \dots & \dots\dots \\ a_{i1} & 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & & \dots & \dots\dots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{nn} \end{pmatrix} + \dots$$

$$+ d_\chi \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \dots & \dots & \dots & & \dots & \dots\dots \\ 0 & 0 & \dots & a_{ij} & \dots & 0 \\ \dots & \dots & \dots & & \dots & \dots\dots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{nn} \end{pmatrix} + \dots$$

$$+ d_\chi \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \dots & \dots & \dots & & \dots & \dots\dots \\ 0 & 0 & \dots & 0 & \dots & a_{in} \\ \dots & \dots & \dots & & \dots & \dots\dots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{nn} \end{pmatrix}.$$

So we have the desired result. □

Now we want to go further in the sense of a Generalized Laplace Expansion, that gives us a formula to calculate the determinant of a matrix of order $n$, using products of determinants of submatrices with order $k$ and $n-k$ of $A$, $1 \leq k \leq n$. This expansion was first proved for the determinant and the same arguments were later used to prove the corresponding generalized expansion for the permanent. We now present both of these formulas that can be found in [23] and in [28].

**Definition 1.3.7.** Let $X$ be a $n \times n$ complex matrix, $k$ a natural number, $1 \leq k \leq n$ and $\alpha, \beta \in Q_{k,n}$. We denote by

$$X[\alpha|\beta]$$

the $k \times k$ matrix obtained from $X$ by picking the rows $\alpha(1), \ldots, \alpha(k)$ and the columns $\beta(1), \ldots, \beta(k)$.

We denote by

$$X(\alpha|\beta)$$

the $(n-k)\times(n-k)$ matrix obtained from $X$ by deleting the rows $\alpha(1), \ldots, \alpha(k)$ and the columns $\beta(1), \ldots, \beta(k)$.

**Example 1.3.8.** Suppose $A = \begin{pmatrix} 2 & -1 & 0 & 3 \\ 1 & 1 & 1 & 0 \\ 0 & -2 & -1 & 3 \\ -2 & -2 & 6 & 1 \end{pmatrix}$ and $\alpha = (1,3)$, $\beta = (2,3)$. Then

$$A[\alpha|\beta] = \begin{pmatrix} -1 & 0 \\ -2 & -1 \end{pmatrix} \; A(\alpha|\beta) = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}.$$

For $\alpha \in Q_{k,n}$, denote $|\alpha| = \alpha(1)+\ldots+\alpha(k)$. We now look at some results that are already known.

**Theorem 1.3.9** ([23], [28])**.** *[Generalized Laplace Expansion for Determinants and Permanents] Fixing* $\alpha \in Q_{k,n}$

$$\det X = (-1)^{|\alpha|} \sum_{\beta \in Q_{k,n}} (-1)^{|\beta|} \det(X[\alpha|\beta]) \det(X(\alpha|\beta)). \tag{1.6}$$

*and*

$$\operatorname{per} X = \sum_{\beta \in Q_{k,n}} \operatorname{per}(X[\alpha|\beta]) \operatorname{per}(X(\alpha|\beta)). \tag{1.7}$$

**Example 1.3.10.** Let $n = 4$ and $k = 2$, then $Q_{2,4} = \{(1,2), (1,3), (1,4), (2,3), (2,4), (3,4)\}$ fixing $\alpha = (1,2)$, $A = \begin{pmatrix} 1 & -1 & 0 & 1 \\ 2 & 1 & 1 & 2 \\ 1 & -1 & 0 & 0 \\ 3 & 1 & 2 & -1 \end{pmatrix}$ and using the previous formula for the determinant,

$$\begin{aligned}
\det A &= (-1)^3 \sum_{\beta \in Q_{2,4}} (-1)^{|\beta|} \det(A[(1,2)|\beta]) \det(A((1,2)|\beta)) \\
&= -(-\det(A[(1,2)|(1,2)]) \det(A((1,2)|(1,2))) \\
&+ \det(A[(1,2)|(1,3)]) \det(A((1,2)|(1,3))) \\
&- \det(A[(1,2)|(1,4)]) \det(A((1,2)|(1,4))) \\
&- \det(A[(1,2)|(2,3)]) \det(A((1,2)|(2,3))) \\
&+ \det(A[(1,2)|(2,4)]) \det(A((1,2)|(2,4))) \\
&- \det(A[(1,2)|(3,4)]) \det(A((1,2)|(3,4)))) \\
&= \det\begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} \det\begin{pmatrix} 0 & 0 \\ 2 & -1 \end{pmatrix} - \det\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \det\begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix} \\
&+ \det\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \det\begin{pmatrix} -1 & 0 \\ 1 & 2 \end{pmatrix} + \det\begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix} \det\begin{pmatrix} 1 & 0 \\ 3 & -1 \end{pmatrix} \\
&- \det\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \det\begin{pmatrix} 1 & 0 \\ 3 & 2 \end{pmatrix} + \det\begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix} \det\begin{pmatrix} 1 & -1 \\ 3 & 1 \end{pmatrix} = -8
\end{aligned}$$

If we analyse both of the proofs, that can be found in [28] and [23] we can see that the similarity between them is due to the fact that the determinant and the permanent of a direct sum of matrices is the product of the determinants, or the permanents, of the block summands. That is

$$\det\begin{pmatrix} A & O \\ O & B \end{pmatrix} = \det(A \oplus B) = \det(A)\det(B),$$

$$\operatorname{per}\begin{pmatrix} A & O \\ O & B \end{pmatrix} = \operatorname{per}(A \oplus B) = \operatorname{per}(A)\operatorname{per}(B).$$

However, if $\chi$ is any other irreducible character, there is no clear general relation between the immanant of $A$ and the immanant of any submatrix of $A$. In fact there is no relation between a character of $S_n$ and a character of $S_m$, for $m < n$. So the Generalized Laplace Expansion formula for any immanant is a little more complicated.

The idea to overcome this drawback is to calculate the immanant of $A$ using matrices of order $n$ in which some of the entries are equal to the entries of $A$ and others are replaced by zeros.

Let $1 \leq k \leq n$, $\alpha \in Q_{k,n}$, we denote by $\operatorname{Im}\alpha$ the image of $\alpha$.

**Definition 1.3.11.** Let $1 \leq k \leq n$, $\alpha, \beta \in Q_{k,n}$. We denote by $S_{\alpha,\beta}$ the subset of $S_n$ defined has

$$S_{\alpha,\beta} = \{\sigma \in S_n : \sigma(\operatorname{Im}\alpha) = \operatorname{Im}\beta\}.$$

**Example 1.3.12.** Let $n = 5$, $k = 3$, then $\alpha, \beta \in Q_{3,5}$. Suppose $\alpha = (1, 3, 5)$ and $\beta = (2, 4, 5)$. It is easy to check that if $\sigma = (12)(34)$ and $\tau = (1234)$, then $\sigma, \tau \in S_{\alpha,\beta}$.

Fixing $\alpha \in Q_{k,n}$, we have the following result.

**Lemma 1.3.13.** *For every $\alpha \in Q_{k,n}$, the set $\{S_{\alpha,\beta} : \beta \in Q_{k,n}\}$ is a partition of $S_n$.*

*Proof.* We first prove that

1. Let $\beta, \gamma \in Q_{k,n}$, if $\beta \neq \gamma$ then $S_{\alpha,\beta} \cap S_{\alpha,\gamma} = \emptyset$.
   Suppose $\sigma \in S_{\alpha,\beta} \cap S_{\alpha,\gamma}$. Then $\sigma(\operatorname{Im}\alpha) = \operatorname{Im}\beta = \operatorname{Im}\gamma$, and thus $\operatorname{Im}\beta = \operatorname{Im}\gamma$. Since $\beta, \gamma \in Q_{k,n}$, it follows that $\beta = \gamma$.
   Now we prove that

2. $S_n = \displaystyle\bigcup_{\beta \in Q_{k,n}} S_{\alpha,\beta}$.

   Take $\pi \in S_n$ with $\pi(\operatorname{Im}\alpha) = \{j_1, \ldots, j_k\}$ and suppose $j_1 < \ldots < j_k$.
   Let $\gamma \in Q_{k,n}$ such that $\gamma(i) = j_i$, for $i = 1, \ldots, k$. Therefore $\pi \in S_{\alpha,\gamma}$ and $S_n \subseteq \displaystyle\bigcup_{\beta \in Q_{k,n}} S_{\alpha,\beta}$.

The other inclusion is trivial. □

Now, for every $\alpha \in Q_{k,n}$ denote by $\overline{\operatorname{Im}\alpha}$ the complement of $\operatorname{Im}\alpha$ that is

$$\overline{\operatorname{Im}\alpha} = \{1, 2, \ldots, n\} \setminus \operatorname{Im}\alpha.$$

**Lemma 1.3.14.** *Let $1 \leq k \leq n$ and $\alpha, \beta \in Q_{k,n}$. If $\sigma \in S_{\alpha,\beta}$ then*

$$\sigma(\overline{\operatorname{Im}\alpha}) = \overline{\operatorname{Im}\beta}.$$

*Proof.* Suppose that $l \in \overline{\operatorname{Im}\alpha}$ and $\sigma(l) = j_l \in \operatorname{Im}\beta$. We have $\sigma \in S_{\alpha,\beta}$ so we can find $i \in \operatorname{Im}\alpha$ such that $\sigma(i) = j_l = \sigma(l)$. But $i \neq l$. This is a contradiction, because $\sigma \in S_n$, and therefore is injective. $\square$

Since $\{S_{\alpha,\beta} : \beta \in Q_{k,n}\}$ is a partition of $S_n$ we have

$$|\{S_{\alpha,\beta} : \beta \in Q_{k,n}\}| = |Q_{k,n}| = \frac{n!}{k!(n-k)!}.$$

In an analogous way, we can prove the same results if we fix $\beta$ instead of $\alpha$. That is, if we consider the set $\{S_{\alpha,\beta} : \alpha \in Q_{k,n}\}$.

Now, we can conclude that for every $\alpha, \beta \in Q_{k,n}$ the value

$$\sum_{\sigma \in S_{\alpha,\beta}} \chi(\sigma) \prod_{t=1}^{n} a_{t\sigma(t)}$$

*does not depend* on the values of the following entries of the matrix $A$:

I. The $k$ rows of $A$ that are in $\operatorname{Im}\alpha$ and the $n-k$ columns of $A$ in $\overline{\operatorname{Im}\beta}$.

II. The $n-k$ rows of $A$ that are in $\overline{\operatorname{Im}\alpha}$ and the $k$ columns of $A$ with index in $\operatorname{Im}\beta$.

We now denote by

$$A\{\alpha|\beta\} = (a^{+}_{ij})$$

the matrix of order $n$ obtained by replacing in the matrix $A$ every entry in I and II by zeros.

**Example 1.3.15.** Suppose

$$A = \begin{pmatrix} 8 & -2 & 1 & 3 & -5 \\ 1 & 2 & -3 & 3 & -2 \\ 3 & 1 & -2 & 4 & -1 \\ 3 & 1 & -1 & 2 & 7 \\ 3 & -3 & 4 & 5 & 1 \end{pmatrix}.$$

We have that $n = 5$, suppose $k = 2$ and let $\alpha = (2,3)$ and $\beta = (1,5)$. Then

$$A\{(2,3)|(1,5)\} = \begin{pmatrix} 0 & -2 & 1 & 3 & 0 \\ 1 & 0 & 0 & 0 & -2 \\ 3 & 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 2 & 0 \\ 0 & -3 & 4 & 5 & 0 \end{pmatrix}.$$

**Lemma 1.3.16.** *With the previously established notation, we have that for each* $\alpha, \beta \in Q_{k,n}$*,*

$$\sum_{\sigma \in S_{\alpha,\beta}} \chi(\sigma) \prod_{t=1}^{n} a_{t\sigma(t)} = d_\chi(A\{\alpha|\beta\}).$$

*Proof.* Using the definition of the immanant and the fact that $S_n = \bigcup_{\gamma \in Q_{k,n}} S_{\alpha,\gamma}$, we have that

$$\begin{aligned} d_\chi(A\{\alpha|\beta\}) &= \sum_{\sigma \in S_n} \chi(\sigma) \prod_{t=1}^{n} a^+_{t\sigma(t)} \\ &= \sum_{\sigma \in \bigcup_{\gamma \in Q_{k,n}} S_{\alpha,\gamma}} \chi(\sigma) \prod_{t=1}^{n} a^+_{t\sigma(t)} \\ &= \sum_{\gamma \in Q_{k,n}} \sum_{\sigma \in S_{\alpha,\gamma}} \chi(\sigma) \prod_{t=1}^{n} a^+_{t\sigma(t)}. \end{aligned}$$

Now take $\delta \in Q_{k,n}$ such that $\delta \neq \beta$ and $\sigma \in S_{\alpha,\delta}$. Then $\prod_{t=1}^{n} a^+_{t\sigma(t)} = 0$, because at least one of the factors is zero, by the definition of the matrix $A\{\alpha|\beta\}$.
Therefore

$$\sum_{\sigma \in S_{\alpha,\gamma}} \chi(\sigma) \prod_{t=1}^{n} a^+_{t\sigma(t)} = 0,$$

for every $\gamma \in Q_{k,n} \setminus \{\beta\}$.

Moreover, for $A\{\alpha|\beta\}$, if $\sigma \in S_{\alpha,\beta}$ then $a^+_{t\sigma(t)} = a_{t\sigma(t)}$. So

$$d_\chi(A\{\alpha|\beta\}) = \sum_{\sigma \in S_{\alpha,\beta}} \chi(\sigma) \prod_{t=1}^{n} a_{t\sigma(t)}.$$

This concludes the proof. □

**Proposition 1.3.17.** *Let* $A$ *be an* $n \times n$ *complex matrix and suppose* $1 \leq k \leq n$*. Suppose* $\alpha \in Q_{k,n}$*. Then*

$$d_\chi(A) = \sum_{\beta \in Q_{k,n}} d_\chi(A\{\alpha|\beta\}).$$

*Proof.*

$$
\begin{aligned}
d_\chi(A) &= \sum_{\sigma\in S_n} \chi(\sigma) \prod_{t=1}^{n} a_{t\sigma(t)} \\
&= \sum_{\beta\in Q_{k,n}} \sum_{\sigma\in S_{\alpha,\beta}} \chi(\sigma) \prod_{t=1}^{n} a_{t\sigma(t)} \\
&= \sum_{\beta\in Q_{k,n}} d_\chi(A\{\alpha|\beta\}). \qquad (1.8)
\end{aligned}
$$

This concludes our proof. □

Now we construct matrices of order $n$ using matrices of order $k$ and order $n-k$. In general this could be done by using the usual direct sum of matrices. We introduce a generalization of this concept.

**Definition 1.3.18.** Let $\alpha, \beta \in Q_{k,n}$, and let $A$ be a $k \times k$ matrix and let $B$ be a $(n-k) \times (n-k)$ matrix. Denote by $\bar{\alpha}$ be the unique element of $Q_{n-k,n}$ with $\operatorname{Im} \overline{\alpha} = \overline{\operatorname{Im} \alpha}$.

We define

$$A \bigoplus_{\alpha|\beta} B = (x_{ij}),$$

as a $n \times n$ matrix such that

- $x_{ij} = 0$ if $i \in \operatorname{Im} \alpha$ and $j \notin \operatorname{Im} \beta$;
- $x_{ij} = 0$ if $i \notin \operatorname{Im} \alpha$ and $j \in \operatorname{Im} \beta$;
- $x_{ij} = a_{\alpha^{-1}(i)\beta^{-1}(j)}$ if $i \in \operatorname{Im} \alpha$ and $j \in \operatorname{Im} \beta$;
- $x_{ij} = b_{\overline{\alpha}^{-1}(i)\overline{\beta}^{-1}(j)}$ if $i \notin \operatorname{Im} \alpha$ and $j \notin \operatorname{Im} \beta$.

In a sense, we place $A$ in rows $\alpha$ and columns $\beta$ and we place $B$ in rows $\bar{\alpha}$ and columns $\bar{\beta}$.

**Example 1.3.19.** Let

$$A = \begin{pmatrix} 1 & -2 \\ 3 & -1 \end{pmatrix} \qquad B = \begin{pmatrix} -2 & 1 & 3 \\ 1 & -1 & 2 \\ -3 & 4 & 6 \end{pmatrix},$$

then $k = 2$ and $n = 5$. Suppose $\alpha = (2, 3)$ and $\beta = (1, 5)$. So we have

$$A \bigoplus_{\alpha|\beta} B = \begin{pmatrix} 0 & -2 & 1 & 3 & 0 \\ 1 & 0 & 0 & 0 & -2 \\ 3 & 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 2 & 0 \\ 0 & -3 & 4 & 6 & 0 \end{pmatrix}.$$

If $\alpha = \beta = (1, \ldots, k)$, this is the usual direct sum of $A$ and $B$, that is

$$A \bigoplus_{(1,\ldots,k)|(1,\ldots,k)} B = \begin{pmatrix} A & O \\ O & B \end{pmatrix}.$$

Also it is easy to see that if $\alpha \neq \beta$ then

$$A \bigoplus_{\alpha|\beta} B \neq A \bigoplus_{\beta|\alpha} B.$$

The following result is easy to verify.

**Lemma 1.3.20.** *Let $X$ be an $n \times n$ complex matrix and let $1 \leq k \leq n$ , $\alpha, \beta \in Q_{k,n}$. Then we have*

$$X\{\alpha|\beta\} = X[\alpha|\beta] \bigoplus_{\alpha|\beta} X(\alpha|\beta).$$

*Proof.* First notice that the matrices $X[\alpha|\beta]$ and $X(\alpha|\beta)$ have order $k$ and $n - k$, respectively. So both matrices $X\{\alpha|\beta\}$ and $X[\alpha|\beta] \bigoplus_{\alpha|\beta} X(\alpha|\beta)$ have order $n$.
The $(i, j)$-entry of $X\{\alpha|\beta\}$ is zero if

I. $i \in \operatorname{Im} \alpha$ and $j \in \overline{\operatorname{Im} \beta}$, or

II. $i \in \overline{\operatorname{Im} \alpha}$ and $j \in \operatorname{Im} \beta$.

All the other entries of $X\{\alpha|\beta\}$ are equal to the entries of the matrix $X$. So rephrasing the last sentences, we have that the $(i, j)$-entry of $X\{\alpha|\beta\}$ is equal to

- 0 if $i \in \operatorname{Im} \alpha$ and $j \notin \operatorname{Im} \beta$;

- $0$ if $i \notin \operatorname{Im} \alpha$ and $j \in \operatorname{Im} \beta$;
- $x_{\alpha^{-1}(i)\beta^{-1}(j)}$ if $i \in \operatorname{Im} \alpha$ and $j \in \operatorname{Im} \beta$;
- $x_{\overline{\alpha}^{-1}(i)\overline{\beta}^{-1}(j)}$ if $i \notin \operatorname{Im} \mathrm{id}$ and $j \notin \operatorname{Im} \beta$,

These entries are exactly the same as the entries of the $n \times n$ complex matrix $X[\alpha|\beta] \bigoplus_{\alpha|\beta} X(\alpha|\beta)$ □

Now we can state *the Laplace expansion for immanants.*

**Theorem 1.3.21** (Generalized Laplace Expansion)**.** *Let $X$ be an $n \times n$ complex matrix, let $1 \leq k \leq n$ , and $\alpha$ a fixed element in $Q_{k,n}$. Suppose $\chi$ is an irreducible character of $S_n$. Then*

$$d_\chi(X) = \sum_{\beta \in Q_{k,n}} d_\chi(X[\alpha|\beta] \bigoplus_{\alpha|\beta} X(\alpha|\beta)) = \sum_{\beta \in Q_{k,n}} d_\chi(X\{\alpha|\beta\}). \tag{1.9}$$

*and*

$$d_\chi(X) = \sum_{\beta \in Q_{k,n}} d_\chi(X[\beta|\alpha] \bigoplus_{\beta|\alpha} X(\beta|\alpha)) = \sum_{\beta \in Q_{k,n}} d_\chi(X\{\beta|\alpha\}). \tag{1.10}$$

**Example 1.3.22.** Let $A$ be a matrix of order 4. Let $k = 2$ and $\alpha = (1, 2)$ in $Q_{2,4}$. Then, we have

$$\begin{aligned} d_\chi(A) &= d_\chi \begin{pmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{12} & a_{22} & 0 & 0 \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & a_{43} & a_{44} \end{pmatrix} + d_\chi \begin{pmatrix} a_{11} & 0 & a_{13} & 0 \\ a_{12} & 0 & a_{23} & 0 \\ 0 & a_{32} & 0 & a_{34} \\ 0 & a_{42} & 0 & a_{44} \end{pmatrix} \\ &+ d_\chi \begin{pmatrix} a_{11} & 0 & 0 & a_{14} \\ a_{12} & 0 & 0 & a_{24} \\ 0 & a_{32} & a_{33} & 0 \\ 0 & a_{42} & a_{43} & 0 \end{pmatrix} + d_\chi \begin{pmatrix} 0 & a_{12} & a_{13} & 0 \\ 0 & a_{22} & a_{23} & 0 \\ a_{31} & 0 & 0 & a_{34} \\ a_{41} & 0 & 0 & a_{44} \end{pmatrix} \\ &+ d_\chi \begin{pmatrix} 0 & a_{12} & 0 & a_{14} \\ 0 & a_{22} & 0 & a_{24} \\ a_{31} & 0 & a_{33} & 0 \\ a_{41} & 0 & a_{43} & 0 \end{pmatrix} + d_\chi \begin{pmatrix} 0 & 0 & a_{13} & a_{14} \\ 0 & 0 & a_{23} & a_{24} \\ a_{31} & a_{32} & 0 & 0 \\ a_{41} & a_{42} & 0 & 0 \end{pmatrix} \end{aligned}$$

Now we show that Proposition 1.3.6 is a particular case of the Generalized Laplace Expansion.

Take $k = 1$, then $Q_{1,n} = \{(1), (2), \ldots, (n)\}$. Suppose $\alpha, \beta \in Q_{1,n}$, with $\alpha = (i)$ and $\beta = (j)$. First, notice that for every $X \in M_n(\mathbb{C})$, we have that

$$d_\chi(X\{\alpha|\beta\}) = d_\chi(X\{(i)|(j)\}) = x_{ij} d_\chi(X(i|j)).$$

So, using the previous theorem

$$\begin{aligned} d_\chi(X) &= \sum_{\alpha \in Q_{k,n}} d_\chi(X\{\alpha|\beta\}) \\ &= \sum_{i=1}^{n} d_\chi(X\{(i)|(j)\}) \\ &= \sum_{i=1}^{n} x_{ij} d_\chi(X(i|j)). \end{aligned}$$

The last equality is due to the multilinearity of the immanant.

We also want to prove that the generalized Laplace expansion for determinants is a particular case of formula 1.9. For that purpose, we list some properties of the matrix $X\{\alpha|\beta\}$.

**Proposition 1.3.23.** *Let $\alpha, \beta, \alpha', \beta' \in Q_{k,n}$. Then we have*

1. $X\{\alpha|\beta\}[\alpha|\beta] = X[\alpha|\beta]$.

2. $X\{\alpha|\beta\}(\alpha|\beta) = X(\alpha|\beta)$.

3. *If $\beta \neq \beta'$, then both matrices $X\{\alpha|\beta\}[\alpha|\beta']$ and $X\{\alpha|\beta\}(\alpha|\beta')$ have a zero column.*

4. *If $\alpha \neq \alpha'$, then both matrices $X\{\alpha|\beta\}[\alpha'|\beta]$ and $X\{\alpha|\beta\}(\alpha'|\beta)$ have a zero row.*

*Proof.* 1. Let us denote as before $X\{\alpha|\beta\} = (x^+_{ij})$, $i, j = 1, \ldots, n$. Then

$$X\{\alpha|\beta\}[\alpha|\beta] = \begin{pmatrix} x^+_{\alpha(1)\beta(1)} & x^+_{\alpha(1)\beta(2)} & \cdots & x^+_{\alpha(k)\beta(k)} \\ x^+_{\alpha(2)\beta(1)} & x^+_{\alpha(2)\beta(2)} & \cdots & x^+_{\alpha(2)\beta(k)} \\ \cdots & \cdots & \cdots & \cdots \\ x^+_{\alpha(k)\beta(1)} & x^+_{\alpha(k)\beta(2)} & \cdots & x^+_{\alpha(k)\beta(k)} \end{pmatrix}.$$

By definition of $X\{\alpha|\beta\}$, for every $s, t = 1, \ldots, k$,

$$x^+_{\alpha(s)\beta(t)} = x_{\alpha(s)\beta(t)},$$

so

$$X\{\alpha|\beta\}[\alpha|\beta] = \begin{pmatrix} x_{\alpha(1)\beta(1)} & x_{\alpha(1)\beta(2)} & \cdots & x_{\alpha(k)\beta(k)} \\ x_{\alpha(2)\beta(1)} & x_{\alpha(2)\beta(2)} & \cdots & x_{\alpha(2)\beta(k)} \\ \cdots & \cdots & \cdots & \cdots \\ x_{\alpha(k)\beta(1)} & x_{\alpha(k)\beta(2)} & \cdots & x_{\alpha(k)\beta(k)} \end{pmatrix} = X[\alpha|\beta].$$

2. With similar arguments we can prove that $X\{\alpha|\beta\}(\alpha|\beta) = X(\alpha|\beta)$.

3. Now suppose that $\beta' \neq \beta$, then

$$X\{\alpha|\beta\}[\alpha|\beta'] = \begin{pmatrix} x^+_{\alpha(1)\beta'(1)} & x^+_{\alpha(1)\beta'(2)} & \cdots & x^+_{\alpha(k)\beta'(k)} \\ x^+_{\alpha(2)\beta'(1)} & x^+_{\alpha(2)\beta'(2)} & \cdots & x^+_{\alpha(2)\beta'(k)} \\ \cdots & \cdots & \cdots & \cdots \\ x^+_{\alpha(k)\beta'(1)} & x^+_{\alpha(k)\beta'(2)} & \cdots & x^+_{\alpha(k)\beta'(k)} \end{pmatrix}.$$

Since $\beta' \neq \beta$, there is $j$ such that $j = \beta'(t)$ and $j \notin \operatorname{Im}\beta$. So the elements of the $j$-th column of $X\{\alpha|\beta\}[\alpha|\beta']$ are $x^+_{\alpha(s)j}$, that are equal to zero by the definition of $X\{\alpha|\beta\}$.

4. Analogous to 3.

This concludes the proof. □

We can now check that this formula generalizes the known Laplace formulas for the determinant and the permanent (see [23] and [28]).

If $\chi = \operatorname{sgn}$, then $d_{\operatorname{sgn}} = \det$. For $\alpha \in Q_{k,n}$, recall that $|\alpha| = \alpha(1) + \ldots + \alpha(k)$.

Fixing $\alpha \in Q_{k,n}$

$$\begin{aligned}
\det X &= \sum_{\beta \in Q_{k,n}} \det(X\{\alpha|\beta\}) \\
&= (-1)^{|\alpha|} \sum_{\beta \in Q_{k,n}} \sum_{\gamma \in Q_{k,n}} (-1)^{|\gamma|} \det(X\{\alpha|\beta\}[\alpha|\gamma]) \det(X\{\alpha|\beta\}(\alpha|\gamma)) \\
&= (-1)^{|\alpha|} \sum_{\beta \in Q_{k,n}} (-1)^{|\beta|} \det(X\{\alpha|\beta\}[\alpha|\beta]) \det(X\{\alpha|\beta\}(\alpha|\beta)) \\
&= (-1)^{|\alpha|} \sum_{\beta \in Q_{k,n}} (-1)^{|\beta|} \det(X[\alpha|\beta]) \det(X(\alpha|\beta)). \qquad (1.11)
\end{aligned}$$

The first equality follows from Theorem 1.9, the second from 1.11, the third and fourth from Proposition 1.3.23.

This is exactly the expression of the the Generalized Laplace Expansion for determinants. With similar arguments we can prove the result for the Generalized Laplace Expansion of the permanent.

Now we turn to another kind of properties of the immanants, that is the relation between immanants and decomposable symmetrized tensors. This question arises because both of them are associated with an irreducible character of the permutation group. These relations will allow us to state the famous Binet-Cauchy theorem that was first stated for determinants but it is true for every immanant, with the convenient notation.
The following results can be found in [12] and [25].

**Proposition 1.3.24.** *Let $A \in M_m(\mathbb{C})$ and $V$ a Hilbert space. Let $u_1, u_2, \ldots, u_m, v_1, v_2, \ldots, v_m$ be vectors in $V$ such that*

$$(a_{ij}) = \langle u_i, v_j \rangle,$$

*$i, j = 1, 2, \ldots m$.*

*Let $\chi$ be an irreducible character of $S_m$. Then*

$$\frac{m!}{\chi(\mathrm{id})} \langle u_1 * u_2 * \ldots * u_m, v_1 * v_2 * \ldots * v_m \rangle = d_\chi(A),$$

*with respect to the induced inner product in $V_\chi$.*

*Proof.* Suppose $A$ is an $m \times m$ complex matrix and $(a_{ij}) = \langle u_i, v_j \rangle$. Since $K_\chi$ is an orthogonal projection,

$$\begin{aligned}
& \langle u_1 * u_2 * \ldots * u_m, v_1 * v_2 * \ldots * v_m \rangle = \\
= \;& \langle K_\chi(u_1 \otimes u_2 \otimes \ldots \otimes u_m), K_\chi(v_1 \otimes v_2 \otimes \ldots \otimes v_m) \rangle \\
= \;& \langle u_1 \otimes u_2 \otimes \ldots \otimes u_m, K_\chi^* K_\chi(v_1 \otimes v_2 \otimes \ldots \otimes v_m) \rangle \\
= \;& \langle u_1 \otimes u_2 \otimes \ldots \otimes u_m, K_\chi^2(v_1 \otimes v_2 \otimes \ldots \otimes v_m) \rangle \\
= \;& \langle u_1 \otimes u_2 \otimes \ldots \otimes u_m, K_\chi(v_1 \otimes v_2 \otimes \ldots \otimes v_m) \rangle \\
= \;& \frac{\chi(\mathrm{id})}{m!} \sum_{\sigma \in S_m} \overline{\chi(\sigma)} \prod_{t=1}^{m} \langle u_t, v_{\sigma^{-1}(t)} \rangle \\
= \;& \frac{\chi(\mathrm{id})}{m!} \sum_{\sigma \in S_m} \chi(\sigma) \prod_{t=1}^{m} \langle u_t, v_{\sigma(t)} \rangle \\
= \;& \frac{\chi(\mathrm{id})}{m!} d_\chi(A).
\end{aligned}$$

This concludes our proof. □

**Corollary 1.3.25.** *Let $V$ be an $m$-dimensional Hilbert space and suppose that $E = \{e_1, e_2, \ldots, e_m\}$ is an orthonormal basis of $V$. Let $\chi$ be an irreducible character of $S_m$. Take $T \in \mathcal{L}(V)$ and let $A^T$ be the matrix of the operator $T$ in the basis $E$. Then*

$$d_\chi(A) = \frac{m!}{\chi(\mathrm{id})} \langle K_\chi(T)(e_1 * e_2 * \ldots * e_m), e_1 * e_2 * \ldots * e_m \rangle.$$

*Proof.* We have that $A^T = M(T; E)$, and $E$ is an orthonormal basis of $V$ so the $(i, j)$ - entry of $A$ is

$$(a_{ij}) = \langle T(e_i), e_j \rangle.$$

The result follows from the theorem by setting $v_i = e_i$ and $u_i = T(e_i)$, for every $i = 1, 2, \ldots, m$. □

The next two technical results will allows us to prove the Cauchy-Binet theorem.

**Lemma 1.3.26** (Parseval's Identity)**.** *Let $E = \{e_1, e_2, \ldots, e_n\}$ be an orthonormal basis of the Hilbert space $V$. If $v, w \in V$ then*

$$\langle v, w \rangle = \sum_{i=1}^{n} \langle v, e_i \rangle \langle e_i, w \rangle.$$

Now we present a result that establishes a relation between the induced transformation $K_\chi(T)$ and the immanant $d_\chi$.

**Lemma 1.3.27.** *Let $V$ be an $n$ dimensional Hilbert space and $E = \{e_1, \ldots, e_n\}$ an orthonormal basis of $V$. For $1 \le m \le n$, let $\chi$ be an irreducible character of $S_m$. Take $T \in \mathcal{L}(V)$ and let $A$ be the matrix of the operator $T$ in the basis $E$. For every $\alpha, \beta \in \widehat{\Delta}$, we have*

$$\langle K_\chi(T)(e^*_\alpha), e^*_\beta\rangle = \frac{\chi(\mathrm{id})}{m!} d_\chi(A^T[\alpha|\beta]).$$

*Proof.* By the properties of the operator $K_\chi(T)$,

$$\langle K_\chi(T)(e^*_\alpha), e^*_\beta\rangle = \langle T(e_{\alpha(1)}) * T(e_{\alpha(2)}) * \ldots * T(e_{\alpha(m)}), e_{\beta(1)} * e_{\beta(2)} * \ldots * e_{\beta(m)}\rangle.$$

On the other hand, since $A = M(T; E)$ and $E$ is orthonormal, for every $i, j = 1, 2, \ldots, n$,

$$(a_{ij}) = \langle T(e_j), e_i\rangle.$$

So the $(i, j)$ entries of the matrix $A[\alpha|\beta]$ are

$$(a_{\alpha(i)\beta(j)}) = \langle T(e_{\beta(j)}), e_{\alpha(i)}\rangle.$$

Consequently the $(i, j)$ entries of the matrix $A^T[\alpha|\beta]$ are

$$(a_{\beta(j)\alpha(i)}) = \langle T(e_{\alpha(i)}), e_{\beta(j)}\rangle.$$

So by Theorem 1.3.24 we have that

$$\langle K_\chi(T)(e^*_\alpha), e^*_\beta\rangle = \frac{\chi(\mathrm{id})}{m!} d_\chi(A^T[\alpha|\beta]).$$

□

It is a well known fact that the determinant is a multiplicative multilinear map, that is, for $A$ and $B$ $n \times n$ matrices,

$$\det(AB) = \det(A)\det(B).$$

In fact it is known that the determinant is the only multiplicative generalized matrix function. So this last equality is false if we consider any other immanant. However, there is a result that includes this property which is called the Cauchy-Binet theorem.

**Theorem 1.3.28** (Cauchy-Binet)**.** *Let $A$ and $B$ be $n \times n$ complex matrices. Suppose $\chi$ is an irreducible character of $S_m$. If $\alpha, \beta \in \Omega_\chi$, then:*

$$d_\chi((AB)[\alpha|\beta]) = \frac{\chi(\mathrm{id})}{m!} \sum_{\gamma \in \Omega_\chi} d_\chi(A[\alpha|\gamma]) d_\chi(B[\gamma|\beta]). \tag{1.12}$$

*Proof.* Suppose $E = \{e_1, e_2, \ldots, e_n\}$ an orthonormal basis of $V$, $S, T \in \mathcal{L}(V)$ and

$$A^T = M(S; E) \qquad B^T = M(T; E).$$

Then

$$\begin{aligned}
\frac{\chi(\mathrm{id})}{m!} d_\chi((AB)[\alpha|\beta]) &= \langle K_\chi(ST)(e^*_\alpha), e^*_\beta\rangle \\
&= \langle K_\chi(S)(e^*_\alpha), K_\chi(T^*)(e^*_\beta)\rangle \\
&= \sum_{\gamma \in \Gamma_{m,n}} \langle K_\chi(S)(e^*_\alpha), e^\otimes_\gamma\rangle \langle e^\otimes_\gamma, K_\chi(T^*)(e^*_\beta)\rangle.
\end{aligned}$$

In the first equality we use Lemma 1.3.27, and in the last equality we use Parseval's Identity.
Since $K_\chi$ is hermitian and idempotent and it commutes with $\otimes^m S$ and $\otimes^m T$, in the last expression we may replace $e^\otimes_\gamma$ by $e^*_\gamma$.
Now we have that $e^*_\gamma = 0$ if $\gamma \notin \Omega_\chi$. So

$$\frac{\chi(\mathrm{id})}{m!} d_\chi((AB)[\alpha|\beta]) = \sum_{\gamma \in \Omega_\chi} \left(\frac{\chi(\mathrm{id})}{m!}\right)^2 d_\chi(A[\alpha|\gamma]) \overline{d_\chi(B^*[\beta|\gamma])},$$

applying Lemma 1.3.27 twice. Notice that

$$B^*[\beta|\gamma] = B[\gamma|\beta]^* \qquad d_\chi(C^*) = \overline{d_\chi(C)},$$

we conclude that

$$d_\chi((AB)[\alpha|\beta]) = \frac{\chi(\mathrm{id})}{m!} \sum_{\gamma \in \Omega_\chi} d_\chi(A[\alpha|\gamma]) d_\chi(B[\gamma|\beta]).$$

□

**Example 1.3.29.** We will prove that $\det(AB) = \det(A)\det(B)$, using Cauchy-Binet theorem. In order to do that, take $m = n$ , $\alpha = \beta = \mathrm{id}$ and $\chi$ the alternating character of $S_n$. Then

$$\begin{aligned} d_{\mathrm{sgn}}(AB[\mathrm{id}|\mathrm{id}]) &= \det(AB) \\ &= \frac{1}{m!}\sum_{\gamma\in\Omega_{\mathrm{sgn}}} \det(A[\mathrm{id}|\gamma])\det(B[\gamma|\mathrm{id}]). \end{aligned}$$

Since $\Omega_{\mathrm{sgn}} = \Gamma_{n,n}$ and $\det(A[\mathrm{id}|\gamma]) = 0$ if $\gamma$ is not injective, we have, from the last equality

$$\det(AB) = \frac{1}{m!}\sum_{\sigma\in S_n} \det(A[\mathrm{id}|\sigma])\det(B[\sigma|\mathrm{id}]). \tag{1.13}$$

Now, we analyse the matrices in each summand. For each $\sigma \in S_n$, $A[\mathrm{id}|\sigma]$ is the matrix whose $(i, j)$-entry is equal to the $(i, \sigma(j))$-entry of $A$ and we also have that the $(i, j)$-entry of $B[\sigma|\mathrm{id}]$ is equal to the $(\sigma(i), j)$-entry of the matrix $B$. So

- $\det(A[\mathrm{id}|\sigma]) = \mathrm{sgn}(\sigma)\det(A)$,
- $\det(B[\sigma|\mathrm{id}]) = \mathrm{sgn}(\sigma)\det(B)$.

Substituting this in the last equation, we get

$$\begin{aligned} \det(AB) &= \frac{1}{m!}\sum_{\sigma\in S_n} \det(A)\det(B) \\ &= \det(A)\det(B). \end{aligned}$$

We have already noticed that the permanent is not a multiplicative immanant, and it is easy to provide a counterexample where this property fails.

**Example 1.3.30.** Let $A = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.
We have

$$\mathrm{per}(A)\,\mathrm{per}(B) = -1 \times 2 \neq -6 = \mathrm{per}(AB).$$

There is, however, a relation between $\operatorname{per}(A)\operatorname{per}(B)$ and $\operatorname{per}(AB)$, which we can deduce using the Cauchy-Binet formula. If we apply this formula to the permanent, we have $\chi \equiv 1$ and

$$\begin{aligned} d_1(AB[\mathrm{id}|\mathrm{id}]) &= \operatorname{per}(AB) \\ &= \frac{1}{m!}\sum_{\gamma\in\Omega_1} \operatorname{per}(A[\mathrm{id}|\gamma])\operatorname{per}(B[\gamma|\mathrm{id}]). \end{aligned}$$

Again it is easy to check that $\Omega_1 = \Gamma_{n,n}$. On the other hand, there are summands that are equal. In fact, suppose $\gamma, \delta \in \Gamma_1$, we have that

$$\operatorname{per}(A[\mathrm{id}|\gamma])\operatorname{per}(B[\gamma|\mathrm{id}]) = \operatorname{per}(A[\mathrm{id}|\delta])\operatorname{per}(B[\delta|\mathrm{id}])$$

if $\operatorname{Im}\gamma = \operatorname{Im}\delta = \{i_1, i_2, \ldots, i_l\}$ and $|\gamma^{-1}(i_k)| = |\gamma^{-1}(i_k)| = m_k$ for every $k = 1, 2, \ldots, l$. So the sum in Cauchy-Binet formula can be indexed by $G_{n,n}$ and for each element $\delta \in G_{n,n}$ there are $\frac{m!}{m_1!m_2!\ldots m_l!}$ elements in $\Gamma_{n,n}$ that have the same value for $\operatorname{per}(A[\mathrm{id}|\gamma])\operatorname{per}(B[\gamma|\mathrm{id}])$. For each element $\delta \in G_{n,n}$ denote $u(\delta) = m_1!m_2!\ldots m_l!$. With this notation we can rewrite the formula for $\operatorname{per}(AB)$,

$$\operatorname{per}(AB) = \sum_{\delta\in G_{n,n}} \frac{1}{u(\delta)}\operatorname{per}(A[\mathrm{id}|\delta])\operatorname{per}(B[\delta|\mathrm{id}]),$$

which is exactly the expression stated in [23].

If $\delta = \mathrm{id}$ then $\frac{1}{u(\mathrm{id})}\operatorname{per}(A[\mathrm{id}|\mathrm{id}])\operatorname{per}(B[\mathrm{id}|\mathrm{id}]) = \operatorname{per}(A)\operatorname{per}(B)$ and so

$$\operatorname{per}(AB) = \operatorname{per}(A)\operatorname{per}(B) + \sum_{\substack{\delta\in G_{n,n}\\ \delta\neq\mathrm{id}}} \frac{1}{u(\delta)}\operatorname{per}(A[\mathrm{id}|\delta])\operatorname{per}(B[\delta|\mathrm{id}]).$$

# Chapter 2

# Higher Order Derivatives

*Analysis takes back with one hand what it gives with the other.*

Charles Hermite

There is a well-known formula for the first directional derivative of the determinant function, due to Jacobi:

$$D\det(A)(X) = \operatorname{tr}(\operatorname{adj}(A)X),$$

where $\operatorname{adj}(A)$ is the adjugate matrix.

When looking for extensions of this formula, one can consider higher order derivatives of the determinant function. This was done by R. Bhatia and T. Jain in [9].

One can also notice that $\det(A) = \wedge^n A$ and look for formulas for derivatives of other Grassmann powers of $A$. T. Jain has presented some formulas in [17].

Yet another possible path is to consider the permanent instead of the determinant, and the symmetric tensor powers instead of the Grassmann powers, and try to establish formulas for their derivatives. This study can be found in a very recent paper by R. Bhatia, P. Grover and T. Jain, [8].

In our work we address the problem of generalizing these formulas for all immanants and all symmetric powers of a matrix or an operator. One of the obstacles in this generalization is that the permanent and the determinant are the only immanants that are associated with linear characters, which have better properties. Another difficulty was finding the expression of the immanant of a direct sum of matrices. Finally, it was necessary to find good bases for the symmetry classes. The results have been collected in [10].

In dealing with symmetric powers we take two approaches: first, we consider powers of matrices and present formulas that depend on the entries of these matrices, then we establish formulas for the symmetric powers of operators and show what is the relation between them.

## 2.1 Differential Calculus

We present some basic concepts and results of differential calculus, more details and proofs can be found in [3]. Like in the previous chapters $V$ is an $n$ dimensional Hilbert space over $\mathbb{C}$.

Let $X$ and $Y$ be real Banach spaces, we write $\mathcal{L}(X, Y)$ to represent the space of all bounded linear operator from $X$ to $Y$. Let $U$ be an open subset of $X$.

**Definition 2.1.1.** Suppose $f : U \longmapsto Y$ is a continuous map. The map $f$ is said to be *differentiable* at a point $u$ of $U$ if there exists a bounded linear operator $T \in \mathcal{L}(X, Y)$ such that

$$\lim_{v \to 0} \frac{\|f(u+v) - f(u) - Tv\|}{\|v\|} = 0.$$

We can easily see that if the operator $T$ exists, it is unique. If the map $f$ is differentiable at $u$, the operator $T$ in the previous definition is called the *derivative* or the *Fréchet derivative* of $f$ at $u$.

**Definition 2.1.2.** Let $\phi : M_n(\mathbb{C}) \longrightarrow M_l(\mathbb{C})$ be a differentiable map and $A, X \in M_n(\mathbb{C})$. The *directional derivative* of $\phi$ at $A$ in the direction $X$ is given by

$$D\phi(A)(X) = \frac{d}{dt}\Big|_{t=0} \phi(A + tX) = \lim_{t \to 0} \frac{\phi(A + tX) - \phi(A)}{t}.$$

If $f$ is differentiable at $u$, then for every $v \in X$ the directional derivative of $f$ at $u$ in the direction $v$ exists. However, the existence of directional derivatives in all directions does not imply differentiability.

If $T$ is a linear operator then $T$ is differentiable at all points, and its derivative is equal to itself, i.e.,

$$DT(A)(X) = T(X),$$

for every $A, X \in M_n(\mathbb{C})$.

**Example 2.1.3.** Let $A, X \in M_n(\mathbb{C})$.

1. Let $f(A) = A^3$, then
$$Df(A)(X) = A^2X + AXA + XA^2.$$

2. Let $f(A) = AA^*$, then
$$Df(A)(X) = AX^* + XA^*.$$

In the next proposition we recall the usual rules of differentiation, which remain valid in this context.

**Proposition 2.1.4.** *Let $f, g : M_n(\mathbb{C}) \longrightarrow M_n(\mathbb{C})$ and $A, X \in M_n(\mathbb{C})$. Suppose $f$ and $g$ are differentiable at all points. Then*

1. $D(f+g)(A)(X) = Df(A)(X) + Dg(A)(X)$,
2. $D(g \circ f)(A)(X) = Dg(f(A))(f(X)).Df(A)(X)$,
3. *If $g$ is a linear map, then* $D(g \circ f)(A)(X) = g(f(X)).Df(A)(X)$,
4. $D(fg)(A)(X) = [Df(A)(X)]\, g(A) + f(A)\, [Dg(A)(X)]\,.$

We can also define higher order derivatives of a differentiable map.

**Definition 2.1.5.** Let $f : M_n(\mathbb{C}) \longrightarrow M_n(\mathbb{C})$ be a differentiable map. Suppose $1 \leq k \leq n$ and $A, X^1, \ldots X^k \in M_n(\mathbb{C})$, the *$k$-th derivative of $f$ at $A$ in the directions of* $(X^1, \ldots, X^k)$ is given by the expression
$$D^k f(A)(X^1, \ldots, X^k) := \frac{\partial^k}{\partial t_1 \ldots \partial t_k}\Big|_{t_1=\ldots=t_k=0} \phi(A + t_1X^1 + \ldots + t_kX^k).$$

Later we will need the following proposition, which generalizes the rules of derivation — recall that the derivative of a linear map $g$ at each point is $g$.

**Proposition 2.1.6.** *If $f, g$ and $h$ are maps such that $f \circ g$ and $g \circ h$ are well defined, with $g$ linear, all of them being $k$ times differentiable, then*
$$D^k(f \circ g)(A)(X^1, \ldots, X^k) = D^k f(g(A))(g(X^1), \ldots, g(X^k))$$
*and*
$$D^k(g \circ h)(A)(X^1, \ldots, X^k) = g \circ D^k h(A)(X^1, \ldots, X^k).$$

## 2.2 First Order Derivative of the Immanant

Let $\det : M_n(\mathbb{C}) \longrightarrow \mathbb{C}$ be the multilinear map taking a $n \times n$ complex matrix to its determinant. By the definition of the directional derivative at a point $A \in M_n(\mathbb{C})$, for each $X \in M_n(\mathbb{C})$,

$$D\det(A)(X) = \frac{d}{dt}\Big|_{t=0} \det(A + tX).$$

The famous *Jacobi formula* states that

$$D\det(A)(X) = \operatorname{tr}(\operatorname{adj}(A)X),$$

where $\operatorname{adj}(A)$ stands for the usual *adjugate* of the matrix $A$.

Our first theorem gives a generalization of Jacobi's formula to all immanants.

**Theorem 2.2.1.** *Let $A$ be an $n \times n$ complex matrix. For each $X \in M_n(\mathbb{C})$,*

$$Dd_\chi(A)(X) = \operatorname{tr}(\operatorname{adj}_\chi(A)^T X).$$

Before we prove the theorem, we prove the following lemma.

**Lemma 2.2.2.** *Let $d_\chi : M_n(\mathbb{C}) \longrightarrow \mathbb{C}$ be the immanant function and let $A, X \in M_n(\mathbb{C})$. Suppose $t$ is a variable. Then $d_\chi(A + tX)$ is a polynomial in $t$ with degree less or equal to $n$, and the first derivative of $d_\chi$ at $A$ in the direction of $X$ is the coefficient of $t$ in the polynomial $d_\chi(A + tX)$.*

*Proof.* By the definition of the immanant we have that

$$d_\chi(A + tX) = d_\chi(A) + a_1 t + a_2 t^2 + \ldots + a_n t^n,$$

where $a_i \in \mathbb{C}$. Then

$$\begin{aligned}
Dd_\chi(A)(X) &= \frac{d}{dt}\Big|_{t=0} d_\chi(A + tX) \\
&= \lim_{t \to 0} \frac{d_\chi(A + tX) - d_\chi(A)}{t} \\
&= \lim_{t \to 0} \frac{d_\chi(A) + a_1 t + a_2 t^2 + \ldots + a_n t^n - d_\chi(A)}{t} \\
&= \lim_{t \to 0} a_1 + a_2 t + \ldots + a_n t^{n-1} \\
&= a_1.
\end{aligned}$$

This concludes our proof. $\square$

Now we can prove Theorem 2.2.1.

*Proof.* For each $1 \leq j \leq n$, let $A(j; X)$ be the matrix obtained from $A$ by replacing the $j$-th column of $A$ by the $j$-th column of $X$ and keeping the rest of the columns unchanged. Then the given equality can be restated as

$$Dd_\chi(A)(X) = \sum_{j=1}^{n} d_\chi(A(j; X)). \tag{2.1}$$

On the other hand, using the previous lemma, we note that $Dd_\chi(A)(X)$ is the coefficient of $t$ in the polynomial $d_\chi(A + tX)$. Let us calculate the coefficient in $t$. By the definition of the immanant, we have that

$$\begin{aligned} d_\chi(A + tX) &= \sum_{\sigma \in S_n} \prod_{i=1}^{n} \chi(\sigma)(a_{i\sigma(i)} + tx_{i\sigma(i)}) \\ &= \sum_{\sigma \in S_n} \prod_{i=1}^{n} \chi(\sigma)(a_{\sigma(i)i} + tx_{\sigma(i)i}), \end{aligned}$$

because $\chi(\sigma) = \chi(\sigma^{-1})$, for every $\sigma \in S_n$. So, the coefficient in $t$ is equal to

$$\begin{aligned} & \sum_{\sigma \in S_n} \chi(\sigma)(x_{\sigma(1)1} a_{\sigma(2)2} \dots a_{\sigma(n)n} + \\ & a_{\sigma(1)1} x_{\sigma(2)2} \dots a_{\sigma(n)n} + \dots + a_{\sigma(1)1} a_{\sigma(2)2} \dots x_{\sigma(n)n}) \\ = & \sum_{\sigma \in S_n} \chi(\sigma) x_{\sigma(1)1} a_{\sigma(2)2} \dots a_{\sigma(n)n} + \\ & \sum_{\sigma \in S_n} \chi(\sigma) a_{\sigma(1)1} x_{\sigma(2)2} \dots a_{\sigma(n)n} + \dots + \sum_{\sigma \in S_n} \chi(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \dots x_{\sigma(n)n} \\ = & \; d_\chi(A(1; X)) + d_\chi(A(2; X)) + \dots + d_\chi(A(n; X)) \\ = & \sum_{j=1}^{n} d_\chi(A(j; X)) \end{aligned}$$

That concludes our proof. $\square$

Now we use the Laplace Expansion for Immanants, in the $j$-th column. This says

$$d_\chi(A) = \sum_{i=1}^{n} a_{ij} d_\chi(A(i|j)).$$

Using this and (2.1), we can re-write the expression for the first derivative of $d_\chi(A)$ as

$$Dd_\chi(A)(X) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij} d_\chi(A(i|j)). \tag{2.2}$$

We will obtain, in the next sections, a few expressions generalizing (2.1) and (2.2) for the derivatives of all orders of the immanant.

## 2.3 First Expression

In this section we follow the techniques used in the papers [8] and [9] by R. Bhatia, T. Jain and P. Grover. In these papers we can find several expressions for the higher order derivatives of the determinant and permanent maps.

Let $V$ and $U$ be Hilbert spaces over the field of the complex numbers. We use the symbol $V^n$ to represent the cartesian product of $n$ copies of $V$, i.e.

$$V^n := V \times \ldots \times V.$$

We recall the definition of the $k$-th derivative.

**Definition 2.3.1.** Let $\phi : V^n \longrightarrow U$ be a multilinear map, and take $1 \leq k \leq n$ and $A, X^1, \ldots X^k \in V^n$. The *$k$-th derivative of $\phi$ at $A$ in directions of* $(X^1, \ldots, X^k)$ is given by the expression

$$D^k\phi(A)(X^1, \ldots, X^k) := \frac{\partial^k}{\partial t_1 \ldots \partial t_k}\Big|_{t_1=\ldots=t_k=0} \phi(A + t_1X^1 + \ldots + t_kX^k).$$

For a fixed $A$, $D^k\phi(A)$ is a multilinear map.

We need the following classical result. Let $n_1, n_2, \ldots n_m$ be $m$ positive integers we define

$$\Gamma(n_1, ..., n_m) = \{\alpha : \{1, 2, \ldots, m\} \longrightarrow \mathbb{N} : \alpha(i) \leq n_i, i = 1, 2, \ldots m\}.$$

**Proposition 2.3.2** (Multilinearity Argument)**.** *Let $V_1, ..., V_m, U$ be vector spaces over $\mathbb{C}$ and $\varphi : V_1 \times ... \times V_m \longrightarrow U$ a multilinear map. Suppose that for each $i \in \{1, ..., m\}$ , $u_i \in V_i$ and $u_i = \sum_{j=1}^{n_i} u_{ij}$. Then*

$$\varphi(u_1, ..., u_m) = \varphi(\sum_{j=1}^{n_1} u_{1j}, ..., \sum_{j=1}^{n_m} u_{mj}) = \sum_{\alpha \in \Gamma(n_1, ..., n_m)} \varphi(u_{1\alpha(1)}, ..., u_{m\alpha(m)}).$$

We use the multilinearity argument to obtain the first expression of the higher order derivatives of the immanant. We need to introduce some new notation. Let

- $\Gamma^0_{n,k} = \{\alpha : \{1, ..., n\} \longrightarrow \{0, 1, ..., k\}\}$.
- $S^0_k = S_{\{0,1,...,k\}}$, i.e. the permutation group of the set $\{0, 1, ..., k\}$.
- $S^{'}_k = \{\sigma \in S^0_k : \sigma(0) = 0\}$.

Given $\alpha$ in $\Gamma^0_{n,k}$ and we denote by $|\alpha^{-1}(i)|$ the number of elements of $\{1, ..., n\}$ whose image is equal to $i$, $0 \le i \le k$.

**Definition 2.3.3.** Let $\alpha \in \Gamma^0_{n,k}$, we define the *support of* $\alpha$ as

$$\operatorname{supp} \alpha = \{i \in \{1, 2, \ldots n\} : \alpha(i) \neq 0\}.$$

In $\Gamma^0_{n,k}$ we consider the elements $\alpha$ that are bijective when restricted to $\operatorname{supp} \alpha$, i.e.

$$\Lambda = \{\alpha \in \Gamma^0_{n,k} : |\alpha^{-1}(0)| = n - k\, , \, |\alpha^{-1}(1)| = |\alpha^{-1}(2)| = \ldots = |\alpha^{-1}(k)| = 1\}.$$

In $\Lambda$ we define the following equivalence relation:

$$\alpha \, \rho \, \beta \qquad \text{if and only if} \qquad \operatorname{supp} \alpha = \operatorname{supp} \beta,$$

meaning that two elements of $\Lambda$ are in the same equivalence class if they have the same support.

Let $X$ be a subset of $\{1, ..., n\}$ we denote by $\overline{X}$ the complement of $X$ in $\{1, ..., n\}$.

**Lemma 2.3.4.** *Let $\alpha, \beta \in \Lambda$. Then $\alpha \, \rho \, \beta$ if and only if there is $\sigma \in S'_k$ such that $\beta = \sigma\alpha$*

*Proof.* Suppose $\beta = \sigma\alpha$ then

$$\beta^{-1}(0) = \alpha^{-1}\sigma^{-1}(0) = \alpha^{-1}(0),$$

by the definition of $S'_k$.

Now, for the converse, let

$$\sigma = \begin{pmatrix} 0 & \alpha(1) & ... & \alpha(k) \\ 0 & \beta(1) & ... & \beta(k) \end{pmatrix} \in S'_k.$$

Then for every $i \in \{1, ..., k\}$ we have that

$$\sigma\alpha(i) = \begin{cases} 0, i \in \alpha^{-1}(0) \\ \beta(i), i \notin \alpha^{-1}(0) \end{cases}$$

So $\sigma\alpha = \beta$. □

For each $\beta$ in $\Lambda$ we will write $\Lambda_\beta$ to denote the equivalence class of $\beta$ with respect to the relation $\rho$.

Define also the set of elements $\alpha$ of $\Lambda$ that are strictly increasing in $\operatorname{supp}\alpha$, i.e.

$$\mathcal{Q}_{n,k} = \{\alpha \in \Lambda : \alpha|_{\operatorname{supp}\alpha} \in Q_{k,n}\}.$$

Then we have the following result.

**Proposition 2.3.5.** *The set $\mathcal{Q}_{k,n}$ is a system of representatives of the equivalence classes of $\dfrac{\Lambda}{\rho}$, i.e.*

$$\Lambda = \{\Lambda_\beta : \beta \in \mathcal{Q}_{n,k}\}.$$

*Proof.* We will start by proving that if $\alpha, \beta \in \mathcal{Q}_{n,k}$ and $\alpha \, \rho \, \beta$ then $\alpha = \beta$. We have seen that if $\alpha \, \rho \, \beta$ then there is $\sigma \in S'_k$ such that $\alpha = \sigma\beta$. Let $\gamma = (\alpha|_{\operatorname{supp}\alpha})^{-1}$ and $\theta = (\beta|_{\operatorname{supp}\beta})^{-1}$, with $\alpha \, \rho \, \beta$. It follows from the definition of the set $\Lambda$ that $\gamma$ and $\theta$ are elements of $Q_{k,n}$.

$$\gamma = (\alpha|_{\operatorname{supp}\alpha})^{-1} = (\beta|_{\operatorname{supp}\alpha})^{-1}\sigma^{-1} = (\beta|_{\operatorname{supp}\beta})^{-1}\sigma^{-1} = \theta\sigma^{-1}$$

So $\sigma = \text{id}$ and then $\alpha = \beta$.
Now we have to prove that every class has an element of $\mathcal{Q}_{n,k}$.
If $\omega \in \Lambda$ then let $\overline{\omega^{-1}(0)} = \{i_1, ..., i_k\}, i_1 < ... < i_k$.

Suppose $\nu \in \Gamma^0_{n,k}$ such that $\nu^{-1}(0) = \omega^{-1}(0)$ and $\nu(i_1) = 1, ..., \nu(i_k) = k$, notice that in particular $\nu \in \mathcal{Q}_{n,k}$.

Let $\sigma \in S'_k$, defined as follows.
We have that

$$\sigma\nu(i_1) = \omega(i_1), ..., \sigma\nu(i_k) = \omega(i_k).$$

Then $\sigma\nu = \omega$. Now we can conclude that every element of the set $\Lambda$ is $\rho$ equivalent to an element of the set $\mathcal{Q}_{n,k}$ and that element is unique.
That concludes our proof. □

It is a direct consequence of the previous proposition that it is possible to define a map

$$\begin{aligned} Q_{k,n} &\longrightarrow \mathcal{Q}_{n,k} \\ \alpha &\longrightarrow \lambda_\alpha, \end{aligned}$$

where

$$\lambda_\alpha(i) = \begin{cases} \alpha^{-1}(i), & i \in \operatorname{Im}\alpha \\ 0, & \text{otherwise} \end{cases} .$$

In other words, this means that for every $\gamma \in \Lambda$ there are unique $\sigma \in S'_k$ and $\alpha \in Q_{k,n}$ such that:

$$\gamma = \sigma\lambda_\alpha. \tag{2.3}$$

**Example 2.3.6.** Suppose $k = 3$ and $n = 6$. Let $\alpha = (3, 5, 6) \in Q_{3,6}$. Then

$$\lambda_\alpha : \{1, 2, 3, 4, 5, 6\} \longrightarrow \{0, 1, 2, 3\}$$

$$\lambda_\alpha = (0, 0, 1, 0, 2, 3),$$

that belongs to $\mathcal{Q}_{6,3}$.

We now obtain a generalization of the expression (3.1) for higher order derivatives. First we demonstrate a more general result for multilinear maps using the multilinearity argument. For $A, X^1, \ldots, X^k \in V^n$, we can write

$A = (A_1, \ldots, A_n)$ and $X^i = (X_1^i, \ldots, X_n^i)$ with $A_j, X_j^i \in V$, for every $1 \leq i \leq k$ and $1 \leq j \leq n$. Define

$$U_i = A_i + t_1 X_i^1 + ... + t_k X_i^k.$$

$$u_{ij} = \begin{cases} t_0 A_i, \ j = 0 \\ t_j X_i^j, \ j = 1, ..., k \end{cases},$$

where $t_0 = 1$ for a matter of convention. We also define

$$U_i' = A_i + X_i^1 + ... + X_i^k.$$

$$u_{ij}' = \begin{cases} A_i, \ j = 0 \\ X_i^j, \ j = 1, ..., k \end{cases}.$$

We start by proving the following result.

**Lemma 2.3.7.** *Let $1 \leq k \leq n$ and let $\phi : V^n \longrightarrow U$ be a multilinear map. Suppose $A, X^1, \ldots, X^k \in V^n$ and let $t_1, \ldots, t_k$, be $k$ complex variables. Then*

$$\phi(A + t_1 X^1 + ... + t_k X^k) = \sum_{\gamma \in \Gamma_{n,k}^0} t_1^{|\gamma^{-1}(1)|} ... t_k^{|\gamma^{-1}(k)|} \phi(u'_{1\gamma(1)}, ..., u'_{n\gamma(n)}). \quad (2.4)$$

*Proof.*

$$\begin{aligned} \phi(A + t_1 X^1 + ... + t_k X^k) &= \phi(U_1, ..., U_n) \\ &= \sum_{\gamma \in \Gamma_{n,k}^0} \phi(u_{1\gamma(1)}, ..., u_{n\gamma(n)}) \\ &= \sum_{\gamma \in \Gamma_{n,k}^0} \phi(t_{\gamma(1)} u'_{1\gamma(1)}, ..., t_{\gamma(n)} u'_{n\gamma(n)}) \\ &= \sum_{\gamma \in \Gamma_{n,k}^0} t_1^{|\gamma^{-1}(1)|} ... t_k^{|\gamma^{-1}(k)|} \phi(u'_{1\gamma(1)}, ..., u'_{n\gamma(n)}) \end{aligned}$$

For the second equality we have used the multilinearity argument. For the last equality we use th fact that $\phi$ is a multilinear map. $\square$

Now we can prove a proposition that gives a result for the higher order directional derivatives of a multilinear map.

**Proposition 2.3.8.** *Let* $1 \leq k \leq n$. *Let* $\phi : V^n \longrightarrow U$ *be a multilinear map. Suppose* $A, X^1, \ldots, X^k \in V^n$ *and let* $t_1, \ldots, t_k$, *be* $k$ *complex variables. Then* $D^k\phi(A)(X^1, \ldots, X^k)$ *is the coefficient of* $t_1 \ldots t_k$ *in the polynomial* $\phi(A + t_1X^1 + \ldots + X^k)$.

*Proof.* By the definition of the $k$-th derivative, we have that

$$D^k\phi(A)(X^1, ..., X^k) = \frac{\partial^k}{\partial t_1 ... \partial t_k}\Big|_{t_1=...=t_k=0} \phi(A + t_1X^1 + ... + t_kX^k).$$

Using the previous lemma, we know that

$$\phi(A + t_1X^1 + ... + t_kX^k) = \sum_{\gamma \in \Gamma^0_{n,k}} t_1^{|\gamma^{-1}(1)|} \ldots t_k^{|\gamma^{-1}(k)|} \phi(u'_{1\gamma(1)}, \ldots, u'_{n\gamma(n)}).$$

This is a polynomial in the variables $t_1, \ldots, t_k$ and its derivative is

$$\frac{\partial^k}{\partial t_1 \ldots \partial t_k} \sum_{\gamma \in \Gamma^0_{n,k}} t_1^{|\gamma^{-1}(1)|} \ldots t_k^{|\gamma^{-1}(k)|} \phi(u'_{1\gamma(1)}, \ldots, u'_{n\gamma(n)}) =$$

$$\sum_{\gamma \in \overline{\Gamma}} t_1^{|\gamma^{-1}(1)|-1} \ldots t_k^{|\gamma^{-1}(k)|-1} \phi(u'_{1\gamma(1)}, \ldots, u'_{n\gamma(n)}),$$

where $\overline{\Gamma} = \{\gamma \in \Gamma^0_{n,k} : |\gamma^{-1}(i)| \geq 1, 1 \leq i \leq k\}$.

The last sum is indexed by $\overline{\Gamma}$ because the summands that are indexed by $\Gamma^0_{n,k} \setminus \overline{\Gamma}$ are equal to zero.

Now, taking $t_1 = t_2 = \ldots = t_k = 0$, the only nonzero summands are the ones in which

$$|\gamma^{-1}(1)| = |\gamma^{-1}(2)| = \ldots = |\gamma^{-1}(k)| = 1,$$

for $\gamma \in \overline{\Gamma}$.

This is exactly the coefficient of $t_1 \ldots t_k$ in $\phi(A + t_1X^1 + \ldots + t_kX^k)$. $\square$

We have proved that the higher order derivatives of a multilinear map $\phi$ are certain coefficients of the polynomial $\phi(A + t_1X^1 + \ldots + t_kX^k)$. In the next theorem we give the explicit expression for those coefficients. We first introduce some notation.

**Definition 2.3.9.** Let $\alpha \in Q_{k,n}$. We denote by $A(\alpha; X^1, \ldots, X^k)$ the element of $V^n$ obtained from the element $A$ by replacing coordinate $\alpha(j)$ of $A$ by coordinate $\alpha(j)$ of $X^j$ for every $1 \leq j \leq k$.

**Theorem 2.3.10.** *Let $1 \leq k \leq n$ and let $\phi : V^n \longrightarrow U$ be a multilinear map and $A, X^1, \ldots, X^k \in V^n$. Then*

$$D^k\phi(A)(X^1, \ldots, X^k) = \sum_{\sigma \in S_k} \sum_{\alpha \in Q_{k,n}} \phi\left(A(\alpha; X^{\sigma(1)}, \ldots, X^{\sigma(k)})\right).$$

*Proof.* By Proposition 2.3.8, $D^k\phi(A)(X^1, ..., X^k)$ is the coefficient of $t_1 \ldots t_k$ in the polynomial

$$\sum_{\gamma \in \overline{\Gamma}} t_1^{|\gamma^{-1}(1)|-1} \ldots t_k^{|\gamma^{-1}(k)|-1} \phi(u'_{1\gamma(1)}, \ldots, u'_{n\gamma(n)}),$$

where $\overline{\Gamma} = \{\gamma \in \Gamma^0_{n,k} : |\gamma^{-1}(i)| \geq 1, 1 \leq i \leq k\}$ and

$$|\gamma^{-1}(1)| = |\gamma^{-1}(2)| = \ldots = |\gamma^{-1}(k)| = 1,$$

for $\gamma \in \overline{\Gamma}$. This also means that $\gamma$ is in $\Lambda$. Therefore,

$$D^k\phi(A)(X^1, ..., X^k) = \sum_{\gamma \in \Lambda} \phi(u'_{1\gamma(1)}, ..., u'_{n\gamma(n)}).$$

We have also seen in (2.3) that for every $\gamma \in \Lambda$ there are unique $\sigma \in S'_k$ and $\alpha \in Q_{k,n}$ such that $\gamma = \sigma\lambda_\alpha$, so we have

$$D^k\phi(A)(X^1, ..., X^k) = \sum_{\alpha \in Q_{k,n}} \sum_{\sigma \in S'_k} \phi(u'_{1\sigma\lambda_\alpha(1)}, ..., u'_{n\sigma\lambda_\alpha(n)}).$$

Fixing a coordinate $t$, $1 \leq t \leq n$, we analyse $u'_{t\sigma\lambda_\alpha(t)}$. If $t \notin \operatorname{Im}\alpha$, then $\lambda_\alpha(t) = 0$ and we have that

$$u'_{t\sigma(0)} = u'_{t0} = A_t.$$

If $t \in \operatorname{Im}\alpha$, then $\lambda_\alpha(t) = \alpha^{-1}(t)$ and we have that

$$u'_{t\sigma(\alpha^{-1}(t))} = X_t^{\sigma(\alpha^{-1}(t))}.$$

In this case $t \in \operatorname{Im}\alpha$, so there is an $s$ such that $t = \alpha(s)$ and we have $X_t^{\sigma(\alpha^{-1}(t))} = X_{\alpha(s)}^{\sigma(s)}$. Therefore, for every $\alpha \in Q_{k,n}$ and for every $\sigma \in S_k$,

$$\phi(u'_{1\sigma\lambda_\alpha(1)}, ..., u'_{n\sigma\lambda_\alpha(n)}) = \phi\left(A(\alpha; X^{\sigma(1)}, ..., X^{\sigma(k)})\right).$$

So we have

$$D^k\phi(A)(X^1, \ldots, X^k) = \sum_{\sigma \in S_k} \sum_{\alpha \in Q_{k,n}} \phi\left(A(\alpha; X^{\sigma(1)}, \ldots, X^{\sigma(k)})\right).$$

This concludes our proof. $\square$

Given a matrix $A \in M_n(\mathbb{C})$, we represent by $A_{[i]}$ the $i$-th column of $A$, $i \in \{1, \ldots, n\}$. In particular, we have $A = (A_{[1]}, \ldots, A_{[n]})$.

Now we consider each matrix to be a list of $n$ columns, so that if $V = \mathbb{C}^n$, $V^n = M_n(\mathbb{C})$. With this identification, we have that $A(\alpha; X^1, \ldots, X^k)$ is the matrix of order $n$ obtained from $A$ replacing the $\alpha(j)$ column of $A$ by the $\alpha(j)$ column of $X^j$.

**Example 2.3.11.** Suppose $n = 5$, $A = I_5$, $(X^1)_{ij} = -1$, for every $i, j = 1, \ldots, 5$ and $(X^2)_{ij} = i + j$, for every $i, j = 1, \ldots, 5$. Let $\alpha = (35)$. Then

$$A(\alpha; X^1, X^2) = \begin{pmatrix} 1 & 0 & -1 & 0 & 6 \\ 0 & 1 & -1 & 0 & 7 \\ 0 & 0 & -1 & 0 & 8 \\ 0 & 0 & -1 & 1 & 9 \\ 0 & 0 & -1 & 0 & 10 \end{pmatrix}.$$

The following theorem states the main result of this section.

**Theorem 2.3.12** (First expression)**.** *For every $1 \leq k \leq n$, let $A, X^1, \ldots, X^k$ be $n \times n$ complex matrices. Then*

$$D^k d_\chi(A)(X^1, \ldots, X^k) = \sum_{\sigma \in S_k} \sum_{\alpha \in Q_{k,n}} d_\chi A(\alpha; X^{\sigma(1)}, \ldots, X^{\sigma(k)}).$$

*In particular,*

$$D^k d_\chi(A)(X, \ldots, X) = k! \sum_{\alpha \in Q_{k,n}} d_\chi A(\alpha; X, \ldots, X).$$

*Proof.* We only have to take $\phi = d_\chi$ and $V^n = M_n(\mathbb{C})$.

If $X^1 = \ldots = X^k = X$, then $d_\chi A(\alpha; X^{\sigma(1)}, \ldots, X^{\sigma(k)}) = d_\chi A(\alpha; X, \ldots, X)$ for every $\sigma \in S_k$. So, in this case

$$D^k d_\chi(A)(X, \ldots, X) = k! \sum_{\alpha \in Q_{k,n}} d_\chi A(\alpha; X, \ldots, X).$$

□

We can re-write the last expression for the $k$-th derivative of the immanant map using the concept of mixed immanant, generalizing the respective concepts for the determinant and the permanent.

**Definition 2.3.13.** Let $X^1, \ldots, X^n$ be $n$ matrices of order $n$. We define the *mixed immanant* of $X^1, \ldots, X^n$ as

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

If $X^1 = \ldots = X^t = A$, for some $t \leq n$ and $A \in M_n(\mathbb{C})$, we denote the mixed immanant by $\Delta_\chi(A; X^{t+1}, \ldots, X^n)$.

As with the permanent and the determinant, we have that

$$\Delta_\chi(A, \ldots, A) = d_\chi(A).$$

**Proposition 2.3.14.** *Let $A \in M_n(\mathbb{C})$. Then*

$$\Delta_\chi(A; X^1, \ldots, X^k) := \frac{(n-k)!}{n!} \sum_{\sigma \in S_k} \sum_{\alpha \in Q_{k,n}} d_\chi A(\alpha; X^{\sigma(1)}, \ldots, X^{\sigma(k)}).$$

*Proof.* One simply has to observe that each summand in $\Delta_\chi(A; X^1, \ldots, X^k)$ appears $(n-k)!$ times: once we fix a permutation of the matrices $X^1, \ldots, X^k$, these summands correspond to the possible permutations of the $n-k$ matrices equal to $A$. □

As an immediate consequence of this result, we can obtain another formula for the derivative of order $k$ of the immanant map. This generalizes formula (26) in [9].

**Proposition 2.3.15** (First expression, rewritten)**.** *For every $1 \leq k \leq n$, let $A, X^1, \ldots, X^k$ be $n \times n$ complex matrices. Then*

$$D^k d_\chi(A)(X^1, \ldots, X^k) = \frac{n!}{(n-k)!} \Delta_\chi(A; X^1, \ldots, X^k). \tag{2.5}$$

## 2.4 Second Expression

In order to generalize the second expression of the $k$-th derivative of the determinant and the permanent, we wish to separate in the expression of $Dd_\chi(A)(X^1, \ldots, X^k)$ the entries of the matrices $X^1, \ldots, X^k$ from the entries of $A$. In the two cases that have already been studied, the determinant and the permanent maps, this was easier, because in both cases there are formulas that allow us to express the determinant of a direct sum in terms of determinants of direct summands, and the same happens with the permanent. With other immanants, the best we can do is use formula (1.9), which is what we do in this second expression.

**Definition 2.4.1.** Let $1 \leq k \leq n$, $X^1, \ldots, X^k$ complex matrices of order $n$. Suppose $\sigma \in S_k$, and $\beta \in Q_{k,n}$. Denoting by $\mathbf{0}$ the zero matrix of order $n$, we define

$$X_\beta^\sigma = \mathbf{0}(\beta; X^{\sigma(1)}, \ldots, X^{\sigma(k)}),$$

the matrix whose $\beta(p)$-th column is equal to $X_{[\beta(p)]}^{\sigma(p)}$ and the remaining columns are zero, for $1 \leq p \leq k$.

**Example 2.4.2.** Suppose $n = 4$ , $k = 2$, $\beta = (2, 4)$,

$$X^1 = I_4 \; X^2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}.$$

We have that $S_2 = \{\mathrm{id}, (12)\}$. Then

$$X_{(2,4)}^{\mathrm{id}} = \begin{pmatrix} 0 & 0 & 0 & 4 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

and

$$X_{(2,4)}^{(12)} = \begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 2 & 0 & 1 \end{pmatrix}.$$

**Theorem 2.4.3** (Second Formula)**.** *For every* $1 \leq k \leq n$*, let* $A, X^1, \ldots, X^k$ *be* $n \times n$ *complex matrices. Then*

$$D^k d_\chi A(X^1, \ldots, X^k) = \sum_{\sigma \in S_k} \sum_{\alpha, \beta \in Q_{k,n}} d_\chi(X^\sigma_\beta[\alpha|\beta] \bigoplus_{\alpha|\beta} A(\alpha|\beta)),$$

*in particular*

$$D^k d_\chi A(X, \ldots, X) = k! \sum_{\alpha, \beta \in Q_{k,n}} d_\chi(X[\alpha|\beta] \bigoplus_{\alpha|\beta} A(\alpha|\beta))$$

*Proof.* We have proved that

$$D^k d_\chi A(X^1, \ldots, X^k) = \sum_{\sigma \in S_k} \sum_{\beta \in Q_{k,n}} d_\chi A(\beta; X^{\sigma(1)}, \ldots, X^{\sigma(k)}).$$

By the Laplace expansion for immanants, for every $\beta \in Q_{k,n}$, we have that

$$\begin{aligned} & d_\chi A(\beta; X^{\sigma(1)}, \ldots, X^{\sigma(k)}) = \\ & = \sum_{\alpha \in Q_{k,n}} d_\chi(A(\beta; X^{\sigma(1)}, \ldots, X^{\sigma(k)})\{\alpha|\beta\}) \\ & = \sum_{\alpha \in Q_{k,n}} d_\chi(A(\beta; X^{\sigma(1)}, \ldots, X^{\sigma(k)})[\alpha|\beta] \bigoplus_{\alpha|\beta} A(\beta; X^{\sigma(1)}, \ldots, X^{\sigma(k)})(\alpha|\beta)). \end{aligned}$$

Now we just notice that

$$A(\beta; X^{\sigma(1)}, \ldots, X^{\sigma(k)})[\alpha|\beta] = X^\sigma_\beta[\alpha|\beta]$$

and

$$A(\beta; X^{\sigma(1)}, \ldots, X^{\sigma(k)})(\alpha|\beta) = A(\alpha|\beta).$$

This concludes the proof of the formula. □

## 2.5 Formulas for the $k$-th Derivative of $K_\chi(A)$

The main goal of this chapter is to generalize higher order derivative formulas for the antisymmetric and symmetric tensor powers obtained by R. Bhatia, T. Jain and P. Grover. In the previous section we have already generalized the formulas for the determinant and the permanent functions to all immanants. Now, we intend to calculate formulas that generalize the ones that have been calculated for the $m$-th induced power of a $n \times n$ matrix $A$, which is usually represented by $\vee^m A$ and for the $m$-th compound of $A$, $\wedge^m A$, these are also called the permanental compound and the determinantal compound of $A$, respectively. Before we can do this, we need quite a bit of definitions, including the very definition of this matrix, which is a little bit more complicated than in the particular cases that have already been studied.

Let $E = \{e_1, e_2, \ldots, e_n\}$ be an orthonormal basis of the Hilbert space $V$. Then

$$\mathcal{E}' := \{e^*_\alpha : \alpha \in \widehat{\Delta}\}, \tag{2.6}$$

is basis of $V_\chi$, induced by the basis $E$. It is also known that this basis is orthogonal if $\chi$ is a linear character.

In general, if we consider the induced inner product on $\otimes^m V$, if $\chi$ does not have degree one, there are no known orthonormal bases of $V_\chi$ formed by decomposable symmetrized tensors. Let

$$\mathcal{E} = \{v_\alpha : \alpha \in \widehat{\Delta}\}$$

be the orthonormal basis of the $m$-th $\chi$-symmetric tensor power of the vector space $V$ obtained by applying the Gram-Schmidt orthonormalization procedure to $\mathcal{E}'$. Let $B$ be the $t \times t$ change of basis matrix, from $\mathcal{E}$ to $\mathcal{E}' = \{e^*_\alpha : \alpha \in \widehat{\Delta}\}$, where $t = \dim(V_\chi)$. This means that for each $\alpha \in \widehat{\Delta}$,

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

We note that this matrix $B$ *does not depend on the choice of the orthonormal basis of* $V$, since the set $\widehat{\Delta}$ is independent of the vectors of $E$, and has a natural order (the lexicographic order), which the basis $\mathcal{E}$ inherits. Moreover, the Gram-Schmidt process only depends on the numbers $\langle e^*_\alpha, e^*_\beta \rangle$ and these are given by formula

$$\langle e^*_\alpha, e^*_\beta \rangle = \frac{\chi(\mathrm{id})}{m!} \sum_{\sigma \in S_m} \chi(\sigma) \prod_{t=1}^{m} \langle e_{\alpha(t)}, e_{\beta\sigma(t)} \rangle.$$

Hence, they only depend on the values of $\langle e_i, e_j \rangle = \delta_{ij}$ and are thus independent of the vectors themselves.

Now we want to define $K_\chi(A)$, the $m$-th $\chi$-symmetric tensor power of the matrix $A$.

**Definition 2.5.1.** Let $E = \{e_1, e_2, \ldots, e_n\}$ be an orthonormal basis of the Hilbert space $V$ and $A$ an $n \times n$ complex matrix. Consider the linear endomorphism $T$ such that $A = M(T, E)$. Let $1 \leq m \leq n$ and $\chi$ an irreducible character of $S_m$. Suppose $\mathcal{E}$ is the orthonormal basis of $V_\chi$ applying the Gram-Schmidt orthonormalization to the induced basis. We define the *$m$-th $\chi$-symmetric tensor power of the matrix* $A$ as the $t \times t$ complex matrix

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

where $t = |\widehat{\Delta}|$ and $|Q_{m,n}| \leq t$.

We now notice that *this matrix does not depend on the choice of the orthonormal basis $E$ of $V$*. This is an immediate consequence of the formula in lemma (1.3.27).

**Proposition 2.5.2.** *Suppose $\alpha, \beta \in \widehat{\Delta}$, the $(\alpha, \beta)$ entry of $K_\chi(A)$ is*

$$\frac{\chi(\mathrm{id})}{m!} \sum_{\gamma,\delta \in \widehat{\Delta}} b_{\gamma\beta}\overline{b_{\delta\alpha}} d_\chi(A[\delta|\gamma]).$$

*Proof.* Since the basis $\mathcal{E}$ is orthonormal, the $(\alpha, \beta)$-entry of $K_\chi(A)$ is given by:

$$\begin{aligned}
\langle K_\chi(T)v_\beta, v_\alpha\rangle &= \sum_{\gamma,\delta \in \widehat{\Delta}} \langle b_{\gamma\beta} K_\chi(T) e^*_\gamma, b_{\delta\alpha} e^*_\delta\rangle \\
&= \sum_{\gamma,\delta \in \widehat{\Delta}} b_{\gamma\beta}\overline{b_{\delta\alpha}} \langle K_\chi(T) e^*_\gamma, e^*_\delta\rangle \\
&= \frac{\chi(\mathrm{id})}{m!} \sum_{\gamma,\delta \in \widehat{\Delta}} b_{\gamma\beta}\overline{b_{\delta\alpha}} d_\chi(A^T[\gamma|\delta]) \\
&= \frac{\chi(\mathrm{id})}{m!} \sum_{\gamma,\delta \in \widehat{\Delta}} b_{\gamma\beta}\overline{b_{\delta\alpha}} d_\chi(A[\delta|\gamma]^T) \\
&= \frac{\chi(\mathrm{id})}{m!} \sum_{\gamma,\delta \in \widehat{\Delta}} b_{\gamma\beta}\overline{b_{\delta\alpha}} d_\chi(A[\delta|\gamma]).
\end{aligned}$$

In the second equality we use the properties of the induced inner product, in the third, Lemma 1.3.27, in the fourth the fact that $d_\chi(A^T[\gamma|\delta]) = d_\chi(A[\delta|\gamma]^T)$ and in the last equality we use the fact that $d_\chi(X) = d_\chi(X^T)$. This concludes our proof. $\square$

This definition admits, as special cases, the $m$-th compound and the $m$-th induced power of a matrix, as defined in [25, p. 236]. The matrix $K_\chi(A)$ is called the *induced matrix* in [25, p. 235], in the case when the character has degree one.

**Definition 2.5.3.** Let $A$ be an $n \times n$ complex matrix and $\chi$ an irreducible character of $S_m$. We denote by $\operatorname{imm}_\chi(A)$ the $t \times t$ matrix for $t = |\widehat{\Delta}|$, with rows and columns indexed by $\widehat{\Delta}$, whose $(\gamma, \delta)$ entry is

$$d_\chi(A[\gamma|\delta]).$$

We call the elements of this matrix *immanantal minors* indexed by $\widehat{\Delta}$. The usual minors are obtained by considering the alternating character, in which case $\widehat{\Delta} = Q_{m,n}$. With this definition, we can rewrite the previous equation as

$$K_\chi(A) = \frac{\chi(\mathrm{id})}{m!} B^* \operatorname{imm}_\chi(A) B. \tag{2.7}$$

Finally, denote by $\operatorname{miximm}_\chi(X^1, \ldots, X^n)$ the $t \times t$ complex matrix with rows and columns indexed by $\widehat{\Delta}$, whose $(\gamma, \delta)$ entry is $\Delta_\chi(X^1[\gamma|\delta], \ldots, X^n[\gamma|\delta])$, so that $\operatorname{miximm}_\chi(A, \ldots, A) = \operatorname{imm}_\chi(A)$. We use the same shorthand as with the mixed immanant: for $k \le n$,

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

We will present two formulas for the higher order derivatives of $K_\chi(A)$. The first formula is written as a matrix equality and the second formula is an expression where we split the entries of $A$ from the entries of $X^1, X^2, \ldots, X^k$. Before our main formulas, we recall a general result about derivatives, which we have stated in section 1.

**Lemma 2.5.4.** *If $f$ and $g$ are two maps such that $f \circ g$ is well defined, with $g$ linear, then*

$$D^k(f \circ g)(A)(X^1, \ldots, X^k) = D^k f(g(A))(g(X^1), \ldots, g(X^k)).$$

**Theorem 2.5.5.** *According to our previous notation, we have*

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

*and, using the notation we have already established, the* $(\alpha, \beta)$ *entry of this matrix is*

$$\frac{\chi(\mathrm{id})}{(m-k)!} \sum_{\gamma,\delta\in\widehat{\Delta}} b_{\gamma\beta}\overline{b_{\delta\alpha}}\Delta_\chi(A[\delta|\gamma]; X^1[\delta|\gamma], \dots, X^k[\delta|\gamma])$$

*Proof.* Notice that the map $A \mapsto A[\delta|\gamma]$ is linear, so we can apply the previous lemma in order to compute the derivatives of the entries of the matrix $K_\chi(A)$. The $(\alpha, \beta)$ entry of the $k$-th derivative of the $m$-th $\chi$-symmetric tensor power of $A$, i.e., the $(\alpha, \beta)$ entry of the matrix $D^k K_\chi(A)(X^1, \dots, X^k)$ is:

$$\frac{\chi(\mathrm{id})}{k!} \sum_{\gamma,\delta\in\widehat{\Delta}} b_{\gamma\beta}\overline{b_{\delta\alpha}} D^k d_\chi(A[\delta|\gamma])(X^1[\delta|\gamma], \dots, X^k[\delta|\gamma]).$$

To abbreviate notation, for fixed $\gamma, \delta \in \widehat{\Delta}$, we will write $C := A[\delta|\gamma]$, and $Z^i := X^i[\delta|\gamma]$, $i = 1, \dots, k$. Using formula (2.5), we get

$$\begin{aligned} D^k d_\chi(A[\delta|\gamma])(X^1[\delta|\gamma], \dots, X^k[\delta|\gamma]) &= D^k d_\chi(C)(Z^1, \dots, Z^k) \\ &= \frac{m!}{(m-k)!}\Delta_\chi(C; Z^1, \dots, Z^k). \end{aligned}$$

So the $(\alpha, \beta)$ entry of $D^k K_\chi(A)(X^1, \dots, X^k)$ is

$$\frac{\chi(\mathrm{id})}{m!} \sum_{\gamma,\delta\in\widehat{\Delta}} b_{\gamma\beta}\overline{b_{\delta\alpha}} \frac{m!}{(m-k)!}\Delta_\chi(C; Z^1, \dots, Z^k) =$$

$$\frac{\chi(\mathrm{id})}{(m-k)!} \sum_{\gamma,\delta\in\widehat{\Delta}} b_{\gamma\beta}\overline{b_{\delta\alpha}}\Delta_\chi(A[\delta|\gamma]; X^1[\delta|\gamma], \dots, X^k[\delta|\gamma])$$

According to the definition of $\mathrm{miximm}_\chi(A; X^1, \dots, X^k)$, we have

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

This concludes our proof. $\square$

**Corollary 2.5.6.** *According to our previous notation, we have that* $D^k K_\chi(A)(X^1, \dots, X^k)$ *is equal to*

$$\frac{\chi(\mathrm{id})}{m!} \sum_{\gamma,\delta\in\widehat{\Delta}} b_{\gamma\beta}\overline{b_{\delta\alpha}} \sum_{\sigma\in S_k} \sum_{\rho,\tau\in Q_{k,m}} d_\chi(X[\delta|\gamma]^\sigma_\tau[\rho|\tau] \bigoplus_{\rho|\tau} A[\delta|\gamma](\rho|\tau)).$$

*Proof.* Again,to abbreviate notation, for fixed $\gamma, \delta \in \widehat{\Delta}$, we will write $C := A[\delta|\gamma]$, and $Z^i := X^i[\delta|\gamma]$, $i = 1, \ldots, k$.
Using the formula in Theorem 2.4.3, we have that

$$D^k d_\chi(C)(Z^1, \ldots, Z^k) = \sum_{\sigma \in S_k} \sum_{\rho,\tau \in Q_{k,m}} d_\chi(Z^\sigma_\tau[\rho|\tau] \bigoplus_{\rho|\tau} C(\rho|\tau)).$$

Recall that

$$Z^\sigma_\tau = \mathbf{0}(\tau; Z^{\sigma(1)}, \ldots, Z^{\sigma(k)}),$$

where $\mathbf{0}$ denotes the zero matrix of order $m$.

So, the $(\alpha, \beta)$ entry of the $k$-th derivative of $K_\chi(A)$ is:

$$\frac{\chi(\mathrm{id})}{m!} \sum_{\gamma,\delta \in \widehat{\Delta}} b_{\gamma\beta} \overline{b_{\delta\alpha}} \sum_{\sigma \in S_k} \sum_{\rho,\tau \in Q_{k,m}} d_\chi(Z^\sigma_\tau[\rho|\tau] \bigoplus_{\rho|\tau} C(\rho|\tau)) =$$

$$\frac{\chi(\mathrm{id})}{m!} \sum_{\gamma,\delta \in \widehat{\Delta}} b_{\gamma\beta} \overline{b_{\delta\alpha}} \sum_{\sigma \in S_k} \sum_{\rho,\tau \in Q_{k,m}} d_\chi(X[\delta|\gamma]^\sigma_\tau[\rho|\tau] \bigoplus_{\rho|\tau} A[\delta|\gamma](\rho|\tau)).$$

This concludes our proof. □

The formula obtained for the higher order derivatives of $K_\chi(A)(X^1, \ldots, X^k)$ generalizes the expressions obtained by Bhatia, Jain and Grover ([9], [15]). We will demonstrate this for the derivative of the $m$-th compound, establishing that, from the formula in Theorem 2.5.5, we can establish formula (2.5) in [17], from which the main formula for the derivative of the $m$-th compound of $A$ is obtained.

Let $\chi = \mathrm{sgn}$. Then

$$K_\chi(A)(X^1, \ldots, X^k) = \wedge^m(A)(X^1, \ldots, X^k).$$

In this case $\widehat{\Delta} = Q_{m,n}$ and the basis $\{e^\wedge_\alpha : \alpha \in Q_{m,n}\}$ is orthogonal and it is easy to see (by direct computation or using formula (1.2)) that every vector has norm $1/\sqrt{m!}$. So the matrix $B$ of order $\binom{n}{m}$ is diagonal and its diagonal entries are equal to $\sqrt{m!}$.

We now notice two properties that we will use in our computations:

I. For any matrices $X \in M_k(\mathbb{C})$, $Y \in M_{n-k}(\mathbb{C})$ and functions $\alpha, \beta \in Q_{k,n}$,

$$\det X \bigoplus_{\alpha|\beta} Y = (-1)^{|\alpha|+|\beta|} \det X \det Y.$$

This is a consequence of formula (1.11). We again notice that if $\gamma \neq \beta$, the matrices

$$(X \bigoplus_{\alpha|\beta} Y)[\alpha|\gamma] \text{ and } (X \bigoplus_{\alpha|\beta} Y)(\alpha|\gamma)$$

have a zero column. Now, using the Laplace expansion for the determinant along $\alpha$,

$$\begin{aligned} \det X \bigoplus_{\alpha|\beta} Y &= (-1)^{|\alpha|} \sum_{\gamma \in Q_{k,n}} (-1)^{|\gamma|} \det((X \bigoplus_{\alpha|\beta} Y)[\alpha|\gamma]) \det((X \bigoplus_{\alpha|\beta} Y)(\alpha|\gamma)) \\ &= (-1)^{|\alpha|+|\beta|} \det((X \bigoplus_{\alpha|\beta} Y)[\alpha|\beta]) \det((X \bigoplus_{\alpha|\beta} Y)(\alpha|\beta)) \\ &= (-1)^{|\alpha|+|\beta|} \det X \det Y. \end{aligned}$$

II. For $\alpha, \beta \in Q_{m,n}$ and $\rho, \tau \in Q_{k,m}$, we have

$$\sum_{\sigma \in S_k} \det(X[\alpha|\beta]^{\sigma}_{\tau}[\rho|\tau]) = k!\Delta(X^1[\alpha|\beta][\rho|\tau], \dots, X^k[\alpha|\beta][\rho|\tau])$$

To check this, consider the columns of the matrices involved. Remember that

$$X[\alpha|\beta]^{\sigma}_{\tau} = \mathbf{0}(\tau, X^{\sigma(1)}[\alpha|\beta], \dots, X^{\sigma(k)}[\alpha|\beta]).$$

For given $\sigma \in S_k$ and $j \in \{1, 2, \dots, k\}$, we have:

$$\begin{aligned} \text{entry } (i,j) \text{ of } X[\alpha|\beta]^{\sigma}_{\tau}[\rho|\tau] &= \text{entry } (\rho(i), \tau(j)) \text{ of } X[\alpha|\beta]^{\sigma}_{\tau} \\ &= \text{entry } (\rho(i), \tau(j)) \text{ of } X^{\sigma(j)}[\alpha|\beta]_{\tau(j)} \\ &= \text{entry } (i,j) \text{ of } X^{\sigma(j)}[\alpha|\beta][\rho|\tau]. \end{aligned}$$

Therefore,

$$X[\alpha|\beta]^{\sigma}_{\tau}[\rho|\tau] = [X^{\sigma(1)}[\alpha|\beta][\rho|\tau]_{[1]} \, \dots \, X^{\sigma(k)}[\alpha|\beta][\rho|\tau]_{[k]})$$

and the matrices that appear in the first sum are the same as the ones that appear in the mixed discriminant.

We are now ready to prove the result. If we replace in Theorem 2.5.5

$d_\chi = \det$, we have that that the $(\alpha, \beta)$ entry of $D^k \wedge^m (A)(X^1, \ldots, X^k)$ is

$$\begin{aligned}
& \frac{1}{m!} \sum_{\gamma,\delta \in Q_{m,n}} b_{\gamma\beta}\overline{b_{\delta\alpha}} \sum_{\sigma \in S_k} \sum_{\rho,\tau \in Q_{k,m}} \det(X[\delta|\gamma]^\sigma_\tau[\rho|\tau] \bigoplus_{\rho|\tau} A[\delta|\gamma](\rho|\tau)) \\
= \ & \frac{1}{m!} m! \sum_{\sigma \in S_k} \sum_{\rho,\tau \in Q_{k,m}} \det(X[\alpha|\beta]^\sigma_\tau[\rho|\tau] \bigoplus_{\rho|\tau} A[\alpha|\beta](\rho|\tau)) \\
= \ & \sum_{\sigma \in S_k} \sum_{\rho,\tau \in Q_{k,m}} (-1)^{|\rho|+|\tau|} \det(A[\alpha|\beta](\rho|\tau)) \det(X[\alpha|\beta]^\sigma_\tau[\rho|\tau]) \\
= \ & k! \sum_{\rho,\tau \in Q_{k,m}} (-1)^{|\rho|+|\tau|} \det(A[\alpha|\beta](\rho|\tau)) \Delta(X^1[\alpha|\beta][\rho|\tau], \ldots, X^k[\alpha|\beta][\rho|\tau]).
\end{aligned}$$

We denoted by

$$\Delta(B_1, ..., B_n)$$

the mixed discriminant. The formula we obtained is formula (2.5) in [17], if you take into account that in this paper the roles of the letters $k$ and $m$ are interchanged.

Using similar arguments we can obtain the formula for the $k$-th derivative of $\vee^m(A)(X^1, \ldots, X^k)$ in [15].

## 2.6 Formulas for $k$-th Derivatives of $K_\chi(T)$

We now present a formula for higher order derivatives of $K_\chi(T)$ that generalizes formulas in [7] and [8].

**Definition 2.6.1.** Let $V$ be an $n$ dimensional Hilbert space, let $S^1, \ldots, S^m \in \mathcal{L}(V)$ and let $\chi$ be an irreducible character of $S_m$. We define an operator on $\otimes^m V$ as

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

**Proposition 2.6.2.** *Let $S^1, \ldots, S^m \in \mathcal{L}(V)$ and let $\chi$ be an irreducible character of $S_m$. The space $V_\chi$ is invariant for the operator $S^1 \tilde{\otimes} S^2 \tilde{\otimes} \ldots \tilde{\otimes} S^m$.*

*Proof.* We only have to prove that $S^1 \tilde{\otimes} S^2 \tilde{\otimes} \ldots \tilde{\otimes} S^m$ sends decomposable symmetrized tensors to elements of $V_\chi$. Let $u_1 * \ldots * u_m \in V_\chi$.

$$S^1\tilde{\otimes}S^2\tilde{\otimes}\ldots\tilde{\otimes}S^m(u_1 * u_2 * \ldots * u_m) =$$

$$\begin{aligned}
&= \frac{1}{m!}\sum_{\sigma\in S_m} S^{\sigma(1)}\otimes S^{\sigma(2)}\otimes\cdots\otimes S^{\sigma(m)}(\sum_{\tau\in S_m}\chi(\tau)P(\tau)(u_1\otimes u_2\otimes\ldots\otimes u_m))\\
&= \frac{1}{m!}\sum_{\sigma\tau\in S_m}\chi(\tau)P(\tau)S^{\sigma(1)}\otimes S^{\sigma(2)}\otimes\cdots\otimes S^{\sigma(m)}(u_1\otimes u_2\otimes\ldots\otimes u_m)\\
&= \frac{1}{m!}\sum_{\sigma\tau\in S_m}\chi(\tau)P(\tau)S^{\sigma(1)}(u_1)\otimes S^{\sigma(2)}(u_2)\otimes\cdots\otimes S^{\sigma(m)}(u_m)\\
&= \frac{1}{m!}\sum_{\sigma\in S_m} S^{\sigma(1)}(u_1) * S^{\sigma(2)}(u_2) * \cdots * S^{\sigma(m)}(u_m).
\end{aligned}$$

This belongs to $V_\chi$. □

We denote the restriction of the map $S^1\tilde{\otimes}S^2\tilde{\otimes}\ldots\tilde{\otimes}S^m$ to $V_\chi$ by

$$S^1 * S^2 * \cdots * S^m$$

and call it the *symmetrized $\chi$-symmetric tensor product* of the operators $S^1, S^2, \ldots, S^m$. We remark that the notation chosen to represent the symmetrized $\chi$-symmetric tensor product does not convey the fact that the product depends on the character $\chi$.

In [8] the following proposition was proved.

**Proposition 2.6.3.** *Let $V$ be an $n$-dimensional Hilbert space. For $1 \le k \le m \le n$, suppose $T, S^1, \ldots, S^m \in \mathcal{L}(V)$. Then*

$$D^k(\otimes^m T)(S^1,\ldots,S^k) = \frac{m!}{(m-k)!}\underbrace{T\tilde{\otimes}\cdots\tilde{\otimes}T}_{m-k\ copies}\tilde{\otimes}S^1\tilde{\otimes}\cdots\tilde{\otimes}S^k. \tag{2.8}$$

*If $k > m$ all derivatives are zero.*

*Proof.* In this context we can use Theorem 2.3.10, which states that this derivative is the coefficient of $t_1 \ldots t_k$ in the polynomial

$$\otimes^m(T + t_1S^1 + \ldots + t_kS^k),$$

given by the expression

$$\sum_{\sigma\in S_k}\sum_{\alpha\in Q_{k,m}} \otimes^m(T(\alpha; S^{\sigma(1)},\dots S^{\sigma(k)})).$$

For a given $\alpha \in Q_{k,m}$, the summand $\otimes^m(T(\alpha; S^{\sigma(1)},\dots S^{\sigma(k)}))$ appears $(m-k)!$ times in the expression of $T\tilde{\otimes}\cdots\tilde{\otimes}T\tilde{\otimes}S^1\tilde{\otimes}\cdots\tilde{\otimes}S^k$, since $T$ appears $m-k$ times as a factor. Thus,

$$\frac{(m-k)!}{m!}\sum_{\sigma\in S_k}\sum_{\alpha\in Q_{k,m}} \otimes^m(T(\alpha; S^{\sigma(1)},\dots S^{\sigma(k)})) = T\tilde{\otimes}\cdots\tilde{\otimes}T\tilde{\otimes}S^1\tilde{\otimes}\cdots\tilde{\otimes}S^k.$$

This proves the formula. □

From this proposition we can deduce a formula for $D^kK_\chi(T)(S^1,\dots,S^k)$, using the same techniques that were used in the previous proof.

**Theorem 2.6.4.** *Let $V$ be an $n$ dimensional Hilbert space, $1\le k\le m\le n$, $T, S^1,\dots,S^m \in \mathcal{L}(V)$ and $\chi$ an irreducible character of $S_m$. Then*

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

*In particular, if $k=m$ then*

$$D^mK_\chi(T)(S^1,\dots,S^m) = m!S^1*S^2\cdots*S^m.$$

*For $k>m$ 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 $V_\chi$. We have that for any operators $T_1, T_2, \dots T_m$ of $V$

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

Both maps $L\mapsto Q^*T$ and $T\mapsto LQ$ are linear, so we can apply formula (2.8) and a derivation rule from Proposition 2.1.6 and get

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

This concludes the proof. □

In the previous section we have calculated several formulas for the $k$-th derivative of $K_\chi(A)$, in the directions of $(X^1, X^2, \dots, X^k)$. We have seen in Theorem 2.5.5 that the $(\alpha, \beta)$ entry of this matrix is equal to

$$\frac{\chi(\mathrm{id})}{(m-k)!} \sum_{\gamma,\delta \in \widehat{\Delta}} b_{\gamma\beta}\overline{b_{\delta\alpha}} \Delta_\chi(A[\delta|\gamma]; X^1[\delta|\gamma], \dots, X^k[\delta|\gamma]).$$

If we look at the formula obtained in the last theorem it is clear that there are similarities between these two expressions. We now establish a relation between $D^kK_\chi(T)(S^1, \dots, S^k)$ and $D^kK_\chi(A)(X^1, \dots, X^k)$.

Recall that $E = \{e_1, e_2, \dots, e_n\}$ is an orthonormal basis of $V$, $\mathcal{E}' = \{e^*_\alpha : \alpha \in \widehat{\Delta}\}$ is the induced basis of $V_\chi$ and $\mathcal{E} = \{v_\alpha : \alpha \in \widehat{\Delta}\}$ is the orthonormal basis of $V_\chi$ obtained by applying the Gram-Schmidt orthogonalization to the basis $\mathcal{E}'$.

**Proposition 2.6.5.** *Let $T, S^1, S^2, \dots, S^k$ be operators on the Hilbert space $V$ and let $E$ be an orthonormal basis of $V$. Let $A, X^1, X^2, \dots, X^k$ be $n \times n$ complex matrices such that $A = M(T; E)$, $X^i = M(S^i; E), i = 1, 2, \dots, k$. Then*

$$D^kK_\chi(A)(X^1, \dots, X^k) = M(D^kK_\chi(T)(S^1, \dots, S^k); \mathcal{E}),$$

*where $\mathcal{E} = \{v_\alpha : \alpha \in \widehat{\Delta}\}$ is the orthonormal basis of $V_\chi$ that is obtained by applying the Gram-Schmidt orthogonalization to the induced basis $\mathcal{E}'$.*

In order to prove the previous proposition we need the following lemma.

**Lemma 2.6.6.** *Suppose $T_1, T_2, \dots, T_m$ are linear operators in the Hilbert space $V$ and let $E = \{e_1, e_2, \dots, e_n\}$ be an orthonormal basis of $V$. Let $A_1, A_2, \dots, A_m$ be $n \times n$ complex matrices such that $A_i = M(T_i; E)$, for $i = 1, 2, \dots, m$. Let $\chi$ be an irreducible character of $S_m$ and $\delta, \gamma \in \widehat{\Delta}$. Then*

$$\langle T_1 \tilde{\otimes} T_2 \tilde{\otimes} \dots \tilde{\otimes} T_m(e^*_\delta), e^*_\gamma \rangle = \frac{\chi(\mathrm{id})}{m!} \Delta_\chi\left(A_1[\gamma|\delta], A_2[\gamma|\delta], \dots, A_m[\gamma|\delta]\right).$$

*Proof.* We have that

$$T_1\tilde{\otimes}T_2\tilde{\otimes}\ldots\tilde{\otimes}T_m(e^*_\delta) = T_1\tilde{\otimes}T_2\tilde{\otimes}\ldots\tilde{\otimes}T_m\left(\sum_{\sigma\in S_m}\chi(\sigma)P(\sigma)(e^{\otimes}_\delta)\right)$$

$$\begin{aligned}
&= \sum_{\sigma\in S_m}\chi(\sigma)P(\sigma)T_1\tilde{\otimes}T_2\tilde{\otimes}\ldots\tilde{\otimes}T_m(e^{\otimes}_\delta)\\
&= \frac{1}{m!}\sum_{\sigma\in S_m}\chi(\sigma)P(\sigma)\sum_{\tau\in S_m}T_{\tau(1)}\otimes T_{\tau(2)}\otimes\ldots T_{\tau(m)}(e_{\delta(1)}\otimes e_{\delta(2)}\otimes\ldots\otimes e_{\delta(m)})\\
&= \frac{1}{m!}\sum_{\sigma,\tau\in S_m}\chi(\sigma)P(\sigma)T_{\tau(1)}(e_{\delta(1)})\otimes T_{\tau(2)}(e_{\delta(2)})\otimes\ldots\otimes T_{\tau(m)}(e_{\delta(m)})\\
&= \frac{1}{m!}\sum_{\tau\in S_m}T_{\tau(1)}(e_{\delta(1)})*T_{\tau(2)}(e_{\delta(2)})*\ldots*T_{\tau(m)}(e_{\delta(m)}).
\end{aligned}$$

The induced inner product in $V_\chi$

$$\langle T_{\tau(1)}(e_{\delta(1)})*T_{\tau(2)}(e_{\delta(2)})*\ldots*T_{\tau(m)}(e_{\delta(m)}), e_{\gamma(1)}*e_{\gamma(1)}\ldots*e_{\gamma(m)}\rangle$$

is by Proposition 1.3.24 equal to

$$\begin{aligned}
&\frac{\chi(\mathrm{id})}{m!}d_\chi\langle T_{\tau(i)}(e_{\delta(i)}), e_{\gamma(j)}\rangle =\\
&= \frac{\chi(\mathrm{id})}{m!}d_\chi(\langle T_{\tau(1)}(e_{\delta(1)}), e_{\gamma(j)}\rangle_{[1]}, \langle T_{\tau(2)}(e_{\delta(2)}), e_{\gamma(j)}\rangle_{[2]}, \ldots, \langle T_{\tau(m)}(e_{\delta(m)}), e_{\gamma(j)}\rangle_{[m]})\\
&= \frac{\chi(\mathrm{id})}{m!}d_\chi(A_{\tau(1)}[\delta|\gamma]_{[1]}, A_{\tau(2)}[\delta|\gamma]_{[2]}, \ldots, A_{\tau(m)}[\delta|\gamma]_{[m]}).
\end{aligned}$$

So,

$$\begin{aligned}
&(T_1\tilde{\otimes}T_2\tilde{\otimes}\ldots\tilde{\otimes}T_m\langle e^*_\delta), e^*_\gamma\rangle =\\
&= \frac{1}{m!}\sum_{\tau\in S_m}\frac{\chi(\mathrm{id})}{m!}d_\chi(A_{\tau(1)}[\delta|\gamma]_{[1]}, A_{\tau(2)}[\delta|\gamma]_{[2]}, \ldots, A_{\tau(m)}[\delta|\gamma]_{[m]})\\
&= \frac{\chi(\mathrm{id})}{m!}\Delta_\chi\left(A_1[\gamma|\delta], A_2[\gamma|\delta], \ldots, A_m[\gamma|\delta]\right).
\end{aligned}$$

This concludes our proof. □

Now we can prove Proposition 2.6.5.

*Proof.* Let $\alpha, \beta \in \widehat{\Delta}$, since the basis $\mathcal{E}$ is orthonormal, the $(\alpha, \beta)$-entry of the matrix that represents $D^k K_\chi(T)(S^1, \ldots, S^k)$ in the basis $\mathcal{E}$ is:

$$\begin{aligned}
& \langle D^k K_\chi(T)(S^1, \ldots, S^k)\langle v_\beta), v_\alpha \rangle = \\
= \; & \frac{m!}{(m-k)!} \left\langle T * T * \ldots * T * S^1 * \ldots * S^k(v_\beta), v_\alpha \right\rangle \\
= \; & \frac{m!}{(m-k)!} \left\langle T \tilde{\otimes} T \tilde{\otimes} \ldots \tilde{\otimes} T \tilde{\otimes} S^1 \tilde{\otimes} \ldots \tilde{\otimes} S^k(v_\beta), v_\alpha \right\rangle \\
= \; & \frac{m!}{(m-k)!} \sum_{\gamma, \delta \in \widehat{\Delta}} b_{\gamma\beta} \overline{b_{\delta\alpha}} \langle T \tilde{\otimes} T \tilde{\otimes} \ldots \tilde{\otimes} T \tilde{\otimes} S^1 \tilde{\otimes} \ldots \tilde{\otimes} S^k(e^*_\delta), e^*_\gamma \rangle.
\end{aligned}$$

By the previous Lemma, we have that

$$\langle T \tilde{\otimes} T \tilde{\otimes} \ldots \tilde{\otimes} T \tilde{\otimes} S^1 \tilde{\otimes} \ldots \tilde{\otimes} S^k(e^*_\delta), e^*_\gamma \rangle = \frac{\chi(\mathrm{id})}{m!} \Delta_\chi \left( A[\gamma|\delta; , X^1[\gamma|\delta], \ldots, X^k[\gamma|\delta] \right).$$

So we can conclude that the $(\alpha, \beta)$-entry of the matrix of the operator $D^k K_\chi(T)(S^1, \ldots, S^k)$ is equal to

$$\frac{\chi(\mathrm{id})}{(m-k)!} \sum_{\gamma, \delta \in \widehat{\Delta}} b_{\gamma\beta} \overline{b_{\delta\alpha}} \Delta_\chi \left( A[\gamma|\delta]; , X^1[\gamma|\delta], \ldots, X^k[\gamma|\delta] \right),$$

which is exactly the $(\alpha, \beta)$-entry of the matrix $D^k K_\chi(A)(X^1, \ldots, X^k)$. $\square$

# Chapter 3

# Variation of Multilinear Induced Operators

*The definition of a good mathematical problem is the mathematics it generates rather then the problem itself.*

Andrew Wiles

Let $V$ be a complex Hilbert space and $\mathcal{L}(V)$ be the vector space of linear operators from $V$ to itself. Let $T$ be an operator in $\mathcal{L}(V)$ and let $\|.\|$ be the vector norm on $V$ defined by the inner product. The *operator norm induced by the vector norm* $\|.\|$ is defined by

$$\|T\| = \sup_{\|x\|=1} \|Tx\|.$$

This norm is called *operator bound norm* or *spectral norm* of $T$. From now on we always consider the spectral norm on $\mathcal{L}(V)$. If $\nu_1 \geq \nu_2 \geq \ldots \geq \nu_n$ are the singular values of $T$, then $\|T\| = \nu_1$.

R. Bhatia and S. Friedland have addressed in [7] the problem of finding the norm of the derivative of the Grassmann power of a matrix, which led to a remarkable formula:

$$\|D \wedge^m (A)\| = p_{m-1}(\nu_1, \ldots, \nu_m) \tag{3.1}$$

where $p_{m-1}$ is the symmetric polynomial of degree $m-1$ in $m$ variables. It should be noted that this formula gives an exact value for the derivative and not just an upper bound.

After this, R. Bhatia has addressed the problem of finding a similar formula for the symmetric power, which was done in [5], finding that this norm was the same as that of the derivative of the $m$-th fold tensor power:

$$\|D \vee^m A\| = m\|A\|^{m-1} = m\nu_1^{m-1}. \tag{3.2}$$

We note that $m\nu_1^{m-1} = p_{m-1}(\nu_1, \ldots, \nu_1)$, where $\nu_1$ appears $m$ times.

Generalizing both these results, R. Bhatia and J. A. Dias da Silva established in [6] a formula for the norm of the first derivative for all $\chi$-symmetric powers of an operator. Being technically more intricate than the previous results, the final formula also involves the value of the symmetric polynomial of degree $m-1$ on $m$ variables calculated on some of the singular values of $A$.

Recently, R. Bhatia, P. Grover and T. Jain have established formulas that generalize (3.1) and (3.2) in a diferent direction: formulas are established for higher order derivatives of symmetric powers and Grassmann powers. Before we briefly describe these results we recall the definition of norm for multilinear maps — the $m$-th directional derivative is a multilinear map.

Let $V$ and $U$ be complex Hilbert spaces and let $\Phi : (\mathcal{L}(V))^m \longrightarrow \mathcal{L}(U)$ be a multilinear operator. The *norm of* $\Phi$ is given by

$$\|\Phi\| = \sup_{\|X^1\|=\ldots=\|X^m\|=1} \|\Phi(X^1, \ldots, X^m)\|.$$

In papers [17] and [8], the authors have established the following values:

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

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

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 $T$ (possibly with repetitions), multiplied by $k!$.

In our work, which was done simultaneously with [8], formulas are established that subsume all the aforementioned formulas. The techniques are similar, but the fact that the symmetry classes we considered were associated with an arbitrary irreducible character made it necessary to use more intricate combinatorial theorems.

## 3.1 Basic Concepts and Results

Recall the operator $K_\chi(T)$, Definition 1.2.26. In this chapter we will obtain exact values for the norm of $k$-th derivative of the operator $f(T) = K_\chi(T)$.

**Definition 3.1.1.** Let $T \in \mathcal{L}(V)$. The operator $T$ is *positive semidefinite* if for every $x \in V$, $\langle x, T(x)\rangle \geq 0$ and $T$ is *positive definite* if for every $x \in V$, $x \neq 0$ $\langle x, T(x)\rangle > 0$.

For any $T \in \mathcal{L}(V)$ the operator $TT^*$ is always positive semidefinite, so its eigenvalues are always nonnegative. The eigenvalues of this operator are called the *singular values* of the operator $T$.

**Definition 3.1.2.** Let $V$ and $U$ be complex Hilbert spaces and let $\Phi$ : $(\mathcal{L}(V))^m \to \mathcal{L}(U)$ be a multilinear operator. We say that $\Phi$ is *positive* if $\Phi(X^1, \ldots, X^m)$ is positive definite for every family $X^1, \ldots X^m \in \mathcal{L}(V)$ of positive definite operators.

**Definition 3.1.3.** Let $m$, $n$ be positive integers $1 \leq m \leq n$, $x_1, x_2, \ldots x_n$ $n$ variables. The *elementary symmetric polynomial* in $n$ variables and degree $m$ is the homogeneous symmetric polynomial defined as

$$p_m(x_1, x_2, \ldots, x_n) = \sum_{\alpha \in Q_{m,n}} x_{\alpha(1)} x_{\alpha(2)} \ldots x_{\alpha(m)}.$$

**Example 3.1.4.**

1. $p_1(x_1, \ldots, x_m) = x_1 + x_2 + ... + x_m$,
2. $p_2(x_1, x_2, x_3) = x_1x_2 + x_1x_3 + x_2x_3$,
3. $p_m(x_1, \ldots, x_m) = x_1x_2 \ldots x_m$.

**Definition 3.1.5.** Let $m, r$ be positive integers, $1 \leq r \leq m$. A *partition* $\pi$ of $m$ is an $r$-tuple of positive integers $\pi = (\pi_1, \ldots, \pi_r)$, such that

- $\pi_1 \geq \ldots \geq \pi_r$,
- $\pi_1 + \ldots + \pi_r = m$.

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 $\pi$ is called the *length of* $\pi$ and is represented by $l(\pi)$.

**Example 3.1.6.** If $m = 10$, $\pi = (4, 2, 2, 1, 1, 0, 0, 0, 0, 0)$, $\overline{\pi} = (10, 0, 0, 0, 0, 0, 0, 0, 0, 0)$ and $\pi' = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$ are partitions of 10. $l(\pi) = 5$, $l(\overline{\pi}) = 1$ and $l(\pi') = 10)$.

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

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

It is known from representation theory (see [25]) 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 $\chi = (1, \ldots, 1)$ then $V_\chi = \wedge^m V$ is the Grassmann space, and if $\chi = (m, 0 \ldots, 0)$, then $V_\chi = \vee^m V$.

**Definition 3.1.7.** For every partition $\pi = (\pi_1, \pi_2, \ldots, \pi_{l(\pi)})$ of $m$ we define $\omega(\pi)$ as

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

**Example 3.1.8.** For $\pi = (4, 2, 2, 1, 1, 0, 0, 0, 0, 0)$, $\overline{\pi} = (10, 0, 0, 0, 0, 0, 0, 0, 0, 0)$ and $\pi' = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$ partitions of 10, we have

$$\omega(\pi) = (1, 1, 1, 1, 2, 2, 3, 3, 4, 5);$$

$$\omega(\overline{\pi}) = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1);$$

$$\omega(\pi') = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10).$$

Now to every element of $\Gamma_{m,n}$ we will associate a partition of $m$.

**Definition 3.1.9.** For each $\alpha \in \Gamma_{m,n}$ let $\operatorname{Im}\alpha = \{i_1, \dots, i_l\}$, suppose that $|\alpha^{-1}(i_1)| \geq \dots \geq |\alpha^{-1}(i_l)|$. The partition of $m$

$$\mu(\alpha) := (|\alpha^{-1}(i_1)|, \dots, |\alpha^{-1}(i_l)|) \tag{3.3}$$

is called the *multiplicity partition* of $\alpha$.

**Example 3.1.10.** Consider $\alpha = (2, 2, 5, 4, 8, 8, 8)$, then

$$\mu(\alpha) = (|\alpha^{-1}(8)|, |\alpha^{-1}(2)|, |\alpha^{-1}(4)|, |\alpha^{-1}(5)|) = (3, 2, 1, 1).$$

For $\beta = (3, 3, 3, 1, 2, 5, 5)$, we see that $\mu(\alpha) = \mu(\beta)$. This means that $\mu$ is not an injective map.

So the maps defined before are not inverses of each other, however there is a weaker relation between them.

**Remark 3.1.11.** The multiplicity partition of $\omega(\pi)$ is equal to the partition $\pi$:

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

We have that $\operatorname{Im}\omega(\pi) = \{1, 2, \dots, l(\pi)\}$ 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 $m$.

**Definition 3.1.12.** Let $\mu$ and $\lambda$ be partitions of the positive integer $m$. The partition $\mu$ *precedes* $\lambda$, 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 can be found in [25], that characterizes the elements of the set $\Omega_\chi$.

**Theorem 3.1.13.** *Let $\chi$ be a partition of $m$ and $\alpha \in \Gamma_{m,n}$. Let $\Omega_\chi$ and $\mu(\alpha)$ be as defined in (1.4) and (3.3). Then $\alpha \in \Omega_\chi$ if and only if $\chi$ majorizes $\mu(\alpha)$.*

## 3.2 Norm of the $k$-the Derivative of $K_\chi(T)$

We recall that the norm of a multilinear operator $\Phi : (\mathcal{L}(V))^k \longrightarrow \mathcal{L}(U)$ is given by

$$\|\Phi\| = \sup_{\|S^1\|=\ldots=\|S^k\|=1} \|\Phi(S^1, \ldots, S^k)\|.$$

The main result of this section is the following theorem.

**Theorem 3.2.1.** *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 $\chi$ be a partition of $m$. Suppose $T$ is an operator in $\mathcal{L}(V)$ and $\nu_1 \geq \nu_2 \geq \ldots \geq \nu_n$ its singular values. Let $T \to K_\chi(T)$ be the map that associates to each element of $\mathcal{L}(V)$ the induced operator $K_\chi(T)$ on the symmetry class $V_\chi$. Then the norm of the derivative of order $k$ of this map is given by the formula*

$$\|D^k K_\chi(T)\| = k!\, p_{m-k}(\nu_{\omega(\chi)}) \tag{3.4}$$

*where $p_{m-k}$ is the elementary symmetric polynomial of degree $m-k$ in $m$ variables.*

The proof of our main result is inspired in the techniques used in [6].

**Proposition 3.2.2** (Polar Decomposition)**.** *Let $T$ be a linear operator in $V$. Then there exist a positive semidefinite operator $P$ and a unitary operator $W$, such that*

$$T = PW.$$

*If $T$ is invertible then this decomposition is unique.*

We will now highlight the most important features of the proof of our main theorem.

First we will use the polar decomposition of 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 $(I, I, \ldots, I)$, 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 $D^k K_\chi(P)(I, I, \ldots, I)$, which is also the norm of $D^k K_\chi(T)$.

**Proposition 3.2.3.** *Let $T \in \mathcal{L}(V)$ , $P$ the positive semidefinite operator and $W$ an unitary operator such that $P = TW$. Then*

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

*Proof.* Let $P = TW$, with $W$ unitary. Then $K_\chi(W)$ is also unitary , because

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

Therefore, we have

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

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

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

□

Now we need to estimate the norm of the operator $D^k K_\chi(P)$. For this, we use a result from [8], a multilinear version of the Russo-Dye theorem.

## A Multilinear Version of the Russo-Dye Theorem

Throughout this chapter we are following the techniques of R. Bhatia and J. A. Dias da Silva that were used to obtain the value of the norm for the first derivative. In their work they used the famous Russo-Dye theorem which is sometimes phrased as *every positive linear operator on* $M_n(\mathbb{C})$ *attains its norm at the identity matrix*.

We intend to calculate the norm of higher order derivatives, so we need a multilinear version of this result. But first we present the classical result and its well known corollary.

Let $\varphi : M_n(\mathbb{C}) \longmapsto M_n(\mathbb{C})$ be a linear operator, $\varphi$ is *unital* if $\varphi(I) = I$, where $I$ is the identity matrix of order $n$.

**Definition 3.2.4.** Suppose $\Phi : (\mathcal{L}(V))^k \longrightarrow \mathcal{L}(U)$ is a multilinear operator and $X^1, \dots, X^k \in \mathcal{L}(V)$. $\Phi$ is said to be *positive* if $\Phi(X^1, \dots, X^k)$ is a positive semidefinite linear operator whenever $X^1, \dots, X^k$ are so.

**Theorem 3.2.5** (Russo-Dye Theorem)**.** *If $\varphi$ is a positive and unital operator on $M_n(\mathbb{C})$, then $\|\varphi\| = 1$.*

**Corollary 3.2.6.** *Let $\varphi$ be a positive linear operator, then $\|\varphi\| = \|\varphi(I)\|$.*

We need some definitions and some technical results to prove the multilinear version of the previous theorem.

First we need the following result from [4]. Let us recall that a linear operator $K$ is said to be a *contraction* if $\|A\| \leq 1$.

**Proposition 3.2.7.** *Let $A$ and $B$ be positive matrices of order $n$. Then the matrix $\begin{pmatrix} A & X \\ X^* & B \end{pmatrix}$ is positive if and only if $X = A^{\frac{1}{2}} K B^{\frac{1}{2}}$, for some contraction $K$.*

Now we can prove a multilinear version of the Russo-Dye theorem. This proof was presented to us by Tanvi Jain and can be found in [8]

**Theorem 3.2.8** (A multilinear Russo-Dye Theorem)**.** *Let $\Phi : (\mathcal{L}(V))^k \to \mathcal{L}(U)$ be a positive multilinear operator. Then*

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

*Proof.* Let $U_1,\dots,U_k$ be unitary operators. First we will show that

$$\|\Phi(U_1,\dots,U_k)\| \le \|\Phi(I,I,\dots,I)\|.$$

For each $i=1,\dots,k$, let $U_i=\sum_{j=1}^{r_i}\lambda_{i_j}P_{i_j}$ be the spectral decomposition for $U_i$. Then

$$\Phi(U_1,\dots,U_k)=\sum_{j_1=1}^{r_1}\sum_{j_2=1}^{r_2}\cdots\sum_{j_k=1}^{r_k}\lambda_{1_{j1}}\lambda_{2_{j2}}\dots\lambda_{k_{jk}}\Phi(P_{1_{j1}},P_{2_{j2}},\dots,P_{k_{jk}}).$$

Let $X=\Phi(I,I,\dots,I)$ and $Y=\Phi(U_1,\dots,U_k)$, then we have

$$\begin{pmatrix} X & Y\\ Y^* & X\end{pmatrix}=$$

$$\sum_{j_1=1}^{r_1}\sum_{j_2=1}^{r_2}\cdots\sum_{j_k=1}^{r_k}\begin{pmatrix}1 & \lambda_{1_{j1}}\lambda_{2_{j2}}\dots\lambda_{k_{jk}}\\ \overline{\lambda_{1_{j1}}\lambda_{2_{j2}}\dots\lambda_{k_{jk}}} & 1\end{pmatrix}\otimes\Phi(P_{1_{j1}},P_{2_{j2}},\dots,P_{k_{jk}}) \tag{3.5}$$

is positive semidefinite and hence by the last proposition we can conclude that

$$\Phi(U_1,\dots,U_k)=\Phi(I,I,\dots,I)^{\frac12}K\Phi(I,I,\dots,I)^{\frac12}$$

for some contraction $K$. Thus

$$\|\Phi(U_1,\dots,U_k)\|\le\|\Phi(I,I,\dots,I)\|.$$

In order to prove the other inequality let $A_1,A_2,\dots,A_k$ be contractions. Then each $A_i=\dfrac{U_i+V_i}{2}$ for some unitaries $U_i,V_i$ and $i=1,2,\dots,n$, and hence

$$\begin{aligned}\|\Phi(A_1,A_2,\dots,A_k)\| &= \|\frac{1}{2}\Phi(U_1+V_1,U_2+V_2,\dots,U_k+V_k)\|\\ &\le \frac{1}{2^k}\sum_{j_1=1}^{2}\sum_{j_2=1}^{2}\cdots\sum_{j_k=1}^{2}\|\Phi(X_{j1},X_{j2},\dots,X_{jk})\|,\end{aligned}$$

where $X_{1i}=U_i$ and $X_{2i}=V_i$.

Since each $\|\Phi(X_{j1},X_{j2},\dots,X_{jk})\|\le\|\Phi(I,I,\dots,I)\|$ and there are $2^k$ summands, we have $\|\Phi(A_1,A_2,\dots,A_k)\|\le\|\Phi(I,I,\dots,I)\|$. □

Now we can apply this result to our particular case.

We have that $D^kK_\chi(P)$ is a positive multilinear operator, since if $X^1, \dots, X^k$ are positive semidefinite, then by the formula in Theorem 2.6.4, $D^k(P)(X^1, \dots, X^k)$ is the restriction of a positive semidefinite operator to an invariant subspace, and thus is positive semidefinite.

Therefore,

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

Now we have to find the maximum eigenvalue of the positive operator $D^kK_\chi(P)(I, I, \dots, I)$. This will be done by finding a basis of $V_\chi$ formed by eigenvectors of $D^kK_\chi(P)(I, I, \dots, I)$. We will see that 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 $D^kK_\chi(P)(I, I, \dots, I)$ (in general, it will not be orthonormal).

**Definition 3.2.9.** Let $1 \le k \le m$ and $\beta \in Q_{m-k,k}$. Define $\otimes^m_\beta P$ as the tensor $X^1 \otimes \cdots \otimes X^m$, in which $X^i = P$ if $i \in \operatorname{Im}\beta$ and $X^i = I$ otherwise.

**Lemma 3.2.10.** *Suppose $P \in \mathcal{L}(V)$ and let $I$ be the identity operator.*

1. *We have*

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

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

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

*Proof.* 1. It is a matter of carrying out the computations: for each $\beta$, the summand $\otimes^m_\beta P$ appears $k!(m-k)!$ times in $P\tilde{\otimes}\cdots\tilde{\otimes}P\tilde{\otimes}I\tilde{\otimes}\cdots\tilde{\otimes}I$, since there are $k$ repetitions of the symbol $I$ and $m-k$ repetitions of the symbol $P$.

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

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

So,

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

This concludes the proof. □

The following proposition gives us the expression for all the eigenvalues of $D^kK_\chi(P)(I,I,\ldots,I)$.

**Proposition 3.2.11.** *Let $\alpha\in\widehat{\Delta}$ and define*

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

*Then $\lambda(\alpha)$ is the eigenvalue of $D^kK_\chi(P)(I,I,\ldots,I)$ associated with the eigenvector $e^*_\alpha$.*

*Proof.* Recall that $E=\{e_1,\ldots,e_n\}$ is an orthonormal basis of eigenvectors for $P$, with eigenvalues $\nu_1,\ldots,\nu_n$. For every $\alpha\in\Gamma_{m,n}$ we have

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

Then

$$\begin{aligned}D^kK_\chi(P)(I,\ldots,I)(e^*_\alpha) &= \frac{m!}{(m-k)!}(P\tilde{\otimes}\cdots\tilde{\otimes}P\tilde{\otimes}I\tilde{\otimes}\cdots\tilde{\otimes}I)(e^*_\alpha)\\ &= \frac{m!}{(m-k)!}\frac{k!(m-k)!}{m!}\sum_{\beta\in Q_{m-k,m}}\otimes^m_\beta P(e^*_\alpha)\\ &= k!\sum_{\sigma\in S_m}\chi(\sigma)\sum_{\beta\in Q_{m-k,m}}\otimes^m_\beta P(e^\otimes_{\alpha\sigma})\\ &= k!\sum_{\sigma\in S_m}\chi(\sigma)p_{m-k}(\nu_{\alpha\sigma})e^\otimes_{\alpha\sigma}\\ &= k!\sum_{\sigma\in S_m}\chi(\sigma)p_{m-k}(\nu_\alpha)e^\otimes_{\alpha\sigma}\\ &= k!p_{m-k}(\nu_\alpha)e^*_\alpha\end{aligned}$$

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

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

**Lemma 3.2.12.** *Suppose $\alpha, \beta \in \Gamma_{m,n}$. If $\alpha$ and $\beta$ are in the same orbit, then*

$$\lambda(\alpha) = \lambda(\beta).$$

*Proof.* If $\alpha$ and $\beta$ 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. □

We have already seen that 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 3.2.13.** *Let $\alpha, \beta$ be elements of $\overline{\Delta} \subseteq G_{m,n}$. Then $\lambda(\alpha) \geq \lambda(\beta)$ if and only if $\alpha$ precedes $\beta$ in the lexicographic order.*

*Proof.* The result follows directly from the expression of the eigenvalues of $D^k K_\chi(P)(I, \ldots, I)$ given in Proposition 3.2.11. □

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

*Proof.* (of Theorem 3.2.1). We have that $\omega(\chi) \in \overline{\Delta}$, so we must have

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

Now let $\alpha \in \overline{\Delta}$. Using the results from Theorem 3.1.13 and Remark 3.1.11, we have that $\chi = \mu(\omega(\chi))$ majorizes $\mu(\alpha)$.
By the definition of multiplicity partition, we have that $\omega(\chi)$ precedes $\alpha$ in the lexicographic order. By Lemma 3.2.13, we then have

$$\lambda(\omega(\chi)) \geq \lambda(\alpha)$$

and

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

This concludes the proof of the theorem. □

Now observe that the formulas obtained by Jain [17] and Grover [15] are particular cases of this last formula.

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

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

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

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

where $p_{m-k}(\nu_1, \nu_2, \cdots, \nu_m)$ 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 [6]. Just notice that if $k = 1$, we have that $Q_{1,m} = \{1, 2, \cdots, m\}$, so

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

## 3.3 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 we recall 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. For this, we also recall the definition of $K_\chi(A)$, the $m$-th $\chi$-symmetric tensor power of the matrix $A$. We fix an orthonormal basis $E$ in $V$, and consider the linear endomorphism $T$ such that $A = M(T, E)$. We have seen that $\mathcal{E}' = (e^*_\alpha : \alpha \in \widehat{\Delta})$ is the induced basis of $V_\chi$. Recall that $\mathcal{E} = (v_\alpha : \alpha \in \widehat{\Delta})$ is the orthonormal basis of the $m$-th $\chi$-symmetric tensor power of the vector space $V$ obtained by applying the Gram-Schmidt orthonormalization procedure to $\mathcal{E}'$. We have that

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

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 [25, p. 236].

Since the basis chosen in $V_\chi$ is orthonormal, the result for the norm of the operator applies to this matrix:

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

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

We have denoted by $\mathrm{imm}_\chi(A)$ the matrix with rows and columns indexed by $\widehat{\Delta}$, whose $(\gamma, \delta)$ entry is $d_\chi(A[\gamma|\delta])$. 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

$$K_\chi(A) = \frac{\chi(\mathrm{id})}{m!} B^* \,\mathrm{imm}_{\overline{\chi}}(A) B. \tag{3.6}$$

We also have

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

Notice the similarity with the formula in Theorem 2.6.4.

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 3.3.1.** *Keeping with the notation established, we have that, for $k \le n$,*

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

*Proof.* We always have $Q_{m,n} \subseteq \widehat{\Delta}$.
We now take $m = n$ and denote $\gamma := (1, 2, \ldots, n) \in Q_{n,n} \subseteq \widehat{\Delta}$ (this is the only element in $Q_{n,n}$). By definition, $d_\chi(A)$ is the $(\gamma, \gamma)$ entry of $\mathrm{imm}_\chi(A)$, and, according to formula (3.6), we have

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

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

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

We denote by $C$ the column $\gamma$ of the matrix $B^{-1}$:

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

Then

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

By formula (1.2), we have that

$$\|e^*_\gamma\|^2 = \frac{\chi(\mathrm{id})}{n!}.$$

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

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

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

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

This concludes the proof. □

This upper bound coincides with the norm of the derivative of the determinant obtained in [9]. When $d_\chi = \mathrm{per}$, the upper bound presented in formula (52) in [8] is, using our notation, $(n!/(n-k)!)\|A\|^{n-k}$. Using our formula, we get the same value: for $\omega(\chi) = (1, 1, \ldots, 1)$,

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

It is also shown that for

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

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

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

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

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

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

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

# Concluding Remarks

*If you want a happy ending, that depends, of course, on where you stop the story...*

Orson Welles

The aim of this dissertation has been to generalize the formulas of higher order derivatives of certain matrix functions and its norms. Throughout this process we have proved different formulas using several different processes and techniques. Sometimes it was purely a matter of using classical results of multilinear algebra, other times we had to apply heavy and new technical results of functional and matrix analysis.

Fortunately, our goal has been attained and in a way that we first did not expect, some of the general formulas look simpler than the ones proved in particular cases. Still, we have various questions that are unanswered.

1. In [8] the authors have proved that their results for norms hold also in the infinite dimensional case. Can we state the same formulas if we consider the infinite dimension case, and if so, which adaptations should be made?

2. We have found an upper bound for the norm of $D^k d_\chi(A)$. It is interesting to find a matrix $A$ where the equality holds and also to check if the inequality is sharp.

3. While generalizing the formulas for $D^k K_\chi(A)(X^1, X^2, \ldots, X^k)$, one of our drawbacks is the fact that we do not know the basis of $V_\chi$ in the general case. However, if the irreducible character $\chi$ is associated with a special partition of $m$, usually called a *hook partition*, there is quite

some work done in these cases and the basis of the space $V_\chi$ is known. The formulas that we obtained might be improved, if we consider these family of irreducible characters.

# Bibliography

[1] N. Akhiezer, I. Glazman, *Theory of Linear Operators in Hilbert Space*, Dover Publications Inc, New York (1993).

[2] R. B. Bapat, Mixed discriminants of positive semidefinite matrices, *Linear Algebra and its Applications*, 126 (1989), 107–124.

[3] R. Bhatia, *Matrix Analysis*, Springer New York (1997).

[4] R. Bhatia, *Positive Definite Matrices*, Princeton University Press, New Jersey (2007).

[5] R. Bhatia, Variation of symmetric tensor powers and permanents, *Linear Algebra and its Applications*, 62 (1984), 269–276.

[6] R. Bhatia, J. A. Dias da Silva, Variation of induced linear operators, *Linear Algebra and its Applications*, 341 (2002), 391–402.

[7] R. Bhatia, S. Friedland, Variation of Grassmann powers and spectra, *Linear Algebra and its Applications*, 40 (1981), 1–18.

[8] R. Bhatia, P. Grover, T. Jain, Derivatives of tensor powers and their norms, *The Electronic Journal of Linear Algebra*, 26 (2013), 604–619.

[9] R. Bhatia, T. Jain, Higher order derivatives and perturbation bounds for determinants *Linear Algebra and its Applications*, 431 (2009), 2102–2108.

[10] S. Carvalho, P. J. Freitas, The $k$-th derivatives of the immanant and the $\chi$-symmetric power of an operator, submitted for publication in *The Electronic Journal of Linear Algebra*.

[11] S. Carvalho, P. J. Freitas, The norm of the $k$-th derivative of the $\chi$-symmetric power of an operator, accepted for publication in *The Electronic Journal of Linear Algebra*.

[12] J. A. Dias da Silva, *Classes Simétricas de Tensores*, Lisboa (1989).

[13] J. A. Dias da Silva, H. Godinho, Generalized derivations and additive theory, *Linear Algebra and its Applications*, 342 (2002), 1–15.

[14] F. R. Gantmacher, *Matrix Theory*, vol I, AMS Chelsea Publishing (1959).

[15] P. Grover, Derivatives of multilinear functions of matrices, F. Nielsen and R. Bhatia, eds., *Matrix Information Geometry*, Springer, Heidelberg (2013), 93–109.

[16] R. Horn, C. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge (1987).

[17] T. Jain, Derivatives for antisymmetric tensor powers and perturbation bounds, *Linear Algebra and its Applications*, 435 (2011), 1111–1121.

[18] O. Kim, J. Chollet, R. Brown, D. Rauschenberg, Orthonormal bases of symmetry classes with computer-generated examples, *Linear and Multilinear Algebra*, 21 (1987), 91–106.

[19] J. Magnus, H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Wiley and Sons, New York (1988).

[20] M. Marcus, Inequalities for matrix functions of combinatorial interest, *SIAM Journal on Applied Mathematics*, 17 (1969), 1023–1031.

[21] M. Marcus, J. Chollet, Construction of orthonormal bases in higher symmetry classes of tensors, *Linear and Multilinear Algebra*, 19 (1986), 133–140.

[22] M. Marcus, H. Minc, Generalized matrix functions, *Trans. Amer. Math. Soc.*, 116 (1965), 316–329.

[23] M. Marcus, H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Dover Publications, New York (1992, reprint of 1964 edition).

[24] M. Marcus, *Finite Dimensional Multilinear Algebra - Part I*, Marcel Dekker, New York (1973).

[25] R. Merris, *Multilinear Algebra*, Gordon and Breech Science Publishers, Singapore (1997).

[26] R. Merris, The structure of higher degree symmetry classes of tensors II, *Linear and Multilinear Algebra*, 6 (1978), 171–178.

[27] R. Merris, Inequalities and identities for generalized matrix functions, *Linear Algebra and Its Applications*, 64 (1985), 223–242.

[28] H. Minc, *Permanents*, Addison-Wesley, Reading MA (1978).

[29] V. S. Sunder, A noncommutative analogue of $|D(X^k)| = |kX^{k-1}|$, *Linear Algebra and its Applications*, 44 (1982), 87–95.

# Index