

Algebra, Logic and Applications Series Volume 8

Multilinear Algebra

Russell Merris



CRC Press
Taylor & Francis Group

MULTILINEAR ALGEBRA

ALGEBRA, LOGIC AND APPLICATIONS

A series edited by

R. Göbel

Universität Gesamthochschule, Essen, Germany

A. Macintyre

The Mathematical Institute, University of Oxford, UK

Volume 1

Linear Algebra and Geometry

A. I. Kostrikin and Yu. I. Manin

Volume 2

Model Theoretic Algebra: With Particular Emphasis on Fields,
Rings, Modules

Christian U. Jensen and Helmut Lenzing

Volume 3

Foundations of Module and Ring Theory: A Handbook for
Study and Research

Robert Wisbauer

Volume 4

Linear Representations of Partially Ordered Sets and
Vector Space Categories

Daniel Simson

Volume 5

Semantics of Programming Languages and Model Theory

M. Droste and Y. Gurevich

Volume 6

Exercises in Algebra: A Collection of Exercises in Algebra,
Linear Algebra and Geometry

Edited by *A. I. Kostrikin*

Volume 7

Bilinear Algebra: An Introduction to the Algebraic Theory
of Quadratic Forms

Kazimierz Szymiczek

Volume 8

Multilinear Algebra

Russell Merris

This book is part of a series. The publisher will accept continuation orders which may be cancelled at any time and which provide for automatic billing and shipping of each title in the series upon publication. Please write for details.

MULTILINEAR ALGEBRA

Russell Merris

California State University, Hayward
USA



CRC Press

Taylor & Francis Group

Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

Copyright © 1997 OPA (Overseas Publishers Association) Amsterdam B.V.
Published in The Netherlands under license by Gordon and Breach Science
Publishers.

All rights reserved.

No part of this book may be reproduced or utilized in any form or by any
means, electronic or mechanical, including photocopying and recording, or by
any information storage or retrieval system, without permission in writing from
the publisher.

Amsteldijk 166
1st Floor
1079 LH Amsterdam
The Netherlands

British Library Cataloguing in Publication Data

A Catalogue record for this book is available from the British Library

Contents

Preface	ix
1. Partitions	1
Applications to Graphs	13
Exercises	21
2. Inner Product Spaces	27
Application to Graphs	44
Exercises	47
3. Permutation Groups	53
Applications to Symmetry	64
Exercises	69
4. Group Representation Theory	75
Exercises	112
5. Tensor Spaces	121
Exercises	145
6. Symmetry Classes of Tensors	151
Applications to the Theory of Enumeration	186
The Enumeration of Graphs	193
Application to NMR Spectroscopy	197
Exercises	202
7. Generalized Matrix Functions	213
An Excursion into Invariant Theory	243
Applications to Graphs	251
Exercises	256

8. The Rational Representations of $GL(n, \mathbb{C})$	265
Application to Graphs	298
Exercises	300
Index of Notation	305
References	311
Index	325

Preface

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. The first four chapters of this book consist of self-contained introductions to a variety of prerequisite notions. Multilinear algebra, proper, begins in Chapter 5 with the development of the tensor product. Ironically, it is there, within sight of the goal, that one encounters what is perhaps the most formidable obstacle. In order to prevail over what Cartan has described as *une débauche d'indices*, one must slog through an obscuring foliage of superscripts and subscripts before reaching the heart, in Chapters 6 and 7, of this elegant and beautiful subject.

Many of the topics developed throughout the book are unified in the final chapter by means of the rational representations of the general linear group. Emerging as characters afforded by these representations, the classical Schur polynomials are one of the keys to the overall unification.

Throughout the book, some of the easier proofs are left to the exercises and some of the more difficult ones to the references. Apart from facilitating the flow of material, it is hoped this approach will encourage the reader to become a more active participant in exploring the subject.

Applications of multilinear algebra can be found in many areas of mathematics and physical science, some of them well beyond the author's interest or comprehension. Among those selected for inclusion in the book, graph theoretic applications are dominant. This does not reflect any particularly close connection between graph theory and multilinear algebra. However, applications to graphs suffice to give the flavor of more general combinatorial applications and, by keeping the focus on a single topic, one is able to probe a little deeper than might otherwise be possible.

Despite the book's broad scope, remarkably little prior experience is expected from the reader. It suffices to be familiar with the contents of the standard third year undergraduate courses in abstract and linear algebra. Ideally suited for a fourth year

'capstone' course, *Multilinear Algebra* is also an attractive choice for a beginning graduate course.

The book began as a series of handwritten lecture notes for an MPhil course at the Quaid-I-Azam University of Islamabad in 1973. A revised typescript was prepared later that same year for a seminar at the Instituto de Física e Matemática in Lisbon. These early versions were designed to supplement a series of lectures given to students whose native language was something other than English. Nevertheless, the lecture notes were circulated widely by the Institute for the Interdisciplinary Applications of Algebra and Combinatorics at the University of California, Santa Barbara. The present text is dedicated to the hearty folks who struggled through that primitive manuscript *without* the benefit of the author's lectures.

That multilinear algebra has flourished in the years since 1973 can be seen by browsing through the references. Much of this activity was stimulated by the appearance in that year of the first part of Marvin Marcus's monumental *Finite Dimensional Multilinear Algebra*. With the appearance of part II in 1975, *FDMA* became the standard reference, eclipsing the earlier classics of Bourbaki (1948) and Greub (1967), and overshadowing the compact treatises of Amir-Moez (1970s) and Oliveira (1973).

Among the individuals who have contributed to the author's scholarly research are José Dias da Silva, Amélia Fonseca, Bob Grone, Tom Pate, Steve Pierce, and Bill Watkins. He is also grateful for the professional competence of editors Donald Degenhardt, Katie Emblen, Matt Giarratano, Rebecca Stubbs and Brian Wyreweden.

CHAPTER 1

Partitions

The integer 6 is said to be “perfect” because it is the sum of its proper divisors: $6 = 1 + 2 + 3$. In this context, $1 + 2 + 3$ is the same as $2 + 3 + 1$ but different from $4 + 2$. In expressing the perfection of 6 what interests us is the *unordered* collection of its proper divisors, the “partition” of 6 whose “parts” are 3, 2, and 1.

DEFINITION 1.1 A **partition** of n of **length** m is an unordered collection of m positive integers that sum to n . The m summands are the **parts** of the partition.

NOTATION 1.2 A partition of n is typically represented by a sequence $\pi = [\pi_1, \pi_2, \dots, \pi_m]$, in which the parts of the partition are arranged so that $\pi_1 \geq \pi_2 \geq \dots \geq \pi_m > 0$. This convention is expressed by the shorthand notation $\pi \vdash n$. The length of π is denoted $L(\pi)$. In the present instance, $L(\pi) = m$.

In ordinary English usage, arranging the parts of a partition from largest to smallest would typically be called “ordering” the parts. This semantic difficulty can be the source of some confusion. It is precisely because a partition is unordered that we are free to arrange its parts any way we like.

EXAMPLES 1.3 The partitions of 5 are [5], [4,1], [3,2], [3,1,1], [2,2,1], [2,1,1,1], and [1,1,1,1,1]. The partitions of 6 having 3 parts are [4,1,1], [3,2,1], and [2,2,2].

□

Already, it seems convenient to introduce another shorthand notation. Rather than [3,1,1], [2,2,1], [2,1,1,1], and so on, we will write $[3,1^2]$, $[2^2,1]$ and $[2,1^3]$, respectively. The partition $[5,5,5,3,3,3,3,2,1,1]$ is abbreviated $[5^3,3^4,2,1^2]$. In this notation superscripts are used, not as exponents, but to denote multiplicities. In particular, $[5^3,3^4,2,1^2]$ is a 10-part partition of 31.



FIGURE 1.1

Partitions are frequently illustrated by means of so-called **Ferrers diagrams**.¹ If π is a partition of n having m parts, the corresponding Ferrers diagram, $F(\pi)$, consists of m rows of “boxes”. The number of boxes in row i of $F(\pi)$ is π_i . The Ferrers diagrams for $[6, 4, 3^2, 2]$ and $[5^2, 4, 2, 1^2]$ are illustrated in Figure 1.1.

DEFINITION 1.4. Suppose $\pi \vdash n$. The **conjugate** of π is the partition π^* whose j -th part is the number of boxes in *column* j of $F(\pi)$. (So, $F(\pi^*)$ is the transpose of $F(\pi)$.)

The conjugate of $[6, 4, 3^2, 2]$ is $[5^2, 4, 2, 1^2]$ as can easily be seen by glancing at Figure 1.1. The length of π^* is the largest part of π , that is, $L(\pi^*) = \pi_1$. Finally, the number of boxes in the j -th column of $F(\pi)$ is equal to the number of rows of $F(\pi)$ that contain at least j boxes, that is, to the number of parts of π that are bigger than or equal to j . In other words, the j -th part of π^* is

$$\pi_j^* = o(\{i : \pi_i \geq j\}), \quad (1.1)$$

where $o(S)$ denotes the **cardinality** of the set S .

EXAMPLE 1.5 The partition π is said to be **self conjugate** if $\pi = \pi^*$, that is, if $F(\pi)$ is symmetric. There is just one self conjugate partition of 6, namely, $[3, 2, 1]$. The self conjugate partitions of 9 are illustrated in Figure 1.2. $\square$



FIGURE 1.2 The self conjugate partitions of 9.

¹After Norman Macleod Ferrers (1829–1903).

Because $\pi_i \geq \pi_{i+1}$ for any $\pi \vdash n$,

$$\begin{aligned}\pi_i - i &\geq \pi_{i+1} - i \\ &> \pi_{i+1} - (i + 1).\end{aligned}$$

Thus, the integers $\pi_1 - \pi_i + i$, $1 \leq i \leq L(\pi)$, are all different, that is, $o(\{\pi_1 - \pi_i + i : 1 \leq i \leq L(\pi)\}) = L(\pi)$. Similarly, the cardinality of $\{\pi_1 + \pi_i^* - i + 1 : 1 \leq i \leq \pi_1\}$ is π_1 . What may not be so obvious is that these two sets are disjoint.

LEMMA 1.6 Suppose $\pi \vdash n$. Let $N = \pi_1 + L(\pi)$. Then $\{1, 2, \dots, N\}$ is the disjoint union of S and T , where $S = \{\pi_1 - \pi_i + i : 1 \leq i \leq L(\pi)\}$ and $T = \{\pi_1 + \pi_i^* - i + 1 : 1 \leq i \leq \pi_1\}$.

Proof It suffices to show that $S \cap T = \emptyset$. Observe that

$$\pi_1 - \pi_i + i = \pi_1 + \pi_j^* - j + 1,$$

if and only if $i + j - 1 = \pi_i + \pi_j^*$. To see that this is impossible, suppose first that $\pi_i \geq j$. Then, from Equation (1.1), $\pi_j^* = o(\{k : \pi_k \geq j\}) \geq i$, and $\pi_i + \pi_j^* \geq j + i > i + j - 1$. Therefore, we may assume $\pi_i \leq j - 1$, in which case, $\pi_j^* = o(\{k : \pi_k \geq j\}) < i$. But, $\pi_i \leq j - 1$ and $\pi_j^* \leq i - 1$ imply $\pi_i + \pi_j^* \leq i + j - 2 < i + j - 1$. $\square$

We now discuss "ordering" the different partitions of n .

DEFINITION 1.7 Let $(a) = (a_1, a_2, \dots, a_r)$ and $(b) = (b_1, b_2, \dots, b_s)$ be two sequences of real numbers satisfying $a_1 \geq a_2 \geq \dots \geq a_r \geq 0$ and $b_1 \geq b_2 \geq \dots \geq b_s \geq 0$. Then (a) **majorizes** (b) , written $(a) \succ (b)$, if

$$\sum_{i=1}^t a_i \geq \sum_{i=1}^t b_i, \quad 1 \leq t \leq r, \quad (1.2)$$

and

$$\sum_{i=1}^r a_i = \sum_{i=1}^s b_i. \quad (1.3)$$

EXAMPLE 1.8 Suppose n is a fixed positive integer. If $\pi \vdash n$, then $\pi = [\pi_1, \pi_2, \dots, \pi_m]$ is a nonincreasing sequence of positive real numbers. If $\rho = [\rho_1, \rho_2, \dots, \rho_k]$ is another partition of n , then $\pi_1 + \pi_2 + \dots + \pi_m = n = \rho_1 + \rho_2 + \dots + \rho_k$, and Equation (1.3) is satisfied automatically.

Suppose $n = 8$. If $\pi = [5, 2, 1]$, and $\rho = [3^2, 1^2]$, then $\pi > \rho$ because $5 \geq 3$, $5 + 2 \geq 3 + 3$, $5 + 2 + 1 \geq 3 + 3 + 1$, and $5 + 2 + 1 = 3 + 3 + 1 + 1$. If $\pi = [5, 2, 1]$ and $\rho = [4^2]$, then neither partition majorizes the other. Thus, majorization is a partial order. Figure 1.3 exhibits the “Hasse Diagram” for the partitions of 6 partially ordered by majorization. $\square$

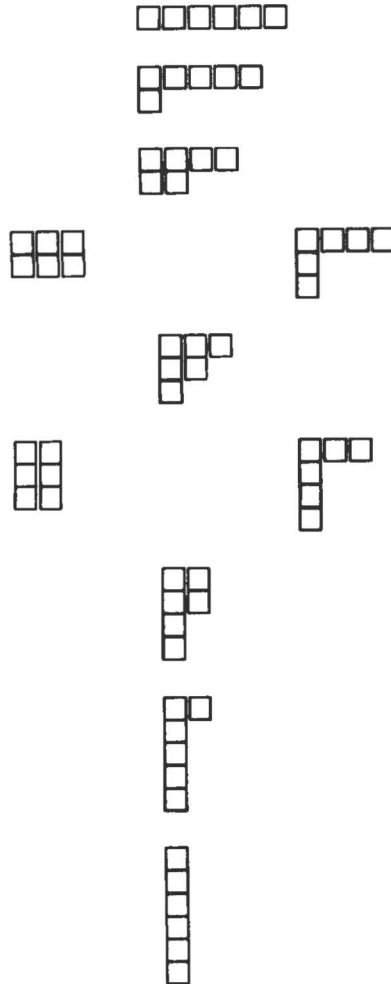


FIGURE 1.3 The partitions of 6 partially ordered by majorization.

DEFINITION 1.9 The n -by- n (entrywise) nonnegative matrix $S = (s_{ij})$ is **doubly stochastic** if its rows and columns all sum to 1, that is, if

$$\sum_{j=1}^n s_{ij} = 1, \quad 1 \leq i \leq n, \quad \text{and} \quad \sum_{i=1}^n s_{ij} = 1, \quad 1 \leq j \leq n.$$

THEOREM 1.10 Let $(a) = (a_1, a_2, \dots, a_n)$ and $(b) = (b_1, b_2, \dots, b_n)$ be two sequences of real numbers satisfying $a_1 \geq a_2 \geq \dots \geq a_n \geq 0$ and $b_1 \geq b_2 \geq \dots \geq b_n \geq 0$. Then (a) majorizes (b) if and only if there is a doubly stochastic matrix S such that $(b) = (a)S$.

Theorem 1.10 is stated for the case in which both sequences have the same length. Because adding zeros to the end of the shorter sequence does not affect majorization, this hypothesis does not impose any real restriction. A proof can be found in [Hardy, Littlewood & Pólya (1967), pp. 47–49] or [Marshall & Olkin (1979), p. 22].

EXAMPLE 1.11 We saw in Example 1.8 that $[5, 2, 1] \succ [3^2, 1^2]$. As an illustration of Theorem 1.10, observe that $(3, 3, 1, 1) = (5, 2, 1, 0)S$, where

$$S = \frac{1}{6} \begin{pmatrix} 2 & 3 & 1 & 0 \\ 4 & 0 & 0 & 2 \\ 0 & 3 & 1 & 2 \\ 0 & 0 & 4 & 2 \end{pmatrix}.$$

If S is an n -by- n doubly stochastic matrix then [Birkhoff (1946)] there exist permutation matrices $P_1, P_2, \dots, P_k$ and positive real numbers $\theta_1, \theta_2, \dots, \theta_k$ such that $\theta_1 + \theta_2 + \dots + \theta_k = 1$ and

$$S = \theta_1 P_1 + \theta_2 P_2 + \dots + \theta_k P_k.$$

In other words, S is a **convex combination** (or “weighted average”) of permutation matrices. Using these terms, Theorem 1.10 can be restated as follows: (a) majorizes (b) if and only if (b) is a convex combination of rearrangements of (a) . In particular,

$$\begin{aligned} (3, 3, 1, 1) &= \frac{1}{3}(2, 5, 0, 1) + \frac{1}{3}(5, 1, 0, 2) \\ &\quad + \frac{1}{6}(2, 5, 1, 0) + \frac{1}{6}(2, 1, 5, 0). \end{aligned}$$

□

Apart from their intrinsic interest, the partitions of n have a variety of uses, one of which involves symmetric polynomials.

DEFINITION 1.12 A polynomial $f(x_1, x_2, \dots, x_k)$ is **symmetric** in $x_1, x_2, \dots, x_k$ if its value is unchanged by any permutation of the k variables, that is, if $f(x_1, x_2, \dots, x_k) = f(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(k)})$, for every permutation σ of $\{1, 2, \dots, k\}$.²

Perhaps the most natural way to begin a discussion of symmetric polynomials is with the notorious “multinomial theorem”.

THE MULTINOMIAL THEOREM 1.13 *If n is a positive integer, then*

$$(x_1 + x_2 + \dots + x_k)^n = \sum \binom{n}{r_1, r_2, \dots, r_k} x_1^{r_1} x_2^{r_2} \dots x_k^{r_k}, \quad (1.4)$$

where the sum is over all nonnegative integer sequences, $(r_1, r_2, \dots, r_k)$, satisfying $r_1 + r_2 + \dots + r_k = n$, and

$$\binom{n}{r_1, r_2, \dots, r_k} = \frac{n!}{r_1! r_2! \dots r_k!}$$

is the corresponding **multinomial coefficient**.

Proofs can be found in any of the standard books on combinatorics.³

EXAMPLE 1.14 The coefficient of $b^4 c^2$ in $(a + b + c)^6$ is

$$\binom{6}{0, 4, 2} = \frac{6!}{0!4!2!} = \frac{6!}{4!2!} = 15.$$

Because $(a + b + c)^6$ is symmetric in a, b , and c , the coefficients of $a^4 b^2$ and $a^2 c^4$ in $(a + b + c)^6$ must be 15 as well. One “piece” of the multinomial expansion of $(a + b + c)^6$ is $15p(x)$, where

$$p(x) = a^4 b^2 + a^4 c^2 + a^2 b^4 + a^2 c^4 + b^4 c^2 + b^2 c^4. \quad (1.5)$$

□

DEFINITION 1.15 Let k and n be positive integers, and π be a partition of n of length $m \leq k$. The **monomial symmetric function**

$$M_\pi(x_1, x_2, \dots, x_k) = \sum x_1^{r_1} x_2^{r_2} \dots x_k^{r_k}, \quad (1.6)$$

where the sum is over all *different* rearrangements, $(r_1, r_2, \dots, r_k)$, of the k -tuple $(\pi_1, \pi_2, \dots, \pi_m, 0, 0, \dots, 0)$, obtained by appending $k - m$ zeros to the end of π . If $m > k$, then $M_\pi(x_1, x_2, \dots, x_k) = 0$.

²This is why the group of all permutations of $\{1, 2, \dots, k\}$ is called the “symmetric” group.

³See, for example, [Merris (1996)].

If $m = 2, k = 3$, and $\pi = [\pi_1, \pi_2] = [2, 2]$, then the “different rearrangements” of $(2, 2, 0)$ are

$$(2, 2, 0), \quad (2, 0, 2), \quad \text{and} \quad (0, 2, 2),$$

not the six rearrangements of the different looking symbols π_1, π_2 , and 0. Thus,

$$M_{[2,2]}(x, y, z) = x^2y^2 + x^2z^2 + y^2z^2.$$

The “piece” of the multinomial expansion of $(a + b + c)^6$ exhibited in Equation (1.5) is

$$M_{[4,2]}(a, b, c) = a^4b^2 + a^4c^2 + a^2b^4 + a^2c^4 + b^4c^2 + b^2c^4.$$

Any symmetric polynomial is a linear combination of minimally symmetric pieces, namely, the monomial symmetric functions. We shall have more to say about this presently.

EXAMPLE 1.16 There are exactly seven partitions of 6 having three or fewer parts. So, there are seven (nonzero) monomial symmetric functions of degree 6 in the three variables a, b , and c . They are

$$M_{[6]}(a, b, c) = a^6 + b^6 + c^6,$$

$$M_{[5,1]}(a, b, c) = a^5b + a^5c + ab^5 + ac^5 + b^5c + bc^5,$$

$$M_{[4,2]}(a, b, c) = a^4b^2 + a^4c^2 + a^2b^4 + a^2c^4 + b^4c^2 + b^2c^4,$$

$$M_{[3^2]}(a, b, c) = a^3b^3 + a^3c^3 + b^3c^3,$$

$$M_{[4,1^2]}(a, b, c) = a^4bc + ab^4c + abc^4,$$

$$M_{[3,2,1]}(a, b, c) = a^3b^2c + a^3bc^2 + a^2b^3c + a^2bc^3 + ab^3c^2 + ab^2c^3,$$

and

$$M_{[2^3]}(a, b, c) = a^2b^2c^2.$$

Setting $M_\pi = M_\pi(a, b, c)$ we obtain, from the multinomial theorem, that

$$\begin{aligned} (a + b + c)^6 &= M_{[6]} + 6M_{[5,1]} + 15M_{[4,2]} + 20M_{[3^2]} \\ &\quad + 30M_{[4,1^2]} + 60M_{[3,2,1]} + 90M_{[2^3]}. \end{aligned}$$

□

If $\pi = [\pi_1, \pi_2, \dots, \pi_m]$ is some fixed but arbitrary partition of n , define

$$\binom{n}{\pi} = \frac{n!}{\pi_1! \pi_2! \dots \pi_m!}.$$

Using this notation, the multinomial theorem can be restated as follows:

THEOREM 1.17 *If n is a positive integer, then*

$$(x_1 + x_2 + \dots + x_k)^n = \sum_{\pi \vdash n} \binom{n}{\pi} M_{\pi}(x_1, x_2, \dots, x_k). \quad (1.7)$$

We now give special names to the two “extreme” monomial symmetric functions, the ones corresponding to the partitions $[n]$ and $[1^n]$.

NOTATION 1.18 Let $P_n(x_1, x_2, \dots, x_k) = M_{[n]}(x_1, x_2, \dots, x_k)$ and $E_n(x_1, x_2, \dots, x_k) = M_{[1^n]}(x_1, x_2, \dots, x_k)$.

It is easy to recognize P_n ; it is the n -th power sum,

$$P_n(x_1, x_2, \dots, x_k) = x_1^n + x_2^n + \dots + x_k^n.$$

What about E_n ?

EXAMPLE 1.19 Let's choose $k = 4$. Then

$$E_1(a, b, c, d) = M_{[1]}(a, b, c, d) = a + b + c + d;$$

$$E_2(a, b, c, d) = M_{[1^2]}(a, b, c, d) = ab + ac + ad + bc + bd + cd;$$

$$E_3(a, b, c, d) = M_{[1^3]}(a, b, c, d) = abc + abd + acd + bcd; \quad \text{and}$$

$$E_4(a, b, c, d) = M_{[1^4]}(a, b, c, d) = abcd.$$

Evidently, $E_n(a, b, c, d)$ is the sum of all $C(4, n)$ (binomial coefficient $\binom{4}{n}$) products of the x 's taken n at a time. $\square$

If $(r_1, r_2, \dots, r_k)$ is some rearrangement of the sequence $(1, 1, \dots, 1, 0, 0, \dots, 0)$ consisting of n ones followed by $k - n$ zeros, then

$$x_1^{r_1} x_2^{r_2} \dots x_k^{r_k} = x_{i_1} x_{i_2} \dots x_{i_n},$$

where $i_1 < i_2 < \dots < i_n$. Summing over the different rearrangements gives

$$E_n(x_1, x_2, \dots, x_k) = \sum x_{i_1} x_{i_2} \dots x_{i_n}, \quad (1.8)$$

where the summation is over all $C(k, n)$ sequences $(i_1, i_2, \dots, i_n)$ satisfying $1 \leq i_1 < i_2 < \dots < i_n \leq k$.

DEFINITION 1.20 Denote by $\Gamma_{n,k}$ the set of all functions from $\{1, 2, \dots, n\}$ into $\{1, 2, \dots, k\}$. Let $Q_{n,k}$ be the subset of $\Gamma_{n,k}$ consisting of the $C(k, n)$ strictly increasing functions.

There is a natural one-to-one correspondence between the functions $\beta \in \Gamma_{n,k}$ and the integer sequences $(i_1, i_2, \dots, i_n)$ satisfying $1 \leq i_t \leq k$, $1 \leq t \leq n$, namely, $\beta \leftrightarrow (\beta(1), \beta(2), \dots, \beta(n))$. We will feel free to abuse the language by identifying $\Gamma_{n,k}$ with a set of sequences. Thus,

$$\Gamma_{2,3} = \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3)\},$$

and $Q_{2,3} = \{(1, 2), (1, 3), (2, 3)\}$.

Using Definition 1.20, we may rewrite Equation (1.8) as

$$E_n(x_1, x_2, \dots, x_k) = \sum_{\beta \in Q_{n,k}} x_{\beta(1)} x_{\beta(2)} \dots x_{\beta(n)}. \quad (1.9)$$

DEFINITION 1.21 The “extreme” monomial symmetric function, $E_n(x_1, x_2, \dots, x_k)$ is the n -th elementary symmetric function of $x_1, x_2, \dots, x_k$. It is useful to define $E_0(x_1, x_2, \dots, x_k) = 1$.

Elementary symmetric functions are familiar objects. They express the coefficients of a monic polynomial in terms of its roots. If, for example, a, b, c , and d are complex numbers, then (Example 1.19)

$$(x - a)(x - b)(x - c)(x - d) = x^4 - E_1x^3 + E_2x^2 - E_3x + E_4, \quad (1.10)$$

where $E_n = E_n(a, b, c, d)$, $1 \leq n \leq 4$.

FUNDAMENTAL THEOREM OF SYMMETRIC FUNCTIONS 1.22 Any polynomial, symmetric in the variables $x_1, x_2, \dots, x_k$, is a polynomial in the elementary symmetric functions $E_n(x_1, x_2, \dots, x_k)$, $1 \leq n \leq k$.

Proof Let $f = f(x_1, x_2, \dots, x_k)$ be a symmetric polynomial of (total) degree p . Write $f = f_0 + f_1 + \dots + f_p$, where $f_i = f_i(x_1, x_2, \dots, x_k)$ is the (homogeneous) part of f consisting of all terms of degree i . It will suffice to show that f_i is a polynomial in the elementary symmetric functions for a fixed but arbitrary i .

Suppose

$$cx_1^{r_1} x_2^{r_2} \dots x_k^{r_k} \quad (1.11)$$

is one of the monomial terms that occur in f_i . Then $r_1 + r_2 + \dots + r_k = i$. By symmetry, we may assume that

$$r_1 \geq r_2 \geq \dots \geq r_t > 0 = r_{t+1} = \dots = r_k.$$

Among all partitions of i occurring as the sequence of exponents in the monomials of f_i , assume $[r_1, r_2, \dots, r_t]$ is last in **lexicographic** (dictionary) order. That is, without loss of generality, we may assume r_1 is the largest single exponent that occurs in any monomial in f_i ; r_2 is the maximum second largest exponent among all the monomials that occur in f_i and have r_1 as their largest exponent; r_3 is the maximum third largest exponent among all the monomials that occur in f_i and have r_1 and r_2 as their two largest exponents; and so on.

Consider

$$E_1^{s_1} E_2^{s_2} \dots E_k^{s_k}, \quad (1.12)$$

where $E_n = E_n(x_1, x_2, \dots, x_k)$, $1 \leq n \leq k$. In lexicographic order of its *exponents* the last monomial that occurs in (1.12) is

$$x_1^{s_1} (x_1 x_2)^{s_2} (x_1 x_2 x_3)^{s_3} \dots (x_1 x_2 \dots x_k)^{s_k}. \quad (1.13)$$

We would like to choose s_1, s_2, s_3 , and so on, so that

$$x_1^{s_1} (x_1 x_2)^{s_2} (x_1 x_2 x_3)^{s_3} \dots (x_1 x_2 \dots x_k)^{s_k} = x_1^{r_1} x_2^{r_2} \dots x_k^{r_k}.$$

This requires that

$$\begin{aligned} r_1 &= s_1 + s_2 + s_3 + \dots + s_k, \\ r_2 &= s_2 + s_3 + \dots + s_k, \\ r_3 &= s_3 + \dots + s_k, \\ &\dots \\ r_k &= s_k. \end{aligned}$$

These equations are satisfied when $s_k = r_k$, $s_{k-1} = r_{k-1} - r_k$, $\dots$, $s_2 = r_2 - r_3$, and $s_1 = r_1 - r_2$. If we make these choices, then either

$$f_i - c E_1^{s_1} E_2^{s_2} \dots E_k^{s_k} = 0,$$

or it is a symmetric homogeneous polynomial of degree i , each of whose monomial terms comes before (1.11) in lexicographic order. Because dictionary ordering is a total order, the result follows by induction. $\square$

Suppose $f = f(x_1, x_2, \dots, x_k)$ is a symmetric homogeneous polynomial of degree n . Then f is, simultaneously, a polynomial in the elementary symmetric functions $E_n(x_1, x_2, \dots, x_k)$, $1 \leq n \leq k$, and a linear combination of the monomial symmetric functions $M_\pi(x_1, x_2, \dots, x_k)$, $\pi \vdash n$. Conversely, if c_π , $\pi \vdash n$, are constants, then

$$g(x_1, x_2, \dots, x_k) = \sum_{\pi \vdash n} c_\pi M_\pi(x_1, x_2, \dots, x_k)$$

defines a symmetric homogeneous polynomial of degree n . If

$$c_\pi = \binom{n}{\pi}$$

for all π , then $g(x_1, x_2, \dots, x_k) = (x_1 + x_2 + \dots + x_k)^n$. What about some other choices? An important and interesting example arises when $c_\pi = 1$ for all π .

DEFINITION 1.23 Let $x_1, x_2, \dots, x_k$ be independent variables. Their n -th **homogeneous symmetric function** is defined by

$$H_n(x_1, x_2, \dots, x_k) = \sum_{\pi \vdash n} M_\pi(x_1, x_2, \dots, x_k). \quad (1.14)$$

It is convenient to define $H_0(x_1, x_2, \dots, x_k) = 1$.

EXAMPLES 1.24

$$\begin{aligned} H_2(a, b, c) &= M_{[2]}(a, b, c) + M_{[1^2]}(a, b, c) \\ &= P_2(a, b, c) + E_2(a, b, c) \\ &= (a^2 + b^2 + c^2) + (ab + ac + bc), \end{aligned}$$

$$\begin{aligned} H_3(a, b, c) &= M_{[3]}(a, b, c) + M_{[2,1]}(a, b, c) + M_{[1^3]}(a, b, c) \\ &= (a^3 + b^3 + c^3) + (a^2b + a^2c + ab^2 + ac^2 + b^2c + bc^2) + abc, \end{aligned} \quad (1.15)$$

and

$$\begin{aligned} H_4(a, b, c) &= M_{[4]}(a, b, c) + M_{[3,1]}(a, b, c) + M_{[2^2]}(a, b, c) + M_{[2,1^2]}(a, b, c) \\ &= (a^4 + b^4 + c^4) + (a^3b + a^3c + ab^3 + ac^3 + b^3c + bc^3) \\ &\quad + (a^2b^2 + a^2c^2 + b^2c^2) + (a^2bc + ab^2c + abc^2). \end{aligned}$$

□

From the definition, each monomial of (total) degree n in the variables $x_1, x_2, \dots, x_k$ occurs in $H_n(x_1, x_2, \dots, x_k)$ exactly once. This leads to a formula for $H_n(x_1, x_2, \dots, x_k)$ analogous to Equation (1.9).

DEFINITION 1.25 Denote by $G_{n,k}$ the subset of $\Gamma_{n,k}$ consisting of all $C(n+k-1, n)$ nondecreasing functions from $\{1, 2, \dots, n\}$ into $\{1, 2, \dots, k\}$.

For all n and k , $Q_{n,k} \subset G_{n,k}$. As (lexicographically ordered) sequence sets, $G_{2,3} = \{(1, 1), (1, 2), (1, 3), (2, 2), (2, 3), (3, 3)\}$, and

$$G_{3,3} = \{(1, 1, 1), (1, 1, 2), (1, 1, 3), (1, 2, 2), (1, 2, 3), (1, 3, 3), \\ (2, 2, 2), (2, 2, 3), (2, 3, 3), (3, 3, 3)\}. \quad (1.16)$$

Using Definition 1.25, we can rewrite Equation (1.14) as

$$H_n(x_1, x_2, \dots, x_k) = \sum_{\beta \in G_{n,k}} x_{\beta(1)} x_{\beta(2)} \dots x_{\beta(n)}. \quad (1.17)$$

We now return to the observation that any symmetric polynomial is a linear combination of “minimally symmetric pieces”.

DEFINITION 1.26 Suppose $x_1, x_2, \dots, x_k$ are independent indeterminates (variables) over the field $\mathbb{C}$ of complex numbers. Denote by $\mathbb{C}[x_1, x_2, \dots, x_k]$ the set of polynomials in $x_1, x_2, \dots, x_k$ with complex coefficients. Let $SC_n[x_1, x_2, \dots, x_k]$ be the subset of $\mathbb{C}[x_1, x_2, \dots, x_k]$ consisting of the zero polynomial together with all symmetric homogeneous polynomials of degree n .

THEOREM 1.27 The set $\{M_\pi(x_1, x_2, \dots, x_k) : \pi \vdash n, L(\pi) \leq k\}$ is a basis of the vector space $SC_n[x_1, x_2, \dots, x_k]$.

Proof Let $M_\pi = M_\pi(x_1, x_2, \dots, x_k)$, $\pi \vdash n$. The only thing remaining to be proved is the linear independence of $\{M_\pi : \pi \vdash n, L(\pi) \leq k\}$.

Suppose

$$\sum_{\substack{\pi \vdash n \\ L(\pi) \leq k}} c_\pi M_\pi = 0, \quad (1.18)$$

the zero polynomial. Let $\rho = [\rho_1, \rho_2, \dots, \rho_r]$, $r \leq k$, be a partition of n . Consider the term

$$c_\rho x_1^{\rho_1} x_2^{\rho_2} \dots x_r^{\rho_r}$$

occurring in Equation (1.18). Taking partial derivatives of (1.18) with respect to x_1 , ρ_1 -times, with respect to x_2 , ρ_2 -times, $\dots$, and with respect to x_r , ρ_r -times, we deduce that

$$\rho_1! \rho_2! \dots \rho_r! c_\rho = 0.$$

□

Application to Graphs

Let V be a set. Denote the family of its 2-element subsets by $V^{(2)}$. Then, for example,

$$\{a, b, c\}^{(2)} = \{\{a, b\}, \{a, c\}, \{b, c\}\};$$

$$\{1, 2, 3, 4\}^{(2)} = \{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \{3, 4\}\};$$

and $\{x, y\}^{(2)} = \{\{x, y\}\}$. If $o(V) = n$, then $o(V^{(2)}) = C(n, 2)$.

DEFINITION 1.28 A **graph** consists of two things, a nonempty finite set V , and a (possibly empty) subset E of $V^{(2)}$. If $G = (V, E)$ is a graph, the elements of V are its **vertices** and the elements of E its **edges**. When more than one graph is under consideration, it may be useful to write $V(G)$ and $E(G)$, respectively, for the sets of vertices and edges. If $e = \{u, v\} \in E(G)$, then u and v are **adjacent vertices**, **incident with e** . Two edges are **adjacent** if their set-theoretic intersection consists of a single vertex.

EXAMPLE 1.29 If $V = \{1, 2, 3, 4, 5\}$, then $V^{(2)}$ has 10 elements and 2^{10} subsets. Hence, there are 1024 different graphs with vertex set $\{1, 2, 3, 4, 5\}$.

It is common to draw pictures of graphs in which vertices are represented by points and points representing adjacent vertices are joined by line segments (or arcs). If $E = \{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}\}$, then each of the pictures in Figure 1.4 illustrates $H = (V, E)$. Note that H is not connected; vertex 5 is an “isolated” vertex. $\square$



FIGURE 1.4 Pictures of graph H .

EXAMPLE 1.30 Not only can one graph be illustrated by different pictures, but one picture can represent different graphs! If $W = \{p, q, r, s, t\}$ and $F = \{\{q, s\}, \{q, t\}, \{r, s\}, \{r, t\}, \{s, t\}\}$, then the four pictures in Figure 1.4 also illustrate $K = (W, F)$. $\square$

We are not so much interested in different graphs as in nonisomorphic graphs.

DEFINITION 1.31 Let $G_1 = (V, E)$ and $G_2 = (W, F)$ be graphs. Then G_1 is **isomorphic** to G_2 if there is a one-to-one function $f: V \rightarrow W$ such that vertices u and v are adjacent in G_1 if and only if $f(u)$ and $f(v)$ are adjacent in G_2 , that is, such that $\{u, v\} \in E$ if and only if $\{f(u), f(v)\} \in F$. The function f is an **isomorphism** from G_1 onto G_2 .

If G_1 and G_2 can be illustrated by the same picture, then they are isomorphic. To each point of the picture there corresponds a unique vertex v_1 of G_1 and a unique vertex v_2 of G_2 . The function that sends v_1 to v_2 (for every point of the picture) is an isomorphism. It is more challenging to tell when graphs illustrated by different pictures are isomorphic.

EXAMPLE 1.32 The so-called “Petersen” graph, G_1 , is illustrated in Figure 1.5. It is isomorphic to the graph G_2 , pictured in the same figure. The proof that G_1 and G_2 are isomorphic is “by the numbers”. If $V(G_1) = \{0, 1, 2, \dots, 9\} = V(G_2)$, then $f(i) = i$, $0 \leq i \leq 9$, is an isomorphism. (Check it out: Confirm that i and j are adjacent in G_1 if and only if they are adjacent in G_2 .) Such a pair of labeled figures may be considered a proof of isomorphism (provided, of course, that it “checks out”). $\square$

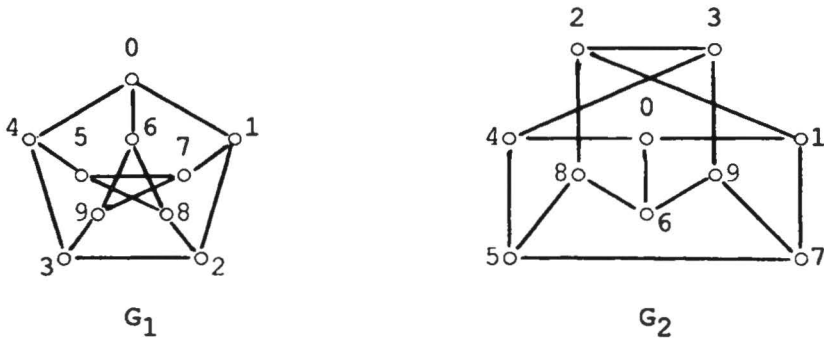


FIGURE 1.5 The Petersen graph.

It is an immediate consequence of the definition that isomorphic graphs have the same numbers of vertices and edges. Consequently, if G_1 and G_2 do not share these properties, they cannot be isomorphic. Properties that isomorphic graphs must share, are called **graph invariants**. We now introduce another graph invariant.

DEFINITION 1.33 Let $G = (V, E)$ be a graph with vertex set $V = \{v_1, v_2, \dots, v_n\}$. The **degree** of $v \in V$, denoted $d(v)$, is the number of edges of G that are incident with v (which is equal to the number of vertices of G that are adjacent to v). When more than one graph is under consideration, it may be useful to write $d(v) = d_G(v)$.

The **degree sequence** is $d(G) = (d_1, d_2, \dots, d_n)$, where $d_1 \geq d_2 \geq \dots \geq d_n \geq 0$ are the degrees of the vertices of G , arranged in nonincreasing order. (We are not necessarily assuming that $d_i = d(v_i)$.)

THEOREM 1.34 *The degree sequence is a graph invariant.*

We can determine from $d(G)$ both n , the number of vertices of G , and m , the number of its edges: n is just the length of the sequence $d(G)$, and m is given by what has come to be known as the “first theorem” of graph theory.

THEOREM 1.35 *Let $G = (V, E)$ be a graph with vertex set $V = \{v_1, v_2, \dots, v_n\}$. If $o(E) = m$, then*

$$\sum_{i=1}^n d(v_i) = 2m.$$

Proof By definition, $d(v)$ is the number of edges incident with vertex v . Thus, in summing the vertex degrees, each edge is counted twice, once at each of its vertices. $\square$



FIGURE 1.6 Nonisomorphic graphs with the same degree sequence.

EXAMPLE 1.36 The nonisomorphic graphs G_1 and G_2 in Figure 1.6 share the degree sequence $(2, 2, 2, 1, 1)$. $\square$

If G is a graph with n vertices and m edges, it follows from Theorem 1.35 that, were it not for isolated vertices (of degree 0), $d(G) = (d_1, d_2, \dots, d_n)$ would be a partition of $2m$. When speaking of the Ferrers diagram of $d(G)$ it will be understood that vertices of degree 0 go unrepresented. Similarly, let $d_j^* = o(\{i : d_i \geq j\})$. Then the **conjugate degree sequence**, $d^*(G) = (d_1^*, d_2^*, \dots, d_k^*)$, is the conjugate of the partition of $2m$ whose parts are the nonzero vertex degrees of G .

THEOREM 1.37 *Let G be a graph with n vertices, m edges, and degree sequence $d(G)$. Then $d^*(G)$ majorizes $d(G)$.*

DEFINITION 1.39 The trace of partition π is $f(\pi) = o(\{i : \pi_i \geq i\})$.

If $F(\pi)$ is the Ferrers diagram corresponding to π , then $f(\pi)$ is the length of its main diagonal.

THEOREM 1.40⁴ Suppose $\pi = [\pi_1, \pi_2, \dots, \pi_n]$ is a partition of the positive integer $2m$. Let $\pi^* = [\pi_1^*, \pi_2^*, \dots, \pi_k^*]$ be its conjugate partition. Then π is graphic if and only if

$$\sum_{i=1}^r \pi_i^* \geq \sum_{i=1}^r (\pi_i + 1), \quad 1 \leq r \leq f(\pi). \quad (1.19)$$

Proof The proof uses the same variations, (a) and (b), of the Ferrers diagram of $d(G)$ that were useful in the proof of Theorem 1.37. (See Figure 1.8.) Because no vertex is adjacent to itself, no row in variation (b) contains its own number. In particular, the (1,1)-entry is not less than 2. Therefore, in addition to all the 1's, the first column of variation (b) contains a number larger than 1, so $d_1^* \geq d_1 + 1$.

Since the (1,1)-entry of variation (b) is at least 2, and since the numbers in the first row are strictly increasing, the (1,2)-entry must be at least 3. If $d_2 \geq 2$ then, because the second vertex is not adjacent to itself, the (2,2)-entry can be no less than 3 as well. Therefore, all the 1's, all the 2's, and at least two numbers no smaller than 3 occur in the first two columns of variation (b). This means $d_1^* + d_2^* \geq d_1 + d_2 + 2 = (d_1 + 1) + (d_2 + 1)$. As long as $d_r \geq r$, we can use the same argument to prove that

$$\sum_{i=1}^r d_i^* \geq \left(\sum_{i=1}^r d_i \right) + r,$$

thus establishing the necessity of Condition (1.19).

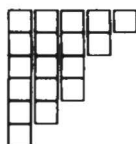
To prove sufficiency, suppose $\pi = [\pi_1, \pi_2, \dots, \pi_n]$ is a partition of $2m$ that satisfies Inequalities (1.19). Consider the extreme case in which $\pi_i^* = \pi_i + 1$, $1 \leq i \leq f(\pi)$. To produce a graph with degree sequence π , begin with the vertex set $V = \{1, 2, \dots, n\}$. "Construct" edges from vertex 1 to each of $2, 3, \dots, n$. Because $n = \pi_1^* = \pi_1 + 1$, we have produced a graph in which $d_1 = \pi_1$, and $d_2 = d_3 = \dots = d_n = 1$. If $f(\pi) = 1$, then $\pi_2 = 1$, and we are finished. Otherwise, construct edges from vertex 2 to each of $3, 4, \dots, \pi_2 + 1$. (This is possible because we have reserved "room" for $\pi_2^* = \pi_2 + 1$ vertices of degree 2 or more.) So far, we have produced a graph in which $d_1 = \pi_1$ and $d_2 = \pi_2$. If $f(\pi) = 2$, we are finished because $F(\pi)$ is completely determined by its first $f(\pi)$ rows and columns. If $\pi_3 \geq 3$, draw edges from vertex 3 to each of $4, 5, \dots, \pi_3 + 1$ (which is possible because $\pi_3^* = \pi_3 + 1$). After three steps, we have $d_1 = \pi_1$,

⁴While this result has been attributed to Hasselbarth [Sierksma & Hoogeveen (1991)], it seems to have been published first in [Ruch & Gutman (1979)].

$d_2 = \pi_2$, and $d_3 = \pi_3$. At the end of $f(\pi)$ steps, we will have produced a graph satisfying $d_i = \pi_i$, $1 \leq i \leq n$.

To complete the proof of sufficiency, two additional facts are required: (1) if ρ is majorized by a graphic partition π , then ρ is graphic; and (2) every partition satisfying the inequalities in (1.19) is majorized by one which is extreme in the sense that equality holds in each of the inequalities. The details are omitted. $\square$

EXAMPLE 1.41 Consider the partition $\pi = [5, 4, 3, 3, 2, 1]$, whose Ferrers diagram, $F(\pi)$ appears in Figure 1.9. Because π is a partition of 18, the first condition of Theorem 1.40 is satisfied: $m = 9$. In this case, the length of the main diagonal of $F(\pi)$ is $3 = f(\pi)$. Glancing at Figure 1.9, we can write down $\pi^* = [6, 5, 4, 2, 1]$. Observe that $\pi_i^* = \pi_i + 1$, for $i = 1, 2, 3$.



$F([5, 4, 3^2, 2, 1])$

FIGURE 1.9

Draw six points in the plane and label them 1, 2, ..., 6. Construct (draw) edges from vertex 1 to vertices 2, 3, 4, 5, and 6, as shown in Figure 1.10(a). This gives a vertex of degree 5 and five vertices of degree 1. Now, draw edges connecting vertex 2 to vertices 3, 4, and 5. Finally, drawing an edge from vertex 3 to vertex 4, one obtains the graph G , illustrated in Figure 1.10(b), having degree sequence $d(G) = \pi$. $\square$

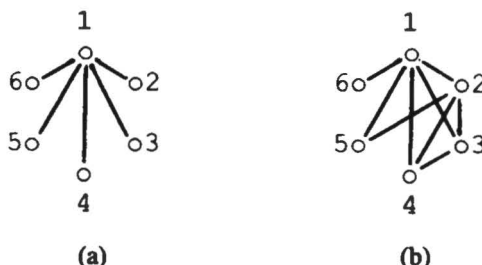


FIGURE 1.10

Example 1.41 illustrates the “greedy” algorithm used in the proof of Theorem 1.40 to construct a graph whose degree sequence is extreme in the sense that equality holds in each of the inequalities in (1.19). We now give a formal name to the graphic partitions that are extreme in this sense.

DEFINITION 1.42 Let $\pi = [\pi_1, \pi_2, \dots, \pi_n]$ be a partition of $2m$. Then π is a **maximal** (graphic) partition if $\pi_i^* = \pi_i + 1$, $1 \leq i \leq f(\pi)$. A graph whose degree sequence is maximal is a **threshold graph**.

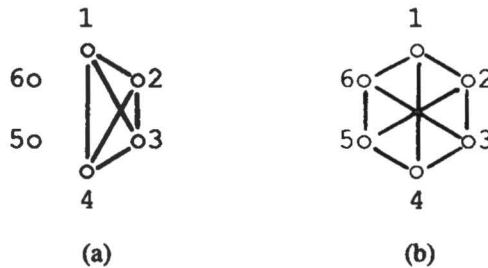


FIGURE 1.11

EXAMPLE 1.43 Let $\pi = [3^6]$. Then $\pi^* = [6^3]$ and, while π is graphic, it is not maximal. Let's see what happens if we try to use the greedy algorithm illustrated in Example 1.41 to construct a graph with degree sequence $[3^6]$. Begin by drawing six points in the plane and labeling them 1, 2, ..., 6. Draw edges from vertex 1 to vertices 2, 3, and 4. Now draw edges from vertex 2 to vertices 3 and 4, producing two vertices of degree 3, two of degree 2, and two of degree 0. When an edge is drawn between vertices 3 and 4, we find ourselves in the position illustrated in Figure 1.11(a). Pretty clearly, a graph with degree sequence $[3^6]$ cannot be obtained from this figure by adding more edges. On the other hand, the existence of a graph with degree sequence $[3^6]$ is established by Figure 1.11(b). $\square$

EXAMPLE 1.44 The connected threshold graphs having $2 \leq m \leq 6$ edges are illustrated in Figure 1.12.

DEFINITION 1.45 Let V be an n -element set. The **complete graph** $K_n = (V, V^{(2)})$ is the graph in which every pair of vertices is adjacent.

Strictly speaking, Definition 1.45 defines the complete graph with vertex set V . However, because any two complete graphs on n vertices are isomorphic, we will abuse the language and speak about *the* complete graph on n vertices. The complete graphs K_3 and K_4 are illustrated in Figure 1.12(b) and (h), respectively.

DEFINITION 1.46 Let $G = (V, E)$ be a graph. The **complement** of G is the graph $G^c = (V, V^{(2)} \setminus E)$.

If G is a graph, then $e = \{u, v\}$ is an edge of G if and only if e is not an edge of G^c . In particular, the complement of K_n is the graph consisting of n isolated vertices, that is, K_n^c has no edges at all.

DEFINITION 1.47 Let $G = (V, E)$ be a graph. A **cycle** in G is a sequence of distinct vertices $v_1, v_2, \dots, v_n, n > 2$, such that $\{v_i, v_{i+1}\} \in E, 1 \leq i < n$, and $\{v_1, v_n\} \in E$. A connected graph without cycle is a **tree**.

Graphs (a), (b), (e), (h), and (i) in Figure 1.12 are trees.

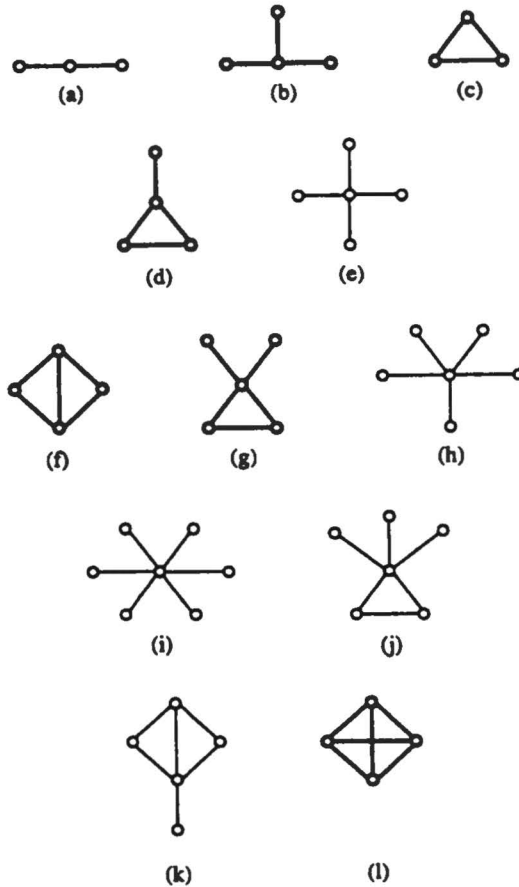


FIGURE 1.12 The Threshold Graphs with $2 \leq m \leq 6$ edges.

Exercises

1. Denote by $p_m(n)$ the number of partitions of n having m parts. Show that
 - a. $p_{n-2}(n) = 2, n \geq 4$.
 - b. $p_{n-3}(n) = 3, n \geq 6$.
 - c. $p_2(n) = \lfloor n/2 \rfloor$, the greatest integer not exceeding $n/2$.
 - d. $p_m(n) = p_{m-1}(n-1) + p_m(n-m), 1 < m < n$.
 - e. Construct a table exhibiting $p_m(n), 1 \leq m \leq n, 1 \leq n \leq 7$.
 - f. The number of partitions of n is the **partition number**

$$p(n) = \sum_{m=1}^n p_m(n).$$

Compute $p(n), 1 \leq n \leq 7$.

2. Explicitly write down
 - a. all 11 partitions of 6.
 - b. all 8 partitions of 7 having 3 or fewer parts.
 - c. all 8 partitions of 7 whose largest part is at most 3.
3. Let $\pi = [6, 4, 2^3]$. Find π^*
 - a. using Ferrers diagrams.
 - b. using Equation (1.1).
4. Which of the following is a self conjugate partition?
 - a. $[5, 4, 3, 2, 1]$
 - b. $[5, 3^2, 1^2]$
 - c. $[4, 3^2, 1]$
 - d. $[5, 3^2, 2, 1^2]$
 - e. $[5, 4^2, 3, 1^2]$
 - f. $[6, 4, 3, 1^2]$
5. Find π^* and use it to confirm Lemma 1.6 when $\pi =$
 - a. $[5, 4, 3, 2, 1]$
 - b. $[5, 3^2, 1^2]$
 - c. $[4, 3^2, 1]$
 - d. $[5, 3^2, 2, 1^2]$
 - e. $[5, 4^2, 3, 1^2]$
 - f. $[6, 4, 3, 1^2]$
6. Find the smallest integer n having three different self conjugate partitions.

7. Suppose $\pi \vdash n$. Show that $\pi_{j+1}^* = \pi_j^* - o(\{i : \pi_i = j\})$.
8. Let $\pi = [\pi_1, \pi_2, \dots, \pi_m]$ and $\rho = [\rho_1, \rho_2, \dots, \rho_k]$ be partitions of n . Show that $\pi \succ \rho$ only if $m \leq k$.
9. Find all the partitions of 7 that
 - a. majorize $[5, 2]$.
 - b. are majorized by $[2^2, 1^3]$.
10. Prove that $\pi \succ \rho$ if and only if $\rho^* \succ \pi^*$.
11. Show that the doubly stochastic matrix S given in Example 1.11 is not unique by finding another one that satisfies $(3, 3, 1, 1) = (5, 2, 1, 0)S$.
12. Show that there are $C(k + n - 1, n)$ nonnegative integer solutions to the equation $r_1 + r_2 + \dots + r_k = n$.
13. When $(a + b + c + d)^{10}$ is expressed as a linear combination of monomial symmetric functions, compute the coefficient of
 - a. $M_{[8, 1^2]}(a, b, c, d)$.
 - b. $M_{[7, 2, 1]}(a, b, c, d)$.
 - c. $M_{[4^2, 2]}(a, b, c, d)$.
 - d. $M_{[3^2, 2, 1^2]}(a, b, c, d)$.
14. Write out in full
 - a. $M_{[4, 1]}(x, y, z)$.
 - b. $M_{[3, 2]}(x, y, z)$.
 - c. $M_{[1^2]}(x, y, z)$.
15. Confirm Equation (1.10) for $a = 1, b = 2, c = 3$, and $d = 4$ by
 - a. using Example 1.19 to compute $E_n(1, 2, 3, 4)$, $1 \leq n \leq 4$.
 - b. computing the product $(x - 1)(x - 2)(x - 3)(x - 4)$.
16. Denote the roots of $p(x) = x^4 - x^2 + 2x + 2$ by a, b, c , and d . Compute the elementary symmetric functions $E_r(a, b, c, d)$, $1 \leq r \leq 4$,
 - a. from the coefficients of $p(x)$. (Hint: Equation (1.10).)
 - b. from the definition of E_r . (Hint: $(x + 1)^2$ divides $p(x)$.)
17. Suppose k is a fixed but arbitrary positive integer. Let $P_n = P_n(x_1, x_2, \dots, x_k)$ and $E_n = E_n(x_1, x_2, \dots, x_k)$, $n \geq 1$, be the n -th power sum and the n -th elementary symmetric function, respectively. It was shown by Isaac Newton (1642–1727) that, for any $n \geq 1$,

$$P_n - P_{n-1}E_1 + P_{n-2}E_2 - \dots + (-1)^{n-1}P_1E_{n-1} + (-1)^n n E_n = 0.$$

Thus, $P_1 - E_1 = 0$, $P_2 - P_1E_1 + 2E_2 = 0$, $P_3 - P_2E_1 + P_1E_2 - 3E_3 = 0$, and so on.

- a. Use Newton's identities to prove that

$$P_4 = E_1^4 - 4E_1^2E_2 + 4E_1E_3 + 2E_2^2 - 4E_4.$$

- b. Use Newton's identities to prove that

$$E_4 = \frac{1}{24}(P_1^4 - 6P_1^2P_2 + 8P_1P_3 + 3P_2^2 - 6P_4).$$

- c. Show that the general formula for E_r as a polynomial in the power sums is $r!E_r = \det(L_r)$, where

$$L_r = \begin{pmatrix} P_1 & 1 & 0 & 0 & \dots & 0 & 0 \\ P_2 & P_1 & 2 & 0 & \dots & 0 & 0 \\ P_3 & P_2 & P_1 & 3 & \dots & 0 & 0 \\ & & & & \dots & & \\ P_r & P_{r-1} & P_{r-2} & P_{r-3} & \dots & P_2 & P_1 \end{pmatrix}.$$

(Hint: Use Cramer's rule on the following matrix version of Newton's identities:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \dots \\ P_1 & -2 & 0 & 0 & \dots \\ P_2 & -P_1 & 3 & 0 & \dots \\ P_3 & -P_2 & P_1 & -4 & \dots \\ & & \dots & & \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \\ E_3 \\ E_4 \\ \vdots \end{pmatrix} = \begin{pmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ \vdots \end{pmatrix}.$$

- d. Prove that any polynomial, symmetric in $x_1, x_2, \dots, x_k$, is a polynomial in the power sum functions $P_n(x_1, x_2, \dots, x_k)$, $1 \leq n \leq k$.
18. If $2 \leq r \leq k$, prove that $E_r(x_1, x_2, \dots, x_k) = E_r(x_1, x_2, \dots, x_{k-1}) + x_k E_{r-1}(x_1, x_2, \dots, x_{k-1})$.
19. Use Equation (1.16) to confirm that Equation (1.17) yields Equation (1.15) when $n = k = 3$, $x_1 = a$, $x_2 = b$, and $x_3 = c$.
20. If $r \geq 2$, prove that $H_r(x_1, x_2, \dots, x_k) = H_r(x_1, x_2, \dots, x_{k-1}) + x_k H_{r-1}(x_1, x_2, \dots, x_{k-1})$.
21. Use Exercise 18 and mathematical induction to prove that

$$\prod_{i=1}^n (x - a_i) = \sum_{r=0}^n (-1)^r E_r(a_1, a_2, \dots, a_n) x^{n-r}.$$

22. Suppose $\alpha \in \Gamma_{m,n}$. Prove that $\alpha \in G_{m,n}$ if and only if $\alpha\sigma \succ \alpha$ for all permutations $\sigma \in S_m$.

23. Denote by $m_i(\pi)$ the multiplicity of i in the partition π , that is, the number of times i occurs as a part of π . Prove that $m_i(\pi) = \pi_i^* - \pi_{i+1}^*$.
24. Suppose $\pi, \rho \vdash n$. Let $\pi + \rho$ be the partition of $2n$, the i -th part of which is $\pi_i + \rho_i$ (with the convention that $\pi_i = 0$ if $i > L(\pi)$). Denote by $\pi \cup \rho$ the partition of $2n$ the parts of which are the parts of π together with the parts of ρ .
- Prove that $(\pi \cup \rho)^* = \pi^* + \rho^*$.
 - Is $(\pi + \rho)^* = \pi^* \cup \rho^*$?
25. Suppose $\pi \vdash n$. Let $\mu_i = \pi_i - i$ and $\nu_i = \pi_i^* - i$, $1 \leq i \leq f(\pi)$. Frobenius used $(\mu|\nu)$ to denote the partition π . Show that the Frobenius notation for π^* is $(\nu|\mu)$.
26. Among the many results known about elementary symmetric functions is that they are Schur concave, that is, $E_r(a) \leq E_r(b)$ whenever (a) majorizes (b) .
- Show that majorization imposes a linear order on the five partitions of 8 having 3 parts.
 - Confirm the Schur concavity of E_r by computing $E_r(\pi)$, $1 \leq r \leq 3$, for each three-part partition of 8.
27. Among the many results known about homogeneous symmetric functions is that they are Schur convex, that is, $H_r(a) \geq H_r(b)$ whenever (a) majorizes (b) .
- Confirm the Schur convexity of H_r by computing $H_r(\pi)$, $1 \leq r \leq 3$, for each three-part partition of 8.
 - If you were to compute $H_4(\pi)$ for each partition π of 24 having 3 parts, which partition would produce the maximum? Which would produce the minimum?
28. Let $E_r = E_r(a_1, a_2, \dots, a_n)$, $1 \leq r \leq n$. Show that $(1 - a_1x)(1 - a_2x) \dots (1 - a_nx) = 1 - E_1x + E_2x^2 - \dots + (-1)^n E_nx^n$.
29. Show that the dimension of $SC_7[x, y, z]$ is 8. (Hint: Exercise 2b.)
30. Compute
- $\dim(SC_7[x_1, x_2, \dots, x_7])$. (Hint: Exercise 1f.)
 - $\dim(SC_7[x_1, x_2, \dots, x_8])$.
31. If A is an m -by- n matrix, denote its i -th row and j -th column sums, respectively, by $r_i(A)$ and $c_j(A)$. Suppose

$$R = (r_1, r_2, \dots, r_m) \text{ and } C = (c_1, c_2, \dots, c_n)$$

are integer vectors satisfying $r_1 \geq r_2 \geq \dots \geq r_m \geq 0$ and $c_1 \geq c_2 \geq \dots \geq c_n \geq 0$. Then ([Gale (1957)] and [Ryser (1957)]) there exists an m -by- n ,

(0,1)-matrix A such that $r_i(A) = r_i$, $1 \leq i \leq m$, and $c_j(A) = c_j$, $1 \leq j \leq n$, if and only if $R^* \succ S$.

- a. Use the Gale-Ryser theorem to prove the existence of a 5-by-4, (0,1)-matrix A having row sum vector $R = (3, 2, 1, 1, 1)$ and column sum vector $(3, 3, 1, 1)$.
- b. Write down such a matrix.

Application Exercises

32. Draw pictures of the 11 nonisomorphic graphs on four vertices.
33. Prove Theorem 1.34.
34. Draw Ferrers diagrams for all the maximal graphic partitions of 6. (Hint: Figure 1.3.)
35. Let $\pi = [4, 2^3, 1]$.
 - a. Show that π satisfies Criteria (1.19).
 - b. Explain why π is not graphic.
36. Prove that

$$\sum_{i=1}^{L(\pi)} i\pi_i = \sum_{i=1}^{\pi_1} C(\pi_i^* + 1, 2).$$

(Hint: Figure 1.8(a)).

37. Confirm that the graphs in Figure 1.12 are threshold graphs.
38. Prove that K_n is a threshold graph, $n \geq 2$.
39. Prove that, apart from isolated vertices, the complement of a threshold graph is a threshold graph.
40. If $T = (V, E)$ is a tree, prove that it has one fewer edges than vertices.

CHAPTER 2

Inner Product Spaces

The purpose of this chapter is to review the more specialized results and techniques from linear algebra that will be needed in Chapters 4–8.¹ To begin, suppose V and W are finite dimensional vector spaces over the field $\mathbb{C}$ of complex numbers. Denote by $L(V, W)$ the set of all linear transformations from V into W . Then, together with addition and scalar multiplication defined by

$$(aS + bT)(v) = aS(v) + bT(v), \quad (2.1)$$

$L(V, W)$ is a vector space. The elements of $L(V, V)$ are called **linear operators**, and the elements of $L(V, \mathbb{C})$ **linear functionals**. The **dual space** of V is the vector space $V' = L(V, \mathbb{C})$, consisting of all linear functionals on V .

Suppose $B = \{e_1, e_2, \dots, e_n\}$ is a basis of V . For each $i \in \{1, 2, \dots, n\}$, define a linear functional $f_i \in V'$ by $f_i(e_j) = \delta_{i,j}$, $1 \leq j \leq n$, and linear extension.² If f is a fixed but arbitrary element of V' , then

$$f = \sum_{i=1}^n f(e_i) f_i, \quad (2.2)$$

as can be seen by evaluating both sides on e_j , $1 \leq j \leq n$. Therefore, $B' = \{f_1, f_2, \dots, f_n\}$ spans V' .

Suppose

$$\sum_{i=1}^n a_i f_i = 0,$$

¹ Among the many fine references to the topics of this chapter are [Marcus & Minc (1964)], [Horn & Johnson (1985)], and [Fiedler (1986)].

² The “Kronecker delta” is defined by $\delta_{i,j} = 1$, if $i = j$, and 0, otherwise.

the identically zero functional. Evaluating both sides on e_j proves that $a_j = 0$, $1 \leq j \leq n$. Therefore, B' is linearly independent.

DEFINITION 2.1 Let $B = \{e_1, e_2, \dots, e_n\}$ be a basis of the vector space V . The set $B' = \{f_1, f_2, \dots, f_n\}$ of linear functionals defined by $f_i(e_j) = \delta_{ij}$, $1 \leq j \leq n$, and linear extension, is called the basis of V' dual to B .

Because V' is a vector space, it too must have a dual.

THEOREM 2.2 Suppose $u \in V$. Let $\hat{u}: V' \rightarrow \mathbb{C}$ be defined by

$$\hat{u}(f) = f(u), \quad f \in V'. \quad (2.3)$$

Then $\hat{u}$ is a linear functional on V' . Moreover, the function $\Psi: V \rightarrow (V')'$ defined by $\Psi(u) = \hat{u}$, $u \in V$, is a vector space isomorphism.³

Because of Theorem 2.2, we will ignore the distinction between V and $(V')'$.

DEFINITION 2.3 An **inner product** on V is a complex valued function (u, v) of two vector variables that satisfies the following three conditions:

- (a) $(u, v) = \overline{(v, u)}$, for all $u, v \in V$.
- (b) $(au + bv, w) = a(u, w) + b(v, w)$, for all $a, b \in \mathbb{C}$, and $u, v, w \in V$.
- (c) $(u, u) > 0$ for all nonzero vectors $u \in V$.

A vector space endowed with a fixed inner product is called an **inner product space**.

DEFINITION 2.4 Denote by $\mathbb{C}^n = \{(x_1, x_2, \dots, x_n): x_i \in \mathbb{C}, 1 \leq i \leq n\}$ the vector space of n -tuples of complex numbers, and by $\mathbb{C}_{m,n}$ the vector space of m -by- n complex matrices. If $A = (a_{ij}) \in \mathbb{C}_{m,n}$, denote its **conjugate transpose** by A^* . That is, A^* is the n -by- m matrix whose (i, j) -entry is $\bar{a}_{ji}$.

EXAMPLE 2.5 If $V = \mathbb{C}_{n,n}$, then the trace function,

$$\text{tr}(A) = \sum_{i=1}^n a_{ii},$$

$A = (a_{ij}) \in \mathbb{C}_{n,n}$, is a linear functional on V .

Suppose $V = \mathbb{C}^n$. If $c_1, c_2, \dots, c_n$ are positive numbers, then

$$(u, v) = \sum_{i=1}^n c_i x_i \bar{y}_i, \quad (2.4)$$

where $u = (x_1, x_2, \dots, x_n)$ and $v = (y_1, y_2, \dots, y_n) \in V$, defines an inner product on V . When $c_1 = c_2 = \dots = c_n = 1$, it is called the **standard inner**

³That is, an invertible transformation from V onto $(V')'$.

product or dot product. When $V = \mathbb{C}_{m,n}$, the standard inner product can be written $(A, B) = \text{tr}(B^*A)$. $\square$

DEFINITION 2.6 Let u and v be vectors in an inner product space V . If $(u, v) = 0$, then u and v are said to be **orthogonal**. A nonempty subset $S \subset V$ is **orthogonal** if the vectors in S are pairwise orthogonal. Finally, S is **orthonormal** (or **o.n.**) provided it is orthogonal and $(u, u) = 1$ for all $u \in S$.

Note that 0 is orthogonal to every vector and that an orthonormal set of vectors is necessarily linearly independent.

DEFINITION 2.7 Let v be a vector in an inner product space V . The **norm** of v , written $\|v\|$, is the nonnegative square root of (v, v) . If $\|v\| = 1$, then v is a **unit vector**.

EXAMPLE 2.8 Let v be a nonzero vector in the inner product space V . Let $u = cv$, where $c = 1/\|v\|$. Then

$$\begin{aligned}\|u\| &= \sqrt{(u, u)} \\ &= \sqrt{(cv, cv)} \\ &= c\|v\| \\ &= 1,\end{aligned}$$

that is, $u = v/\|v\|$ is a unit vector. $\square$

Suppose that $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ is a basis of the inner product space V . We are going to describe an algorithm, called the **Gram-Schmidt process**, for transforming $\mathcal{B}$ into an **orthonormal** basis of V . This is how it works: Let

$$\begin{aligned}u_1 &= \frac{v_1}{\|v_1\|}, \\ u_2 &= \frac{v_2 - (v_2, u_1)u_1}{\|v_2 - (v_2, u_1)u_1\|}, \\ u_3 &= \frac{v_3 - (v_3, u_1)u_1 - (v_3, u_2)u_2}{\|v_3 - (v_3, u_1)u_1 - (v_3, u_2)u_2\|},\end{aligned}$$

and so on. In general,

$$w_k = v_k - \sum_{i=1}^{k-1} (v_k, u_i)u_i,$$

and $u_k = w_k / \|w_k\|$. Now, w_k is not zero because $\{v_1, v_2, \dots, v_k\}$ is linearly independent. Thus, u_k is a unit vector, for all k . Moreover,

$$\begin{aligned} (u_2, u_1) &= (w_2, u_1) / \|w_2\| \\ &= ((v_2, u_1) - (v_2, u_1)(u_1, u_1)) / \|w_2\| \\ &= 0, \end{aligned}$$

because u_1 is a unit vector. Thus, $\{u_1, u_2\}$ is orthonormal. Given that $\{u_1, u_2, \dots, u_{k-1}\}$ is orthonormal,

$$\begin{aligned} (u_k, u_j) &= (w_k, u_j) / \|w_k\| \\ &= \left((v_k, u_j) - \sum_{i=1}^{k-1} (v_k, u_i)(u_i, u_j) \right) / \|w_k\| \\ &= ((v_k, u_j) - (v_k, u_j)) / \|w_k\| \\ &= 0, \end{aligned}$$

$j < k$, that is, $\{u_1, u_2, \dots, u_k\}$ is orthonormal. Finally, note that $\langle v_1, v_2, \dots, v_k \rangle$, the subspace of V spanned by the first k basis vectors, is equal to $\langle u_1, u_2, \dots, u_k \rangle$, $1 \leq k \leq n$.

THEOREM 2.9 *Let V be an inner product space of dimension n . Then V has an orthonormal basis. Moreover, if S is an orthonormal set of vectors in V , then S can be extended to an o.n. basis of V .*

Proof The existence of an o.n. basis follows from the Gram-Schmidt process. If $S = \{u_1, u_2, \dots, u_k\}$ is an orthonormal set, then it is linearly independent. Therefore, S can be extended to a basis, $\{u_1, u_2, \dots, u_k, v_{k+1}, \dots, v_n\}$ of V . Applying the Gram-Schmidt process to this basis does not change any of its first k vectors. $\square$

DEFINITION 2.10 Let W be a subspace of the inner product space V . The **orthogonal complement** of W is

$$W^\perp = \{v \in V : (w, v) = 0 \text{ for all } w \in W\}.$$

THEOREM 2.11 *If W is a subspace of the inner product space V , then $W^\perp$ is a subspace. Moreover, $\dim(W^\perp) + \dim(W) = \dim(V)$.*

REPRESENTATION THEOREM FOR LINEAR FUNCTIONALS 2.12 *Let V be an inner product space. If $f: V \rightarrow \mathbb{C}$ is a linear functional, then there exists a unique vector $w \in V$ such that $f(v) = (v, w)$, for all $v \in V$.*

Proof Let $\{u_1, u_2, \dots, u_n\}$ be an orthonormal basis of V . Define

$$w = \sum_{i=1}^n \overline{f(u_i)} u_i. \quad (2.5)$$

If $v = \sum a_i u_i \in V$, then $a_i = (v, u_i)$, $1 \leq i \leq n$. Hence,

$$\begin{aligned} (v, w) &= \left(v, \sum_{i=1}^n \overline{f(u_i)} u_i \right) \\ &= \sum_{i=1}^n f(u_i) (v, u_i) \\ &= \sum_{i=1}^n a_i f(u_i) \\ &= f(v). \end{aligned}$$

If $(v, w_1) = (v, w_2)$ for all $v \in V$, then $(v, w_1 - w_2) = 0$, for all $v \in V$. Choosing $v = w_1 - w_2$ produces (Definition 2.3(c)) $w_1 - w_2 = 0$. $\square$

Note that the vector w in Equation (2.5) depends on f . If, for example, $\{f_1, f_2, \dots, f_n\}$ is the basis of V' dual to $\{u_1, u_2, \dots, u_n\}$, then $f_j(v) = (v, u_j)$ for all $v \in V$. Hence, the w that works for f_j is u_j .

CAUCHY-SCHWARZ INEQUALITY 2.13 *Let V be an inner product space. If $u, v \in V$, then $|(u, v)| \leq \|u\| \|v\|$, with equality if and only if u and v are linearly dependent.*

Proof If $u = 0$, there is nothing to prove. Otherwise, let

$$w = v - \frac{(v, u)}{\|u\|^2} u. \quad (2.6)$$

Then, because $(w, u) = 0$,

$$\begin{aligned} 0 &\leq (w, w) \\ &= (w, v) \\ &= (v, v) - \frac{(v, u)}{\|u\|^2} (u, v). \end{aligned}$$

This completes the proof because $(v, u)(u, v) = |(u, v)|^2$, and $w = 0$ if and only if $\{u, v\}$ is linearly dependent. $\square$

PARSEVAL'S IDENTITY 2.14 *Let $\{u_1, u_2, \dots, u_n\}$ be an orthonormal basis of the inner product space V . If $v, w \in V$, then*

$$(v, w) = \sum_{i=1}^n (v, u_i)(u_i, w).$$

Proof Let $v = \sum a_i u_i$ and $w = \sum b_j u_j$. Then

$$\begin{aligned} (v, w) &= \left(\sum_{i=1}^n a_i u_i, \sum_{j=1}^n b_j u_j \right) \\ &= \sum_{i=1}^n \sum_{j=1}^n a_i \bar{b}_j (u_i, u_j) \\ &= \sum_{i=1}^n \sum_{j=1}^n a_i \bar{b}_j \delta_{i,j} \\ &= \sum_{i=1}^n a_i \bar{b}_i \\ &= \sum_{i=1}^n (v, u_i)(u_i, w). \end{aligned}$$

□

BESSEL'S INEQUALITY 2.15 *Let $\{u_1, u_2, \dots, u_k\}$ be an orthonormal set of vectors in the inner product space V . If $v \in V$, then*

$$\|v\|^2 \geq \sum_{j=1}^k |(v, u_j)|^2,$$

with equality if and only if $v \in \langle u_1, u_2, \dots, u_k \rangle$, the subspace of V spanned by $\{u_1, u_2, \dots, u_k\}$.

THEOREM 2.16 *Let V be an inner product space and suppose $T \in L(V, V)$. Then there exists a unique $S \in L(V, V)$ such that $(T(v), w) = (v, S(w))$, for all $v, w \in V$.*

Proof Let $\{u_1, u_2, \dots, u_n\}$ be an orthonormal basis of V . Define a function $S: V \rightarrow V$ by

$$S(w) = \sum_{i=1}^n (w, T(u_i)) u_i, \quad w \in V.$$

Then, by Definition 2.3(b), $S \in L(V, V)$. Let $v = a_1 u_1 + a_2 u_2 + \dots + a_n u_n$ be a fixed but arbitrary vector in V . Then, because $a_i = (v, u_i)$,

$$\begin{aligned} (v, S(w)) &= \left(v, \sum_{i=1}^n (w, T(u_i)) u_i \right) \\ &= \sum_{i=1}^n \overline{(w, T(u_i))} (v, u_i) \\ &= \sum_{i=1}^n (T(u_i), w) a_i, \\ &= \left(\sum_{i=1}^n a_i T(u_i), w \right) \\ &= (T(v), w). \end{aligned}$$

Uniqueness is a consequence of the following lemma. □

LEMMA 2.17 Let V be an inner product space. Suppose $T_1, T_2 \in L(V, V)$. If $(T_1(v), w) = (T_2(v), w)$, for all $v, w \in V$, then $T_1 = T_2$.

DEFINITION 2.18 Let V be an inner product space and suppose $T \in L(V, V)$. The unique $S \in L(V, V)$ guaranteed by Theorem 2.16 is denoted T^* and called the **adjoint** of T .

EXAMPLE 2.19 Let $\mathcal{B} = \{u_1, u_2, \dots, u_n\}$ be an orthonormal basis of the inner product space V . If $T \in L(V, V)$, then T is completely (and uniquely) determined by its action on the basis $\mathcal{B}$, that is, by the coefficients a_{ij} in the equations

$$T(u_j) = \sum_{i=1}^n a_{ij} u_i, \quad 1 \leq j \leq n. \quad (2.7)$$

Denote the **matrix representation** of T with respect to the basis $\mathcal{B}$ by $[T] = (a_{ij})$. Because $\mathcal{B}$ is orthonormal, it follows from Equation (2.7) that $a_{ij} = (T(u_j), u_i)$.

Therefore,

$$\begin{aligned}\bar{a}_{ji} &= \overline{(T(u_i), u_j)} \\ &= \overline{(u_i, T^*(u_j))} \\ &= (T^*(u_j), u_i),\end{aligned}$$

the (i, j) -entry of the matrix representation of T^* . Evidently, $[T^*] = [T]^*$. With respect to an orthonormal basis, the matrix representation of T^* is the conjugate transpose of the matrix representation of T . $\square$

DEFINITION 2.20 Let W_1 and W_2 be subspaces of an inner product space V . Their **sum** is $W_1 + W_2 = \{w_1 + w_2 : w_1 \in W_1 \text{ and } w_2 \in W_2\}$; if $W_1 \cap W_2 = \{0\}$, the sum is **direct**, and is written $W_1 \oplus W_2$. If $W_1 \subset W_2^\perp$, then the sum is **orthogonal direct**, and is written $W_1 \perp W_2$.

EXAMPLE 2.21 Let T be a linear operator on the inner product space V . If $T = T^2$, the composition of T with itself, then T is called a **projection**. Denote by $I_V \in L(V, V)$ the **identity operator** on V , that is, $I_V(u) = u$, for all $u \in V$. If $S = I_V - T$ then, because $S^2 = (I_V - T)(I_V - T) = I_V - T - T + T^2 = I_V - T$, S is another projection. Denote the **image** of T by $T(V) = \{T(v) : v \in V\}$. Because $T^2 = T$, $T(w) = w$ for all $w \in T(V)$. Similarly, $S(w) = w$ for all $w \in S(V)$. It follows from $v = T(v) + (I_V - T)(v) = T(v) + S(v)$ that V is the sum $T(V) + S(V)$. Because $w = S(w) = w - T(w) = w - w = 0$, for all $w \in T(V) \cap S(V)$, V is the direct sum $T(V) \oplus S(V)$. $\square$

DEFINITION 2.22 Let V be an inner product space and suppose $T \in L(V, V)$. If $T^* = T$, then T is **hermitian** (or **self adjoint**). A hermitian projection is an **orthogonal projection**. The matrix $A \in \mathbb{C}_{n,n}$ is **hermitian** if $A^* = A$.

EXAMPLE 2.23 Suppose T is an orthogonal projection on the inner product space V . Let $S = I_V - T$. For a fixed but arbitrary $v \in V$, let $v_1 = T(v)$ and $v_2 = S(v)$. Because $T^* = T = T^2$,

$$\begin{aligned}(v_1, v_2) &= (T(v), v - T(v)) \\ &= (v, T(v) - T^2(v)) \\ &= (v, 0) \\ &= 0.\end{aligned}$$

It follows that $V = T(V) \perp S(V)$. $\square$

Let V be an inner product space and suppose $T \in L(V, V)$. Define $[v, w] = (T(v), w)$, for all $v, w \in V$. Under what conditions is $[,]$ an inner product on V ? Of the three conditions in Definition 2.3, $[au + bv, w] = a[u, w] + b[v, w]$ is always satisfied. However,

$$\begin{aligned}\overline{[v, u]} &= \overline{(T(v), u)} \\ &= \overline{(v, T^*(u))} \\ &= (T^*(u), v),\end{aligned}$$

whereas, $[u, v] = (T(u), v)$. Evidently, $[v, u] = \overline{[u, v]}$, for all $u, v \in V$, if and only if $(T^*(u), v) = (T(u), v)$, for all $u, v \in V$, if and only if (Lemma 2.17) $T = T^*$, that is, if and only if T is hermitian.

DEFINITION 2.24 Let T be a linear operator on the inner product space V . If $(T(v), v) > 0$ for every nonzero vector $v \in V$, then T is **positive definite**. If $(T(v), v) \geq 0$ for all $v \in V$, then T is **positive semidefinite**. The notation $T \geq 0$ indicates that T is positive semidefinite hermitian.

We have shown that $[u, v] = (T(u), v)$ defines an inner product on V if and only if T is a positive definite hermitian operator. In fact, as we now see, every inner product on V arises in this way.

THEOREM 2.25 Suppose V is an inner product space. Let $[,]$ be a second inner product on V . Then there exists a unique (positive definite hermitian) operator $T \in L(V, V)$ such that $[u, v] = (T(u), v)$, for all $u, v \in V$.

Proof Let $\{u_1, u_2, \dots, u_n\}$ be a basis of V that is orthonormal with respect to the inner product $(,)$. Define $T \in L(V, V)$ by

$$T(u) = \sum_{i=1}^n [u, u_i] u_i,$$

$u \in V$. If $v = \sum a_j u_j$, then

$$\begin{aligned}(T(u), v) &= \sum_{i=1}^n \sum_{j=1}^n [u, u_i] \bar{a}_j (u_i, u_j) \\ &= \sum_{i=1}^n \bar{a}_i [u, u_i] \\ &= [u, v].\end{aligned}$$

Because $[,]$ is an inner product, T is positive definite hermitian. Uniqueness follows from Lemma 2.17. $\square$

The matrix version of Definition 2.24 is this:

DEFINITION 2.26 Let $A \in \mathbf{C}_{n,n}$ be hermitian. If $x^*Ax > 0$ for every nonzero $x \in \mathbf{C}_{n,1}$, then A is **positive definite**. If $x^*Ax \geq 0$ for all $x \in \mathbf{C}_{n,1}$, then A is **positive semidefinite**. The notation $A \geq 0$ indicates that A is positive semidefinite hermitian.

Observe that A is positive definite if and only if it is positive semidefinite and invertible.

LEMMA 2.27 Let B be an orthonormal basis of the inner product space V . Suppose $T \in L(V, V)$. Then $T \geq 0$ if and only if $[T] \geq 0$, where $[T]$ is the matrix representation of T with respect to B .

DEFINITION 2.28 A linear operator T on an inner product space V is **unitary** if $T^*T = I_V$, that is, if T is invertible and $T^{-1} = T^*$. A matrix $U \in \mathbf{C}_{n,n}$ is **unitary** if it is invertible and $U^{-1} = U^*$, the conjugate transpose of U .

SCHUR'S TRIANGULARIZATION THEOREM 2.29 Let $A \in \mathbf{C}_{n,n}$. Then A is unitarily similar to an upper triangular matrix.

Proof The proof is by induction on n . Because every 1-by-1 matrix is upper triangular, the $n = 1$ case is trivial. Assume $n > 1$. Let λ be an eigenvalue of A afforded by the eigenvector $x \in \mathbf{C}_{n,1}$. Because $x \neq 0$, we may assume $\|x\| = 1$ (with respect to the standard inner product $(y, z) = z^*y$ on $\mathbf{C}_{n,1}$). By Theorem 2.9, $\{x\}$ can be extended to an orthonormal basis $\{x, y_2, \dots, y_n\}$ of $\mathbf{C}_{n,1}$. Let U be the matrix whose first column is x and whose j -th column is y_j , $1 < j \leq n$. Then $U^*U = I_n$, the n -by- n identity matrix. Moreover, the first column of U^*AU is $U^*Ax = \lambda U^*x = \lambda C_1$, where C_1 is the first column of I_n , that is,

$$U^*AU = \begin{pmatrix} \lambda & \# & \# & \dots & \# \\ 0 & & & & \\ 0 & & A_1 & & \\ \vdots & & & & \\ 0 & & & & \end{pmatrix},$$

where A_1 is an $(n-1)$ -by- $(n-1)$ matrix, and the $\#$'s stand for unspecified entries.

It follows from the induction hypothesis that there is an $(n-1)$ -square unitary matrix U_1 such that $U_1^*A_1U_1$ is upper triangular. Let $L = (1) \oplus U_1$, that is,

$$L = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & & & & \\ 0 & & U_1 & & \\ \vdots & & & & \\ 0 & & & & \end{pmatrix}.$$

Then

$$(UL)^*A(UL) = \begin{pmatrix} \lambda & \# & \# & \dots & \# \\ 0 & & & & \\ 0 & & U_1^*A_1U_1 & & \\ \vdots & & & & \\ 0 & & & & \end{pmatrix}$$

is upper triangular. Because it is a product of unitary matrices, UL is unitary, and the proof is complete. $\square$

SPECTRAL THEOREM 2.30 Let $A \in \mathbb{C}_{n,n}$. Then A is unitarily similar to a diagonal matrix if and only if $A^*A = AA^*$.

Proof Let U be a unitary matrix such that $U^*AU = D$ is diagonal. Then $A^*A = (UD^*U^*)(UDU^*) = U(D^*D)U^* = U(DD^*)U^* = (UDU^*)(UD^*U^*) = AA^*$. Conversely (Exercise 34), there is a unitary matrix U such that U^*AU and U^*A^*U are both upper triangular. Because U^*A^*U is upper triangular, its conjugate transpose, U^*AU , must be lower triangular. Therefore, U^*AU is both upper and lower triangular. $\square$

DEFINITION 2.31 A square matrix A (linear operator T) is **normal** if $A^*A = AA^*$ ($T^*T = TT^*$).

If $A = (a_{ij})$ is a hermitian matrix, then $\bar{a}_{ii} = a_{ii}$ is real. Because hermitian matrices are normal, there exists a unitary matrix U such that $U^*AU = D$, a diagonal matrix whose main diagonal entries are the eigenvalues of A . Because $D^* = (U^*AU)^* = U^*A^*U = U^*AU = D$, the eigenvalues of A are all real as well.

The next result is a restatement of the Spectral Theorem for linear operators.

COROLLARY 2.32 Let T be a linear operator on an inner product space V . Then T is normal if and only if there exists an orthonormal basis of V consisting of eigenvectors for T .

DEFINITION 2.33 Let $A \in \mathbb{C}_{n,n}$. The **field of values** or **numerical range** of A is $F(A) = \{x^*Ax : x \in \mathbb{C}_{n,1} \text{ and } \|x\| = 1\}$.

One may view $F(A)$ as the set of those complex numbers that can occur as the $(1,1)$ -entry of a matrix unitarily similar to A .

Let $S = \{c_1, c_2, \dots, c_n\}$ be a set of complex numbers. The **convex hull** of S is the set

$$\{\theta_1 c_1 + \theta_2 c_2 + \dots + \theta_n c_n : \theta_i \geq 0, 1 \leq i \leq n, \text{ and } \theta_1 + \theta_2 + \dots + \theta_n = 1\}$$

consisting of all convex combinations of the elements of S .

THEOREM 2.34 *If $A \in \mathbf{C}_{n,n}$ is normal, then $F(A)$ is the convex hull of its eigenvalues. In particular, if A is hermitian with eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$, then $F(A) = [\lambda_n, \lambda_1]$.*

Proof Sketch Suppose $A \in \mathbf{C}_{n,n}$. It follows from Schur's Triangularization Theorem that the eigenvalues of A are elements of $F(A)$. By a theorem of Toeplitz and Hausdorff,⁴ $F(A)$ contains the convex hull of each of its finite subsets. Therefore, normal or not, the numerical range of A contains the convex hull of its eigenvalues.

With respect to the dot product, $V = \mathbf{C}_{n,1}$ is an inner product space. If $A \in \mathbf{C}_{n,n}$, then $x \rightarrow Ax$ defines a linear operator on V . If A is normal, there is an orthonormal basis $B = \{x_1, x_2, \dots, x_n\}$ of V consisting of eigenvectors of A , say, $Ax_i = \lambda_i x_i$, $1 \leq i \leq n$. If $x = \sum a_i x_i \in V$, then

$$x^* Ax = \sum_{i=1}^n \lambda_i |a_i|^2. \quad (2.8)$$

If x is a unit vector, then $\sum |a_i|^2 = 1$. This proves that the numerical range of A is contained in the convex hull of its eigenvalues.⁵ $\square$

Let A be a hermitian matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. Then, from Theorem 2.34,

$$\lambda_1 = \max_{\|u\|=1} u^* Au \quad (2.9)$$

and

$$\lambda_n = \min_{\|u\|=1} u^* Au, \quad (2.10)$$

where the maximum and minimum are over the unit vectors $u \in V = \mathbf{C}_{n,1}$. This proves the following:

COROLLARY 2.35 *Let $A \in \mathbf{C}_{n,n}$ be hermitian. Then $A \geq 0$ if and only if its eigenvalues are all nonnegative.*

In fact, Equations (2.9) and (2.10) are but the simplest examples of the following elegant result [Fischer (1905)]. (A generalization to a wider class of operators can be found in [Courant & Hilbert (1953)].)

⁴The convexity of $F(A)$, $A \in \mathbf{C}_{n,n}$, was proved independently by [Toeplitz (1918)] and [Hausdorff (1919)]. For a discussion of why the theorem is true, see [Davis (1971)].

⁵Related results can be found in [Fiedler (1981)], [Grone, Johnson, Sa & Wolkowicz (1987)], and [Johnson (1976)].

COURANT-FISHER THEOREM 2.36 Let $A \in \mathbb{C}_{n,n}$ be a hermitian matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. Then

$$\lambda_k = \min_W \max_{\substack{u \in W^\perp \\ \|u\|=1}} u^* A u, \quad (2.11)$$

where the minimum is over the $(k-1)$ -dimensional subspaces W of $V = \mathbb{C}_{n,1}$. Alternatively,

$$\lambda_k = \max_W \min_{\substack{u \in W^\perp \\ \|u\|=1}} u^* A u, \quad (2.12)$$

where the maximum is over the $(n-k)$ -dimensional subspaces of V .

Proof When $k = 1$ or $k = n$, Equation (2.11) follows from Equations (2.9)–(2.10). Suppose $n > k > 1$. Let $u_1, u_2, \dots, u_n$ be an orthonormal family of eigenvectors of A that afford $\lambda_1, \lambda_2, \dots, \lambda_n$, respectively. If U is a unitary matrix whose first $k-1$ columns are $u_1, u_2, \dots, u_{k-1}$, then $U^* A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_{k-1}) \oplus B$, where the spectrum of B is $\lambda_k \geq \lambda_{k+1} \geq \dots \geq \lambda_n$. It follows from Equation (2.9) that

$$\begin{aligned} \lambda_k &= \max_{\|v\|=1} v^* B v \\ &= \max u^* A u, \end{aligned}$$

where the second maximum is over the unit vectors u in the orthogonal complement of $W = \langle u_1, u_2, \dots, u_{k-1} \rangle$. In other words, there exists a $(k-1)$ -dimensional subspace W of V such that λ_k is equal to the maximum value of $u^* A u$ over the unit vectors in $W^\perp$. Therefore, λ_k is at least as large as the minimum of this maximum over all $(k-1)$ -dimensional subspaces. That is, $\lambda_k \geq \min \max u^* A u$.

If U is the n -by- n matrix whose i -th column is the eigenvector u_i , $1 \leq i \leq n$, then $U^* A U = D$, where $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$. Let W be a fixed but arbitrary $(k-1)$ -dimensional subspace of V , and set

$$\begin{aligned} M &= \max_{\substack{u \in W^\perp \\ \|u\|=1}} u^* A u \\ &= \max_{\substack{u \in X^\perp \\ \|u\|=1}} u^* D u, \end{aligned}$$

where $X = \{U^* w : w \in W\}$. If M_1 is the maximum of $u^* D u$ over the unit vectors $u \in X^\perp$ whose last $n-k$ components are zero, then $M \geq M_1$. If

$v = (a_1, a_2, \dots, a_k, 0, 0, \dots, 0)^t$, is such a unit vector (which must exist because $\dim(X^\perp) = n - k + 1$), then

$$\begin{aligned}
 M_1 &\geq v^* D v \\
 &= \sum_{i=1}^k \lambda_i |a_i|^2 \\
 &\geq \lambda_k \sum_{i=1}^k |a_i|^2 \\
 &= \lambda_k \|v\|^2 \\
 &= \lambda_k,
 \end{aligned}$$

because v is a unit vector. Therefore, $M \geq \lambda_k$. Because W was arbitrary, the minimum of all such M 's cannot be less than λ_k , that is, $\min \max u^* A u \geq \lambda_k$. This completes the proof of Equation (2.11). The proof of Equation (2.12) is similar. $\square$

DEFINITION 2.37 Suppose A and B are hermitian n -by- n matrices, or hermitian operators on an inner product space V . If $A - B \geq 0$, then A dominates B , written $A \geq B$.

If $A \geq B$, then $u^* A u \geq u^* B u$ for every unit vector $u \in \mathbb{C}_{n,1}$. Therefore (Equation (2.11)), the k -th largest eigenvalue of A ,

$$\lambda_k(A) \geq \lambda_k(B), \quad (2.13)$$

the k -th largest eigenvalue of B .

If $A \in \mathbb{C}_{n,n}$ is hermitian, let $d(A)$ be the n -tuple consisting of the main diagonal entries of A (multiplicities included) arranged in nonincreasing order, and denote by $s(A)$ the n -tuple consisting of the eigenvalues of A (multiplicities included) arranged in nonincreasing order.

THEOREM 2.38 If $A = (a_{ij}) \geq 0$, then $s(A)$ majorizes $d(A)$.

Proof Let P be a permutation matrix such that the main diagonal of $P^* A P$ is $d(A)$. Let $U = (u_{ij})$ be a unitary matrix such that $U^* (P^* A P) U = D$, where $D = \text{diag}(s(A))$. Then $P^* A P = U D U^*$. Comparing the (i, i) -entries of these (equal) matrices, we see that

$$\begin{aligned}
(P^*AP)_{ii} &= \sum_{j=1}^n u_{ij} \lambda_j \bar{u}_{ij} \\
&= \sum_{j=1}^n |u_{ij}|^2 \lambda_j, \quad 1 \leq i \leq n.
\end{aligned} \tag{2.14}$$

Because U is unitary,

$$\begin{aligned}
(UU^*)_{ii} &= \sum_{j=1}^n |u_{ij}|^2 \\
&= 1,
\end{aligned}$$

$1 \leq i \leq n$. Similarly,

$$\sum_{i=1}^n |u_{ij}|^2 = 1,$$

$1 \leq i \leq n$. Therefore, $Q = (q_{ij})$ defined by $q_{ij} = |u_{ij}|^2$, is doubly stochastic. Because Equations (2.14) can be written in the form $d(A) = s(A)Q'$, the result follows from Theorem 1.10. $\square$

COROLLARY 2.39 *Let $A = (a_{ij}) \in \mathbb{C}_{n,n}$ be a hermitian matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ and nonincreasing diagonal sequence $d(A) = (b_1, b_2, \dots, b_n)$. Then*

$$\sum_{i=1}^k \lambda_i \geq \sum_{i=1}^k b_i, \quad 1 \leq k \leq n. \tag{2.15}$$

Proof By Corollary 2.35, $H = A - \lambda_n I_n \geq 0$. By Theorem 2.38,

$$\begin{aligned}
s(H) &= (\lambda_1 - \lambda_n, \lambda_2 - \lambda_n, \dots, \lambda_n - \lambda_n) \\
&\succ (b_1 - \lambda_n, b_2 - \lambda_n, \dots, b_n - \lambda_n) \\
&= d(H).
\end{aligned}$$

$\square$

Corollary 2.39 was first proved in [Schur (1923)]. The following lovely converse appeared in [Horn (1954)]: Let $(a) = (a_1, a_2, \dots, a_n)$ and $(b) = (b_1, b_2, \dots, b_n)$ be nonincreasing sequences of real numbers. If

$$\sum_{i=1}^k a_i \geq \sum_{i=1}^k b_i, \quad 1 \leq k \leq n,$$

with equality when $k = n$, then there exists a real, symmetric, n -by- n matrix A with eigenvalues $a_1, a_2, \dots, a_n$ and diagonal entries $b_1, b_2, \dots, b_n$. (An excellent reference for these and related results is [Marshall & Olkin (1979)].)

DEFINITION 2.40 Hadamard's function is defined by

$$h(A) = \prod_{i=1}^n a_{ii}, \quad A = (a_{ij}) \in \mathbf{C}_{n,n}.$$

HADAMARD'S THEOREM 2.41 If $A \geq 0$, then $h(A) \geq \det(A)$, with equality if and only if A has a zero row and column, or A is diagonal.

Proof Sketch In the same (1923) paper that contained Corollary 2.39, Schur proved that the elementary symmetric functions are what we now call Schur concave: If $(a) = (a_1, a_2, \dots, a_n)$ and $(b) = (b_1, b_2, \dots, b_n)$ are nonincreasing sequences of *positive* real numbers such that $(a) \succ (b)$, then $E_r(a) \leq E_r(b)$, $1 < r \leq n$, with equality if and only if $(a) = (b)$. Applying the $r = n$ case of this result to $s(A) \succ d(A)$, we obtain

$$\begin{aligned} \det(A) &= \prod_{i=1}^n \lambda_i \\ &\leq \prod_{i=1}^n a_{ii} \\ &= h(A), \end{aligned}$$

with equality if and only if $\lambda_i = a_{ii}$ for all i (if and only if A is diagonal), or $a_{ii} = 0$ for some i (in which case every entry in row and column i of A is 0). $\square$

Theorem 2.41 was proved in [Hadamard (1893)].⁶ We shall have more to say about Hadamard's Theorem in Chapter 7.

⁶In 1944, Jacques Hadamard had this to say about Theorem 2.41: "To continue about my failures, I shall mention one which I particularly regret. It concerns the celebrated Dirichlet problem which I, for years, tried to solve in the same initial direction as Fredholm did ... [In] 1893, I had been attracted by a question in algebra (on determinants). When solving it, I had no suspicions of any definite use it might have, only *feeling* that it deserved interest; then in 1900 appeared Fredholm's theory, for which the result obtained in 1893 happens to be essential."

THE INTERLACING INEQUALITIES 2.42 Let $A \in \mathbb{C}_{n,n}$ be a hermitian matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Suppose B is an r -by- r principal submatrix of A . If the eigenvalues of B are $\eta_1 \geq \eta_2 \geq \cdots \geq \eta_r$, then $\lambda_k \geq \eta_k \geq \lambda_{n+k-r}$, $1 \leq k \leq r$.

Proof If $r = n$, equality holds for all k . If $r = 1$, the result follows from Theorem 2.34. Thus, we may proceed under the assumption that $2 \leq r \leq n - 1$. Because the eigenvalues of A are unchanged by a permutation similarity, we may assume that B is the *leading* r -by- r principal submatrix of A (the submatrix obtained by deleting rows and columns $r + 1, r + 2, \dots, n$). From Equation (2.9), λ_1 is the maximum of $u^* A u$ over all unit vectors $u \in \mathbb{C}_{n,1}$, while η_1 is the maximum of $v^* B v$ over all unit vectors $v \in \mathbb{C}_{r,1}$. That is η_1 is the maximum of $u^* A u$ over all unit vectors $u \in \mathbb{C}_{n,1}$ the last $n - r$ components of which are zero. Therefore, $\lambda_1 \geq \eta_1$. If $k > 1$, a similar argument based on Equation (2.11) shows that $\lambda_k \geq \eta_k$, $r \geq k \geq 2$.

Turning to the inequalities $\eta_k \geq \lambda_{n+k-r}$, we may continue to assume that B is the leading r -by- r principal submatrix of A . Consider the case $r = n - 1$, so that the inequalities become $\eta_k \geq \lambda_{k+1}$, $1 \leq k \leq n - 1$. From Equation (2.12),

$$\begin{aligned} \lambda_{k+1} &= \max_W \min_{\substack{u \in W^\perp \\ \|u\|=1}} u^* A u \\ &\leq \max_W \min_{\substack{u \in W^\perp \\ \|u\|=1 \\ u \perp E_n}} u^* A u, \end{aligned}$$

where the maxima are over the subspaces W of $\mathbb{C}_{n,1}$ of dimension $n - (k + 1)$, and $E_n \in \mathbb{C}_{n,1}$ is the column vector whose only nonzero entry is a 1 in row n . If $u \perp E_n$, then $u^* A u = v^* B v$, where $v \in \mathbb{C}_{n-1,1}$ is the column vector obtained from u by deleting the 0 from row n .

If W is a fixed but arbitrary $(n - k - 1)$ -dimensional subspace of $\mathbb{C}_{n,1}$, then $\dim(\{u \in \mathbb{C}_{n,1} : u \in W^\perp \text{ and } u \perp E_n\})$ is either $k + 1$ or k depending on whether or not $E_n \in W$. Because every $(k + 1)$ -dimensional subspace of $\mathbb{C}_{n-1,1}$ contains k -dimensional subspaces,

$$\begin{aligned} \max_W \min_{\substack{u \in W^\perp \\ \|u\|=1 \\ u \perp E_n}} u^* A u &\leq \max_V \min_{\substack{v \in V^\perp \\ \|v\|=1}} v^* B v \\ &= \eta_k, \end{aligned}$$

where V ranges over the subspaces of $\mathbb{C}_{n-1,1}$ of dimension $(n - 1) - k$.

If $r < n - 1$, then B is a principal submatrix of an $(n - 1)$ -by- $(n - 1)$ principal submatrix of A and the result follows by induction. $\square$

LEMMA 2.43 Suppose $A \in \mathbf{C}_{n,n}$. Then $A \geq 0$ if and only if there exists a matrix $B \in \mathbf{C}_{n,n}$ such that $A = B^*B$.

Proof Suppose $A = B^*B$. Then $A^* = A$. If $x \in \mathbf{C}_{n,1}$, then $x^*Ax = x^*B^*Bx = \|B(x)\|^2 \geq 0$. Conversely, suppose $A \geq 0$. By the Spectral Theorem, there is a unitary matrix U such that

$$UAU^* = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n).$$

By Corollary 2.35, the λ 's are all nonnegative. Let $B = CU$, where $C = \text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n})$. Then $A = B^*B$. $\square$

COROLLARY 2.44 Let $A \in \mathbf{C}_{n,n}$ be a hermitian matrix. Then $A \geq 0$ if and only if the determinants of its principal submatrices are all nonnegative, and A is positive definite if and only if the determinants of its leading principal submatrices are all positive.

Proof It follows from Corollary 2.35 and the interlacing inequalities that each principal submatrix of a positive (semi)definite hermitian matrix is positive (semi)definite. Because $\det(B^*B) = |\det(B)|^2$, the necessity of the conditions is a consequence Lemma 2.43. Sufficiency can be proved using the interlacing inequalities, Corollary 2.35, and mathematical induction. $\square$

Suppose $A \in \mathbf{C}_{n,n}$. If $A \geq 0$ then, as in the proof of Lemma 2.43, there is a unitary matrix U such that $UAU^* = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$. If r is any positive real number, define the positive semidefinite matrix

$$A^r = U^* \text{diag}(\lambda_1^r, \lambda_2^r, \dots, \lambda_n^r)U. \quad (2.16)$$

Application to Graphs

The next definition opens the way for these ideas and techniques to be applied in the study of graphs.

DEFINITION 2.45 Let $G = (V, E)$ be a graph with vertex set $V = \{v_1, v_2, \dots, v_n\}$. Define

$$a_{ij} = \begin{cases} 1 & \text{if } \{v_i, v_j\} \in E, \\ 0 & \text{otherwise.} \end{cases}$$

The **adjacency matrix** of G is the n -by- n matrix $A(G) = (a_{ij})$. If $D(G) = \text{diag}(d(v_1), d(v_2), \dots, d(v_n))$ is the diagonal matrix of vertex degrees, then the **Laplacian matrix** of G is $L(G) = D(G) - A(G)$.

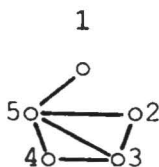


FIGURE 2.1

EXAMPLE 2.46 Let G be the graph illustrated in Figure 2.1. Then

$$L(G) = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 2 & -1 & 0 & -1 \\ 0 & -1 & 3 & -1 & -1 \\ 0 & 0 & -1 & 2 & -1 \\ -1 & -1 & -1 & -1 & 4 \end{pmatrix}.$$

□

Let $Y_n \in \mathbb{C}_{n,1}$ be the column n -tuple, each of whose entries is 1. Then $L(G)Y_n = 0$ expresses the fact that every row sum of $L(G)$ is zero. It also means that $L(G)$ is singular; for any graph G ,

$$\det(L(G)) = 0. \quad (2.17)$$

THEOREM 2.47 If G is a graph, then $L(G) \geq 0$.

Proof It is clear from the definitions that both $A(G)$ and $L(G)$ are symmetric and, therefore, hermitian. The proof that $L(G)$ is positive semidefinite is left to the exercises. □

Observe (Definition 2.45) that $L(G)$ depends not only on $G = (V, E)$, but on the numbering of the vertex set $V = \{v_1, v_2, \dots, v_n\}$. If L_1 is the version of $L(G)$ based on one numbering of the vertices and L_2 is the version based on another, then there exists an n -by- n permutation matrix P such that

$$L_2 = P^{-1}L_1P, \quad (2.18)$$

where, of course, $P^{-1} = P^* = P^t$. In fact, more is true.

THEOREM 2.48 Let G_1 and G_2 be graphs. Then G_1 is isomorphic to G_2 if and only if there is a permutation matrix P such that $L(G_2) = P^{-1}L(G_1)P$.

It follows from Theorem 2.48 that G_1 and G_2 are isomorphic only if $L(G_1)$ and $L(G_2)$ are (unitarily) similar. By the Spectral Theorem, two hermitian matrices

are unitarily similar if and only if they have the same eigenvalues. Evidently, the multiset of eigenvalues of $L(G)$ is a graph invariant.

If G is a graph on n vertices, let

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{n-1} \geq 0 = \lambda_n$$

be the eigenvalues of $L(G)$. Define $s(G) = (\lambda_1, \lambda_2, \dots, \lambda_{n-1})$, the $(n-1)$ -tuple obtained from $s(L(G))$ by deleting $\lambda_n = 0$. Then G_1 is isomorphic to G_2 only if $s(G_1) = s(G_2)$.

THEOREM 2.49 *If G is a graph, then $s(G) \succ d(G)$, that is, the spectrum of $L(G)$ majorizes the degree sequence of G .⁷*

Proof This is an immediate consequence of Theorem 2.38 and the definitions. $\square$

EXAMPLE 2.50 Let G be the graph in Figure 2.1. Then $d(G) = (4, 3, 2, 2, 1)$, $s(L(G)) = (5, 4, 2, 1, 0)$, and $s(G) = (5, 4, 2, 1)$. Observe that

$$5 \geq 4,$$

$$5 + 4 \geq 4 + 3,$$

$$5 + 4 + 2 \geq 4 + 3 + 2,$$

$$5 + 4 + 2 + 1 \geq 4 + 3 + 2 + 2,$$

and $5 + 4 + 2 + 1 = 4 + 3 + 2 + 2 + 1$. $\square$

Recall (Theorem 1.37) that the conjugate degree sequence, $d^*(G)$, also majorizes $d(G)$. In fact, there is some empirical evidence for the following:

CONJECTURE 2.51 *Let G be a graph. Then $d^*(G) \succ s(G)$, that is, the conjugate degree sequence majorizes the Laplacian spectrum.*

In Example 2.50, $d(G) = (4, 3, 2, 2, 1)$. Therefore, $d^*(G) = (5, 4, 2, 1) = s(G)$, affording a case of equality in Conjecture 2.51. As we now see, this is no coincidence. It follows from the fact that $d(G)$ is a maximal graphic partition.

THEOREM 2.52⁸ *Let G be a connected graph. Then $d^*(G) = s(G)$ if and only if G is a threshold graph.*

⁷ An improvement of this result can be found in [Grone (1995)].

⁸ Theorem 2.52 was first proved in [Merris (1994a)]. Threshold graphs were introduced independently by [Chvátal & Hammer (1977)] and [Henderson & Zalcstein (1979)]. Consult [Arikati & Peled (1994)], [Golumbic (1978)], [Hammer, Ibaraki & Simeone (1981)], [Orlin (1977)], [Peled (1980)], and [Ruch & Gutman (1979)] for more about these "maximal" graphs.

Exercises

1. Prove that $\hat{u}: V' \rightarrow \mathbb{C}$, defined by Equation (2.3), is a linear functional on V' .
2. Prove that the mapping Ψ in Theorem 2.2 is a vector space isomorphism.
3. Let $(,)$ be an inner product on V . Let W be a subspace of V . Explain why the restriction of $(,)$ to W is an inner product on W .
4. Prove that Equation (2.4) defines an inner product on $\mathbb{C}^n$.
5. Let V be a (complex) vector space of dimension n . Explain how Example 2.5 can be used to construct an inner product on V . (Hint: Start with a fixed but arbitrary basis of V .)
6. Suppose that $S = \{v_1, v_2, \dots, v_k\}$ is an orthogonal set of nonzero vectors. Prove that S is linearly independent.
7. Prove Theorem 2.11.
8. If W is a subspace of the inner product space V , show that $(W^\perp)^\perp = W$.
9. Exhibit the orthonormal basis of $\mathbb{C}^3$ obtained by applying the Gram-Schmidt process to the basis $\{(0, 1, 1), (1, 0, 1), (1, 1, 0)\}$.
10. What is the result of applying the Gram-Schmidt process to the linearly dependent set $\{(1, 0), (0, 1), (1, 1)\}$?
11. Let w be the vector defined in Equation (2.6). Prove that $(w, u) = 0$.
12. Let v and $u \neq 0$ be linearly *dependent* vectors in an inner product space V . Prove that

$$v = \frac{(v, u)}{\|u\|^2} u.$$

13. Prove Bessel's Inequality.
14. Prove that $\operatorname{tr}(A^*A) \operatorname{tr}(B^*B) \geq |\operatorname{tr}(A^*B)|^2$, for all $A, B \in \mathbb{C}_{m,n}$.
15. Let V be an inner product space of dimension n . Suppose $A = (a_{ij}) \in \mathbb{C}_{n,n}$. Prove that there exist vectors $v_1, v_2, \dots, v_n$ and $w_1, w_2, \dots, w_n$ such that $a_{ij} = (v_i, w_j)$.
16. Let V be an inner product space. If $S, T \in L(V, V)$, prove that $(ST)^* = T^*S^*$.
17. Let T be a projection on V . Prove that $\dim(T(V)) = \operatorname{tr}(T)$. (Since the trace of a matrix is preserved under similarity, we can unambiguously define the trace of T to be the trace of the matrix representation of T with respect to any fixed basis B of V . Another approach would be to define the trace of T to be the sum of its eigenvalues.)

18. Let W be a subspace of the inner product space V .
 - a. Prove that there exists an orthogonal projection $P \in L(V, V)$ such that $W = P(V)$.
 - b. Let $T \in L(V, V)$. If T commutes with P , show that T holds W invariant, that is, $T(w) \in W$ for all $w \in W$.
 - c. Suppose T holds W invariant. Must T commute with P ?
19. Suppose $V = W_1 + W_2$. Prove that $V = W_1 \oplus W_2$ if and only if $\dim(V) = \dim(W_1) + \dim(W_2)$.
20. Let $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ be a basis of the inner product space V . Suppose $S, T \in L(V, V)$. If $(T(v_i), v_j) = (v_i, S(v_j))$, $1 \leq i, j \leq n$, prove that $S = T^*$.
21. Let T be a linear operator on the inner product space V . Prove $(T^*)^* = T$, that is, prove $(T^*(v), w) = (v, T(w))$, for all $v, w \in V$.
22. Let $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ be a basis of V . Prove that there is an inner product on V with respect to which $\mathcal{B}$ is orthonormal.
23. Prove Lemma 2.17.
24. Let $\mathcal{B} = \{u_1, u_2, \dots, u_n\}$ be an o.n. basis of the inner product space V . If $v = \sum a_i u_i$, let $[v] \in \mathbb{C}_{n,1}$ be the column vector whose i -th entry is a_i , that is, $[v]$ is the coordinate representation of v with respect to $\mathcal{B}$.
 - a. If $T \in L(V, V)$, prove that $(T(u), v) = [v]^*[T][u]$, where $[T]$ is the matrix representation of T with respect to $\mathcal{B}$ (Example 2.19).
 - b. Prove Lemma 2.27.
25. Prove that $T \in L(V, V)$ is positive definite hermitian if and only if it is positive semidefinite hermitian and invertible.
26. Show that every orthogonal projection is positive semidefinite.
27. Suppose $A \in \mathbb{C}_{2,2}$.
 - a. If $x^*Ax \geq 0$ for all $x \in \mathbb{C}_{2,1}$, prove that A is hermitian.
 - b. If A is real and $x^*Ax \geq 0$ for all real 2-by-1 matrices x , must A be symmetric?
28. Let $\{u_1, u_2, \dots, u_n\}$ be an orthonormal basis of the inner product space V . Exhibit a hermitian operator $T \in L(V, V)$ that satisfies $(T(u_i), u_i) > 0$, $1 \leq i \leq n$, but such that T is not positive semidefinite.
29. Prove that $U \in \mathbb{C}_{n,n}$ is unitary if and only if its columns are orthonormal with respect to the standard inner product of $\mathbb{C}^n$.
30. Let $\mathcal{B}$ be an o.n. basis of the inner product space V . Let $[T]$ be the matrix representation of $T \in L(V, V)$ with respect to $\mathcal{B}$. Prove that T is unitary if and only if $[T]$ is unitary.

31. Let T be a unitary operator on the inner product space V . Suppose $\{u_1, u_2, \dots, u_n\}$ is an orthonormal set of vectors in V . Prove that $\{T(u_1), T(u_2), \dots, T(u_n)\}$ is orthonormal.
32. Let T be a linear operator on the inner product space V . Prove that there exists an orthonormal basis of V with respect to which the matrix representation of T is upper triangular.
33. Prove Corollary 2.32.
34. Let $A, B \in \mathbb{C}_{n,n}$ be a pair of commuting matrices (that is, $AB = BA$). Prove that there exists a unitary matrix $U \in \mathbb{C}_{n,n}$ such that U^*AU and U^*BU are both upper triangular. (Hint: Prove that A and B have a common eigenvector.)
35. Give a proof of the Spectral Theorem that does not rely on Exercise 34. (Hint: If A is normal and U is unitary, show that U^*AU is normal. Then prove that a normal, upper triangular matrix is diagonal.)
36. Prove that a normal matrix is hermitian if and only if its eigenvalues are real.
37. Let T be a hermitian operator on the inner product space V . If the eigenvalues of T are (all) nonnegative, prove that T is positive semidefinite.
38. Prove that a normal matrix is unitary if and only if its eigenvalues all have absolute value 1.
39. Let T be a linear operator on the inner product space V . Prove that $T \geq 0$ if and only if there exists an $S \in L(V, V)$ such that $T = S^*S$.
40. Let V be an inner product space. Suppose $T \in L(V, V)$. If $(T(v), v) \geq 0$ for all $v \in V$, show that $T = T^*$. (Thus, the phrase “positive semidefinite hermitian” is redundant.)
41. Let $v_1, v_2, \dots, v_n$ be vectors in an inner product space V . Define $a_{ij} = (v_i, v_j)$, $1 \leq i, j \leq n$. Then $A = (a_{ij})$ is the **Gram matrix** based on $v_1, v_2, \dots, v_n$.
 - a. Show that $A \geq 0$ if and only if A is a gram matrix.
 - b. Show that A is positive definite hermitian if and only if it is a gram matrix based on linearly independent vectors.
42. Let k be a positive integer. Suppose $r = 1/k$. Show that
 - a. $(A^r)^k = A$.
 - b. $(A^k)^r = A$.
43. Prove that $A \in \mathbb{C}_{n,n}$ is positive semidefinite hermitian if and only if there is a *lower triangular* matrix B such that $A = B^*B$.
44. Prove that $A \in \mathbb{C}_{n,n}$ is positive semidefinite hermitian if and only if there is an *upper triangular* matrix B such that $A = B^*B$.

45. Suppose $A = (a_{ij}) \in \mathbf{C}_{n,n}$ is positive definite hermitian. Define $d_i = (a_{ii})^{-1/2}$, $1 \leq i \leq n$, $D = \text{diag}(d_1, d_2, \dots, d_n)$, and $B = DAD$.
- Show that $(\det(A))^{1/n} \leq (\text{tr}(A))/n$. (Hint: Use eigenvalues and the arithmetic-geometric mean inequality.)
 - Show that the “correlation matrix” $B = DAD$ is positive definite hermitian.
 - Show that $\det(A) \leq h(A)$ if and only if $\det(B) \leq 1$.
 - Show that $\det(B) \leq 1$.
46. Suppose $A = (a_{ij}) \geq 0$. Prove that
- $a_{ii} \geq 0$, $1 \leq i \leq n$.
 - $a_{ii}a_{jj} \geq |a_{ij}|^2$, $1 \leq i, j \leq n$.
 - if $a_{ii} = 0$, then every entry in the i -th row and column of A is zero.
47. Let $S, T \in L(V, V)$ be orthogonal projections. Show that $S \geq T$ if and only if $ST = T$.
48. Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ be the eigenvalues of a hermitian matrix $A \in \mathbf{C}_{n,n}$. Show that $\lambda_1 I_n - A \geq 0$ and $A - \lambda_n I_n \geq 0$.
49. Suppose $A \geq B$. If $\text{tr}(A) = \text{tr}(B)$, prove that $A = B$.
50. Let $A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$. Show that $F(A)$ is not the convex hull of the eigenvalues of A .
51. Suppose $A \in \mathbf{C}_{n,n}$ and $\alpha \in \mathcal{Q}_{m,n}$. Denote by $A[\alpha|\alpha]$ the principal m -by- m submatrix of A whose (i, j) -entry is the $(\alpha(i), \alpha(j))$ -entry of A . Prove that $(A^*A)[\alpha|\alpha] \geq A[\alpha|\alpha]^*A[\alpha|\alpha]$.
52. Suppose $0 < \varepsilon < \sqrt{2} - 1$. Let $A = \text{diag}(1 + \varepsilon, 1 + \varepsilon^{-1})$.
- Prove that $A \geq J$, the 2-by-2 matrix each of whose entries is 1.
 - Prove that $A^2 \not\geq J^2$. (Necessary and sufficient conditions for $A^k \geq B^k$ to hold for all positive integers k are given in [Kwong (1977)].)
 - Prove that $A^{1/2} \geq J^{1/2}$. (More generally, it is known [Au-Yeung (1973)] that $A \geq B$ implies $A^\theta \geq B^\theta$ whenever $0 < \theta < 1$.)
53. Suppose $A \in \mathbf{C}_{m,m}$ and $B \in \mathbf{C}_{m,n}$. If $A \geq 0$, prove that $B^*AB \geq 0$.
54. Suppose $A \in \mathbf{C}_{n,n}$ is partitioned as

$$A = \begin{pmatrix} P & Q \\ R & S \end{pmatrix},$$

where P and S are square and P is invertible. Then the Schur Complement of P in A is $A/P = S - RP^{-1}Q$.

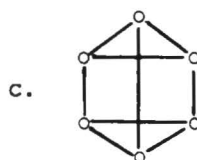
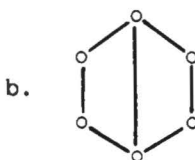
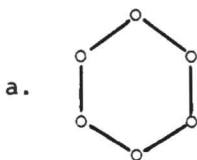
- a. Prove that⁹ $\det(A) = \det(P) \det(A/P)$. (Hint: Compute XAY , where

$$X = \begin{pmatrix} I & 0 \\ -RP^{-1} & I \end{pmatrix} \quad \text{and} \quad Y = \begin{pmatrix} I & -P^{-1}Q \\ 0 & I \end{pmatrix}.)$$

- b. If $A \geq 0$, prove that $A/P \geq 0$. (Hint: If $A \geq 0$, then $R = Q^*$; use part (a) and Exercise 53.)

Application Exercises

55. Let $G = (V, E)$ be a graph with vertex set $V = \{v_1, v_2, \dots, v_n\}$ and Laplacian matrix $L(G)$.
- a. Prove that $xL(G)x^* = \sum |x_i - x_j|^2$, where the sum is over the edges of G , that is, over those ordered pairs (i, j) such that $i < j$ and $\{v_i, v_j\} \in E$.
- b. Prove that $L(G) \geq 0$.
56. Let G be the graph in Example 2.50 (Figure 2.1). Confirm that $s(G) = (5, 4, 2, 1)$.
57. Prove Theorem 2.48.
58. Confirm that $s(G)$ majorizes $d(G)$ if G is the graph



59. Let G be a graph with vertex set $V = \{v_1, v_2, \dots, v_n\}$ and edge set $E = \{e_1, e_2, \dots, e_m\}$. For each edge $e_j = \{v_i, v_k\}$, choose one of v_i, v_k to be the “positive end” of e_j , and the other to be the “negative end”. Thus, G is given an **orientation**. For a fixed but arbitrary orientation of G , define the n -by- m matrix $Q = Q(G) = (q_{ij})$ by

$$q_{ij} = \begin{cases} +1, & \text{if } v_i \text{ is the positive end of } e_j, \\ -1, & \text{if it is the negative end, and} \\ 0, & \text{otherwise.} \end{cases}$$

⁹This identity appeared in [Schur (1917)]. The name “Schur Complement” was introduced in [Haynsworth (1968)]. Also see [Cottle (1974)].

Then Q is an oriented vertex-edge incidence matrix for the graph G .

- a. Prove that $QQ^t = L(G)$, independently of the orientation.
 - b. Prove that $L(G) \geq 0$.
60. Confirm the sufficiency part of Theorem 2.52 for the graphs in Figure 1.12.
61. Let G be a graph. Show that $L(G)$ commutes with $L(G^c)$.
62. Suppose $s(G) = (\lambda_1, \lambda_2, \dots, \lambda_{n-1})$ and $s(G^c) = (\mu_1, \mu_2, \dots, \mu_{n-1})$. Show that $\lambda_i + \mu_{n-i} = n$, $1 \leq i < n$. (Hint: Exercises 34 and 61.)

CHAPTER 3

Permutation Groups

A **permutation of degree n** is a one-to-one function from $\{1, 2, \dots, n\}$ onto itself. The set of all $n!$ permutations of degree n is denoted S_n . Under the operation of function composition, S_n is a group, the **symmetric group of degree n** . The identity of this group is the permutation e_n defined by $e_n(x) = x$ for all $x \in \{1, 2, \dots, n\}$. When n is understood from the context, e may be used in place of e_n .¹

Let $x \in \{1, 2, \dots, n\}$ and $\sigma \in S_n$ be fixed but arbitrary. If $\sigma(x) = x$, then x is a **fixed point** of σ . More generally, let $x = x_1$ and define $x_{i+1} = \sigma(x_i) = \sigma^i(x)$, $i \geq 1$. If k is the smallest positive integer such that $\sigma^k(x) = x$, then

$$C_\sigma(x) = (x_1 x_2 \dots x_k)$$

is the **cycle of σ containing x** . The integer k is the **length** of $C_\sigma(x)$, and $C_\sigma(x)$ is sometimes called a k -cycle. Thus, x is a fixed point of σ if and only if the length of $C_\sigma(x)$ is 1. If $j = i + 1$ or if $i = k$ and $j = 1$, then $t = x_j$ is said to *follow* $s = x_i$ in $C_\sigma(x)$. Cycles $C_\sigma(x)$ and $C_\tau(y)$ are **equivalent** if they have the same length and the same integers in the same (cyclical) order. So, the phrase “ $C_\sigma(x)$ and $C_\tau(y)$ are equivalent” means that t follows s in $C_\sigma(x)$ if and only if t follows s in $C_\tau(y)$, for every choice of s and t in $\{1, 2, \dots, n\}$. If $\{\sigma^i(x) : i \geq 1\} \cap \{\tau^i(y) : i \geq 1\} = \emptyset$, then the cycles $C_\sigma(x)$ and $C_\tau(y)$ are **disjoint**.

EXAMPLE 3.1 Let $\sigma \in S_6$ be the permutation defined by $\sigma(1) = 5$, $\sigma(2) = 4$, $\sigma(3) = 1$, $\sigma(4) = 2$, $\sigma(5) = 3$, and $\sigma(6) = 6$. Then 6 is a fixed point of σ , and $C_\sigma(6) = (6)$. While $C_\sigma(1) = (153)$, $C_\sigma(3) = (315)$, and $C_\sigma(5) = (531)$ are

¹The letter e has already been used to denote an edge of a graph; e_n will be used later to denote the n -th vector in a basis. It is to be hoped that the context will eliminate any confusion these notational abuses might otherwise cause.

equivalent to each other, they are not equivalent to (135). Finally, $C_\sigma(2) = (24)$ and $C_\sigma(4) = (42)$ are equivalent to each other and disjoint from $C_\sigma(x)$ when $2 \neq x \neq 4$. $\square$

LEMMA 3.2 Suppose $x, y \in \{1, 2, \dots, n\}$ and $\sigma, \tau \in S_n$. Then

- (a) either $C_\sigma(x)$ and $C_\sigma(y)$ are disjoint or they are equivalent; and
- (b) either $C_\sigma(x)$ and $C_\tau(x)$ are identical or they are inequivalent.

DEFINITION 3.3 Let $C_\sigma(x), C_\sigma(y), \dots, C_\sigma(z)$, be the inequivalent cycles of $\sigma \in S_n$. Then the **disjoint cycle factorization**² of σ is $C_\sigma(x)C_\sigma(y) \dots C_\sigma(z)$.

Suppose that σ is the permutation in Example 3.1. Because there are three (equivalent) ways to express the 3-cycle, two ways to express $(24) = (42)$, and because the three cycles can be written in any order, there are $3 \times 2 \times (3!) = 36$ different *looking* ways to express its disjoint cycle factorization. (Some examples are: $(153)(24)(6)$, $(315)(6)(24)$, $(42)(531)(6)$, and so on.) Apart from equivalence and the order in which the cycles are written, however, the disjoint cycle factorization of a permutation is unique. Moreover, it is customary when expressing permutations using disjoint cycle notation to suppress the cycles of length 1 (corresponding to fixed points). In particular, the permutation from Example 3.1 is typically written $\sigma = (153)(24)$.

DEFINITION 3.4 The **cycle type** of $\sigma \in S_n$ is the partition of n whose parts are the lengths of the cycles in its disjoint cycle factorization. Two permutations are said to have the same **cycle structure** if their cycle types are the same.

EXAMPLE 3.5 Consider the permutation $\sigma = (1357)(246) \in S_7$ with cycle type $[4,3]$. As we know, $(1357)(246)$ is just one of $4 \times 3 \times 2 = 24$ different looking ways to express the disjoint cycle factorization of σ . We now consider a different question, namely, the number of permutations in S_7 that have cycle type $[4,3]$. Any such permutation can be written in the form $\tau = (abcd)(xyz)$. While $7 \times 6 \times 5 \times 4 = 840$ is the number of ways to “fill up” the 4-cycle $(abcd)$, it is not the number of ways to choose the 4-cycle; it is too big. We have not taken equivalence into account. Because $(abcd) = (bcda) = (cdab) = (dabc)$, there are only $840/4 = 210$ different 4-cycles. Once the 4-cycle is chosen, three numbers remain to play the roles of x, y , and z . These can be arranged in 3-cycles in $3 \times 2 \times 1/3 = 2$ inequivalent ways (namely (xyz) or (xzy)). Therefore, S_7 contains a total of $210 \times 2 = 420$ permutations of type $[4,3]$. $\square$

Note that the 420 permutations enumerated in Example 3.5 have the same cycle

²Despite the fact that the binary operation of S_n is function composition, we will adopt the generic language and speak of it as a “product”.

structure as $(abc)(wxyz)$. Indeed,

$$\frac{7 \times 6 \times 5 \times 4}{4} \times \frac{3 \times 2 \times 1}{3} = \frac{7 \times 6 \times 5}{3} \times \frac{4 \times 3 \times 2 \times 1}{4}.$$

EXAMPLE 3.6 One of the 77 partitions of 12 is $[3^2, 2^3]$. How many permutations in S_{12} have cycle type $[3, 3, 2, 2, 2]$? The generic permutation with this cycle type may be written

$$\sigma = (abc)(xyz)(pq)(rs)(tu).$$

There are $12 \times 11 \times 10/3 = 440$ ways to choose the first 3-cycle. Once it has been chosen, there are $9 \times 8 \times 7/3 = 168$ ways to choose the second. So, $440 \times 168 = 73,920$ is the number of ways to choose a sequence of two 3-cycles. There is a subtle point here that did not arise in our previous discussion: We have double counted the pairs of 3-cycles that could occur in our generic σ , once in the form $(abc)(xyz)$ and again as $(xyz)(abc)$. Compensating for this double counting, we conclude that the pair of 3-cycles can be chosen in $73,920/2 = 36,960$ different ways. (While it is true that $(abcd)(xyz) = (xyz)(abcd)$, we never counted these expressions as different in Example 3.5.)

No matter which six numbers are used in the two 3-cycles, six numbers remain to be distributed among the three 2-cycles. We may choose (pq) in $6 \times 5/2 = 15$ ways; (rs) in $4 \times 3/2 = 6$ ways; and (tu) in $2 \times 1/2 = 1$ way. There would be $15 \times 6 \times 1 = 90$ ways to choose the three 2-cycles if, for example, $(pq)(rs)(tu)$ were different from $(rs)(pq)(tu)$. Because the order in which the 2-cycles are written does not matter, we have counted each triple of 2-cycles $3! = 6$ times. Once the two 3-cycles have been chosen, there are only $90/6 = 15$ ways to choose the three 2-cycles. Hence, the number of permutations in S_{12} of cycle type $[3^2, 2^3]$ is $36,960 \times 15 = 554,400$. (So, fewer than 0.12% of the 479,001,600 permutations in S_{12} have cycle type $[3^2, 2^3]$.) $\square$

The evident generalization of these examples involves permutations of cycle type $\pi = [n^{r_1}, \dots, 3^{r_3}, 2^{r_2}, 1^{r_1}] \vdash n$, an awkward thing to write down. To simplify such expressions, we will sometimes abuse the language and reverse the (usual decreasing) order of the parts, writing instead $\pi = [1^{r_1}, 2^{r_2}, \dots] \vdash n$.

THEOREM 3.7 Let $\pi = [1^{r_1}, 2^{r_2}, \dots] \vdash n$. Then the number of permutations in S_n of cycle type π is

$$n(\pi) = \frac{n!}{1^{r_1}(r_1!)2^{r_2}(r_2!)3^{r_3}(r_3!) \dots n^{r_n}(r_n!)}.$$

Recall that the length of the cycle $C_\sigma(x)$ is the smallest positive integer k such that $\sigma^k(x) = x$. In this context, x is fixed, and k depends on x . We are now interested in the smallest positive integer k such that $\sigma^k(x) = x$, for all $x \in \{1, 2, \dots, n\}$.

DEFINITION 3.8 Let $\sigma \in S_n$. The smallest positive integer k such that $\sigma^k = e_n$ is called the **order** of σ , denoted $o(\sigma)$.

EXAMPLE 3.9 Let $\sigma = (123)(45) \in S_5$. Then $\sigma^2 = (132)$, $\sigma^3 = (45)$, $\sigma^4 = (123)$, $\sigma^5 = (132)(45)$, and $\sigma^6 = e_5$. The degree of σ is 5 and its order is 6. $\square$

THEOREM 3.10 The order of $\sigma \in S_n$ is the least common multiple of the lengths of the cycles in its disjoint cycle factorization.

DEFINITION 3.11 A **transposition** is a permutation of cycle type $[2, 1^{n-2}]$ (or $[1^{n-2}, 2]$).

THEOREM 3.12 Every permutation $\sigma \in S_n$ can be written as a product of (not necessarily disjoint) transpositions. (The factorization into a product of transpositions may not be unique.)

Proof It suffices to show that a k -cycle can be expressed as a product of transpositions: Observe that

$$\begin{aligned}(x_1 x_2 \dots x_k) &= (x_1 x_2)(x_2 x_3) \dots (x_{k-1} x_k) \\ &= (x_1 x_k)(x_1 x_{k-1}) \dots (x_1 x_2).\end{aligned}\tag{3.1}$$

$\square$

THEOREM 3.13 Let $\sigma_1, \sigma_2, \dots, \sigma_r$ and $\tau_1, \tau_2, \dots, \tau_s$ be transpositions. If $\sigma_1 \sigma_2 \dots \sigma_r = \tau_1 \tau_2 \dots \tau_s$, then $(-1)^r = (-1)^s$, that is, r and s are either both odd or both even.

Proof Because (Theorem 3.10) the order of a transposition is 2, every transposition is its own inverse. If $\sigma_1 \sigma_2 \dots \sigma_r = \tau_1 \tau_2 \dots \tau_s$, then the identity permutation

$$e = \sigma_1 \sigma_2 \dots \sigma_r \tau_s \tau_{s-1} \dots \tau_1$$

is a product of $r + s$ transpositions. Suppose $r + s$ is odd. Let t be minimal such that t is odd, and the identity can be written as $e = \mu_1 \mu_2 \dots \mu_t$, a product of t transpositions. Evidently, $t > 1$. Suppose $\mu_t = (xy)$. Let m be the largest integer less than t such that $\mu_m = (xz)$. (Such an m must exist. Otherwise, $e(y) = \mu_1 \mu_2 \dots \mu_t(y) = \mu_1 \mu_2 \dots \mu_{t-1}(x) = x$.) If $m < t - 1$, let $\mu_{m+1} = (pq)$, where $p \neq x \neq q$. If $p \neq z \neq q$, then $\mu_m \mu_{m+1} = \mu_{m+1} \mu_m$. If $p = z$, then $\mu_m \mu_{m+1} = (xz)(zq) = (zq)(xq)$. Thus, without loss of generality we may assume

$$e = \mu_1 \mu_2 \dots \mu_k (xy_1)(xy_2) \dots (xy_{t-k}),$$

where x is fixed by $\mu_1, \mu_2, \dots, \mu_k$.

Now, there is an integer $m < t - k$ such that $y_m = y_{t-k}$. Otherwise, x is sent by the identity to $\mu_1\mu_2\ldots\mu_k(y_{t-k})$, which cannot be x because x is fixed by $\mu_1, \mu_2, \ldots, \mu_k$. Let m be maximal such that $y_m = y_{t-k}$. If $m + 1 < t - k$, then $(xy_m)(xy_{m+1}) = (y_my_{m+1})(xy_m)$. Therefore, after some more rearranging, we may assume

$$\begin{aligned} e &= \mu_1\mu_2\ldots\mu_{t-2}(xy)(xy) \\ &= \mu_1\mu_2\ldots\mu_{t-2}. \end{aligned}$$

This is a contradiction because $t - 2$ is odd and less than t . $\square$

DEFINITION 3.14 A permutation $\sigma \in S_n$ that can be written as a product of an even number of transpositions is said to be **even**. If σ is not even, then it is **odd**. The **alternating character** of S_n is defined by

$$\varepsilon(\sigma) = \begin{cases} 1 & \text{if } \sigma \text{ is even} \\ -1 & \text{if } \sigma \text{ is odd} \end{cases}.$$

Observe that $\varepsilon: S_n \rightarrow \mathbb{C}$ is a homomorphism. Its kernel, A_n , is the **alternating group** of degree n . Thus, A_n is a normal subgroup of S_n consisting of all the even permutations.

DEFINITION 3.15 Let $c_t(\sigma)$ be the number of cycles of length t , and

$$c(\sigma) = \sum_{t=1}^n c_t(\sigma)$$

the total number of cycles in the disjoint cycle factorization of $\sigma \in S_n$.

Using the notation introduced in Definition 3.15, the cycle type of σ is $[1^{c_1(\sigma)}, 2^{c_2(\sigma)}, \ldots]$. In particular,

$$\sum_{t=1}^n t c_t(\sigma) = n. \quad (3.2)$$

Note that $c_1(\sigma)$ (not to be confused with $C_\sigma(1)$) is the number of fixed points of σ .

THEOREM 3.16 If $\sigma \in S_n$, then $\varepsilon(\sigma) = (-1)^{n-c(\sigma)}$.

Proof From Equation (3.1), each k -cycle of σ is the product of $k - 1$ transpositions. Summing the numbers of transpositions over all $c(\sigma)$ cycles of σ (including the cycles of length 1), we deduce that σ is the product of $n - c(\sigma)$ transpositions. Combined with Theorem 3.13 and Definition 3.14, this completes the proof. $\square$

EXAMPLE 3.17 If $\sigma = (12345)(6789)$, then

$$(24)\sigma(24) = (14325)(6789)$$

and

$$(27)\sigma(27) = (17345)(6289).$$

□

THEOREM 3.18 Two permutations, $\sigma, \mu \in S_n$ have the same cycle structure if and only if they are conjugate in S_n , that is, if and only if there is a permutation $\tau \in S_n$ such that $\tau^{-1}\sigma\tau = \mu$.

Proof As in Example 3.17, $(xy)\sigma(xy)$ is the permutation obtained from σ by interchanging the positions of x and y in its disjoint cycle factorization. Now, σ and μ have the same cycle structure if and only if μ can be obtained from σ by a sequence of interchanges, if and only if there exist transpositions $\tau_1, \tau_2, \dots, \tau_r$ such that

$$(\tau_r \tau_{r-1} \dots \tau_1)\sigma(\tau_1 \tau_2 \dots \tau_r) = \mu,$$

if and only if $\tau^{-1}\sigma\tau = \mu$, where $\tau = \tau_1 \tau_2 \dots \tau_r$. The result follows from Theorem 3.12. □

DEFINITION 3.19 A permutation group of degree n is a subgroup of S_n .

DEFINITION 3.20 Let G be permutation group of degree n . Then $x, y \in \{1, 2, \dots, n\}$ are **equivalent modulo G** , written $x \equiv y \pmod{G}$, if there exists a $\sigma \in G$ such that $\sigma(x) = y$.

THEOREM 3.21 For any subgroup G of S_n , equivalence modulo G is an equivalence relation.

It follows from Theorem 3.21, and the general theory of equivalence relations, that equivalence modulo G partitions $\{1, 2, \dots, n\}$ into a disjoint union of equivalence classes.

DEFINITION 3.22 The equivalence classes of $\{1, 2, \dots, n\}$ modulo G are called **orbits** of G . The orbit of G to which x belongs is $O_x = \{\sigma(x) : \sigma \in G\}$. If $x \equiv y \pmod{G}$ for all $x, y \in \{1, 2, \dots, n\}$, that is, if G has just one orbit, then G is said to be **transitive**.

EXAMPLE 3.23 Evidently, S_n is transitive for all n . If $G = \{e_4, (12)(34), (13)(24), (14)(23)\}$ and $H = \{e_4, (12), (34), (12)(34)\}$, then G and H are isomorphic groups. However, G is transitive but H is not; the orbits of H are $O_1 = \{1, 2\} = O_2$ and $O_3 = \{3, 4\} = O_4$. This shows that transitivity is not a group invariant. □

DEFINITION 3.24 Let G be a subgroup of S_n . For each $x \in \{1, 2, \dots, n\}$, the **stabilizer subgroup** $G_x = \{\sigma \in G : \sigma(x) = x\}$.

EXAMPLE 3.25 Let $G = \langle (12)(345) \rangle = \{e_5, (12)(345), (354), (12), (345), (12)(354)\}$, the cyclic group generated by $(12)(345)$. Then $G_2 = \{e_5, (345), (354)\} = \langle (345) \rangle$. Note that O_x is a set of numbers while G_x is a set of permutations. In this case, $O_2 = \{\sigma(2) : \sigma \in G\} = \{2, 1, 2, 1, 2, 1\}$, and $O_3 = \{\sigma(3) : \sigma \in G\} = \{3, 4, 5, 3, 4, 5\}$, multiplicities included. $\square$

LEMMA 3.26 Let G be a permutation group of degree n . If $x \in \{1, 2, \dots, n\}$ then $o(O_x) = o(G)/o(G_x)$. That is, the cardinality of the orbit of x is the index of its stabilizer subgroup.

Proof As in Example 3.25, $O_x = \{\sigma(x) : \sigma \in G\}$ contains $o(G)$ elements, multiplicities included. By definition, x occurs in O_x with multiplicity $o(G_x)$. If $y \in O_x$, then there exists a permutation $\tau \in G$ such that $\tau(x) = y$. It follows that $\sigma(x) = y$ for all σ in the coset $\tau G_x = \{\tau\mu : \mu \in G_x\}$. Conversely, if $\sigma(x) = y = \tau(x)$, then $\tau^{-1}\sigma(x) = x$, in which case $\tau^{-1}\sigma \in G_x$, and $\sigma \in \tau G_x$. In other words, $\{\sigma \in G : \sigma(x) = y\} = \tau G_x$. It follows that y occurs in O_x with multiplicity $o(\tau G_x) = o(G_x)$. Evidently, every element of $O_x = \{\sigma(x) : \sigma \in G\}$ occurs with the same multiplicity. Therefore, $o(O_x) = o(G)/o(G_x)$. $\square$

DEFINITION 3.27 Denote by $F(\sigma)$ the number of **fixed points** of the permutation σ .

Evidently, $F(\sigma) = c_1(\sigma)$, the number of cycles of length one in the disjoint cycle factorization of σ . However, simplifying some of the expressions that come later is worth a little notational redundancy.

BURNSIDE'S LEMMA 3.28 Let G be a permutation group of degree n , affording t orbits. Then t is the average of the numbers of fixed points of the permutations in G , that is,

$$\frac{1}{o(G)} \sum_{\sigma \in G} F(\sigma) = t.$$

Proof Consider the set $S = \{(\sigma, x) : \sigma(x) = x\}$. Of the ordered pairs in S , $F(\sigma)$ begin with σ and $o(G_x)$ end with x . Thus, counting S in two different ways, we obtain

$$\sum_{\sigma \in G} F(\sigma) = \sum_{x=1}^n o(G_x).$$

Applying Lemma 3.26 to the right-hand side of this equation produces

$$\sum_{\sigma \in G} F(\sigma) = o(G) \sum_{x=1}^n 1/o(O_x).$$

If Z is a system of distinct representatives for the orbits of G , then $o(Z) = t$ and

$$\begin{aligned} \sum_{\sigma \in G} F(\sigma) &= o(G) \sum_{y \in Z} \sum_{x \in O_y} 1/o(O_x) \\ &= o(G) \sum_{y \in Z} 1 \\ &= to(G). \end{aligned}$$

□

EXAMPLES 3.29 Suppose $H = \{e_4, (12), (34), (12)(34)\}$. Averaging the numbers of fixed points over the permutations in H , we obtain $(4 + 2 + 2 + 0)/4 = 2$, confirming Example 3.23.

Because S_n is transitive it must be that, on the average, permutations in S_n have one fixed point. □

COROLLARY 3.30 Let G be a permutation group of degree n . Then

$$\frac{1}{o(G)} \sum_{\sigma \in G} F(\sigma) \geq 1,$$

with equality if and only if G is transitive.

Proof This is an immediate consequence of Burnside's Lemma; $t = 1$ if and only if G is transitive. □

There is a surprising analog of Corollary 3.30 for doubly transitive groups.

DEFINITION 3.31 Let G be a permutation group of degree n . Then G is **doubly transitive** if, for all $x_1, x_2, y_1, y_2 \in \{1, 2, \dots, n\}$ satisfying $x_1 \neq x_2$ and $y_1 \neq y_2$, there exists a permutation $\sigma \in G$ such that $\sigma(x_1) = y_1$ and $\sigma(x_2) = y_2$.

THEOREM 3.32 Let G be a permutation group of degree $n > 1$. Then

$$\frac{1}{o(G)} \sum_{\sigma \in G} F(\sigma)^2 \geq 2, \quad (3.3)$$

with equality if and only if G is doubly transitive.

Proof Because the result is trivial when $n = 2$, we may assume $n \geq 3$. As in the proof of Burnside's Lemma, we count a certain set S two different ways. This time, $S = \{(\sigma, x, y) : \sigma(x) = x \text{ and } \sigma(y) = y\}$. Of the 3-tuples in S , $F(\sigma)^2$ begin with σ . On the other hand, any 3-tuple that ends with y must begin with a $\sigma \in G_y$. Moreover, for any such ending and beginning, there will be $F(\sigma)$ middle entries. Therefore,

$$\sum_{\sigma \in G} F(\sigma)^2 = \sum_{y=1}^n \sum_{\sigma \in G_y} F(\sigma). \quad (3.4)$$

Every $\sigma \in G_y$ has at least one fixed point, namely y . Therefore, $F(\sigma) \geq 1$ for all $\sigma \in G_y$. Let $F_y(\sigma) = F(\sigma) - 1$, $\sigma \in G_y$. Then $F_y(\sigma)$ is the number of fixed points of the restriction of σ to

$$\{1, 2, \dots, n\} \setminus y = \{1, 2, \dots, y-1, y+1, \dots, n\}.$$

Substituting for F in the right-hand side of Equation (3.4) and applying Burnside's Lemma, we obtain

$$\begin{aligned} \sum_{\sigma \in G} F(\sigma)^2 &= \sum_{y=1}^n \sum_{\sigma \in G_y} (F_y(\sigma) + 1). \\ &= \sum_{y=1}^n \left(o(G_y) + \sum_{\sigma \in G_y} F_y(\sigma) \right) \\ &\geq 2 \sum_{y=1}^n o(G_y), \end{aligned}$$

with equality if and only if G_y is transitive on $\{1, 2, \dots, n\} \setminus y$ for all $y \in \{1, 2, \dots, n\}$. As in the proof of Burnside's Lemma,

$$\begin{aligned} \sum_{y=1}^n o(G_y) &= o(G) \sum_{y=1}^n 1/o(G_y) \\ &= to(G), \end{aligned}$$

where t is the number of orbits of G . Because doubly transitive groups are transitive, it remains to prove (for $n \geq 3$) that G is doubly transitive if and only if G_y is transitive on $\{1, 2, \dots, n\} \setminus y$ for all $y \in \{1, 2, \dots, n\}$, and this is left to the exercises. $\square$

Cycle type of π	[5]	[4,1]	[3,2]	[3,1 ²]	[2 ² ,1]	[2,1 ³]	[1 ⁵]
$n(\pi)$	24	30	20	20	15	10	1

FIGURE 3.1

EXAMPLE 3.33 Of the permutations in S_5 , there are 24 with cycle type [5], 30 with cycle type [4,1], 20 with cycle type [3,2], and so on. The numbers of permutations of each cycle type appear in Figure 3.1. When $G = S_5$, Inequality (3.3) becomes

$$\begin{aligned}
 \frac{1}{120} \sum_{\sigma \in G} F(\sigma)^2 &= \frac{1}{120} (24 \times 0 + 30 \times 1 + 20 \times 0 + 20 \times 4 + 15 \times 1 \\
 &\quad + 10 \times 9 + 1 \times 25) \\
 &= 2,
 \end{aligned}$$

confirming that S_5 is doubly transitive.³

□

Which abstract finite groups have manifestations (isomorphic images) as transitive permutation groups. The answer is easy: All of them. To see why, it suffices to examine the proof of the following classical result:

CAYLEY'S THEOREM 3.34 *Let G be a finite group. Then G is isomorphic to a transitive permutation group.*

Proof We have defined a permutation of degree n to be a one-to-one function from $A = \{1, 2, \dots, n\}$ onto A . It does not matter, of course, whether the elements of A are expressed in the Hindu-Arabic numeration system, Roman Numerals, binary, hexadecimal, or any other numeration system. In fact, the elements of A don't have to be numbers at all. What matters is that A is a set of cardinality n . For most purposes the elements $1, 2, \dots, n$ are as good as any. However, in the present instance it makes things easier to permute, not the first n positive integers, but the elements of $G = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$.

For each $\tau \in G$, define $\bar{\tau}: G \rightarrow G$ by $\bar{\tau}(\sigma_j) = \tau\sigma_j$, $1 \leq j \leq n$. Since $\tau\sigma_j = \tau\sigma_k$ if and only if $j = k$, $\bar{\tau}$ is a permutation of G . Similarly, $\bar{\tau}_1 = \bar{\tau}_2$ if and

³For a classification of finite doubly transitive groups, see [Kantor (1985)].

only if $\tau_1 = \tau_2$, so the correspondence $\tau \rightarrow \bar{\tau}$ is one-to-one. Because

$$\begin{aligned} [\bar{\tau}_1 \circ \bar{\tau}_2](\sigma) &= \bar{\tau}_1(\bar{\tau}_2(\sigma)) \\ &= \bar{\tau}_1(\tau_2\sigma) \\ &= \tau_1(\tau_2\sigma) \\ &= (\tau_1\tau_2)\sigma \\ &= \overline{\tau_1\tau_2}(\sigma), \end{aligned}$$

for all $\sigma \in G$, $\bar{G} = \{\bar{\tau} : \tau \in G\}$ is a group under function composition, and $\tau \rightarrow \bar{\tau}$ is an isomorphism from G onto $\bar{G}$. Finally, for any pair, i and j , there exists a $\tau \in G$, namely $\tau = \sigma_i\sigma_j^{-1}$, such that $\tau(\sigma_j) = \sigma_i$. Hence, $\bar{G}$ is transitive. $\square$

Consider S_3 , a transitive permutation group of degree 3 and order 6. The proof of Theorem 3.34 produces a transitive isomorphic image, $\bar{S}_3$, of degree 6. (While it is true that S_3 may be viewed as the subgroup of S_6 consisting of those permutations that fix 4, 5, and 6, this subgroup is not the same as the transitive subgroup $\bar{S}_3$.) Evidently, S_3 has (at least) two genuinely different manifestations as transitive permutation groups. In fact, given a finite group G , it is surprisingly easy to determine all of its transitive homomorphic images.

Suppose H is a subgroup of G . Let $G = \sigma_1 H \cup \sigma_2 H \cup \cdots \cup \sigma_r H$, $r = [G : H]$, be the left coset decomposition of H in G . The idea is to use the elements of G to permute $A = \{\sigma_i H : 1 \leq i \leq r\}$. For each $\tau \in G$, define $\hat{\tau} : A \rightarrow A$ by $\hat{\tau}(\sigma_i H) = (\tau\sigma_i)H$, $1 \leq i \leq r$. As in the proof of Cayley's Theorem, the mapping $\tau \rightarrow \hat{\tau}$ is a homomorphism from G onto the transitive permutation group $\hat{G} = \{\hat{\tau} : \tau \in G\}$. Indeed, if $H = \{e\}$, then $\hat{G} = \bar{G}$. When $H \neq \{e\}$, however, the mapping need not be one-to-one.

Observe that $\hat{\tau}(\sigma H) = \sigma H$, if and only if $(\tau\sigma)H = \sigma H$, if and only if $\sigma^{-1}\tau\sigma \in H$, if and only if $\tau \in \sigma H\sigma^{-1}$. Therefore, the kernel of the homomorphism $\tau \rightarrow \hat{\tau}$ is

$$\begin{aligned} K &= \{\tau \in G : \hat{\tau}(\sigma H) = \sigma H, \sigma \in G\} \\ &= \bigcap_{\sigma \in G} \sigma H\sigma^{-1}, \end{aligned}$$

the largest normal subgroup of G contained in H . In particular, $\tau \rightarrow \hat{\tau}$ is one-to-one if and only if $\{e\}$ is the only subgroup of H that is normal in G .

DEFINITION 3.35 Let H be a subgroup of the finite group G . The transitive homomorphic manifestation of G arising from its action on the left cosets of H is said to **correspond** to H .

Suppose, conversely, that G is a permutation group acting transitively on $A = \{a_1, a_2, \dots, a_m\}$. Let $a = a_1$ and take $H = G_a$, the stabilizer subgroup of a . By Lemma 3.26, $m = [G : G_a] = [G : H]$. Moreover, because G is transitive, there exist $\sigma_2, \sigma_3, \dots, \sigma_m \in G$ such that $\sigma_i(a) = a_i$, $2 \leq i \leq m$. Thus, $a_i \rightarrow \sigma_i H$ establishes a natural one-to-one correspondence between A and the set $\{H, \sigma_2 H, \dots, \sigma_m H\}$ consisting of the distinct left cosets of H in G . It is easily verified that the action of G on A is carried over by this correspondence to the action of G on the cosets. Thus, the transitive action of G on A corresponds to $H = G_a$. These observations are summarized in the following.

THEOREM 3.36 *Let $\varphi: G \rightarrow \hat{G}$ be a homomorphism from the finite group G onto a permutation group $\hat{G}$ that acts transitively on a set A . Then the action of $\hat{G}$ on A is identical to the action of G on the left cosets of one of its subgroups H . Moreover, φ is an isomorphism if and only if $K = \{e\}$ is the only subgroup of H that is normal in G .*

Applications to Symmetry

Permutation groups arise naturally in discussions of geometric symmetry. Consider, for example, the square in Figure 3.2(a). Imagine that it has been reproduced on an overhead projection transparency. If the transparency square were aligned on top of the original, then only a single square would be visible. If you were to place the point of a needle at the intersection of the diagonals of the square and rotate (just the transparency) 28 degrees in the clockwise direction, you would see not one square, but two. Therefore, a 28° rotation is *not* a symmetry of the square. If the transparency were rotated 90°, the squares would again be superimposed; again only one square would be visible. A 90° clockwise rotation is a symmetry of the square. In order to discuss all the different symmetries we must be a little more precise about what is meant by a “symmetry”, and much more precise about what is meant by “different”.



FIGURE 3.2

Suppose the vertices of the square are numbered as shown in Figure 3.2(b). (Never mind that a 90° rotation is not a symmetry of the labeled figure. The labels are only there to facilitate discussion. While they rotate with the square, they are not part of it.) A 90° clockwise rotation acts as a permutation of the vertices. Vertex 1 is sent to the position formerly occupied by vertex 2, vertex 2 goes to the place previously held by vertex 4, and so on. It is natural to associate the vertex permutation $\sigma = (1243)$ with a 90° clockwise rotation. What about a 90° counterclockwise rotation? That corresponds to $(1342) = \sigma^{-1}$. Notice that a 90° counterclockwise rotation and a 270° clockwise rotation correspond to the same permutation. What matters in discussions of symmetry is where the figure winds up, not the route it took getting there. Two symmetries are the same if and only if they afford the same permutation. A 90° counterclockwise rotation and a 270° clockwise rotation are different geometric routes to the same symmetry.

Each symmetry of the square corresponds to a unique permutation of its vertices. This suggests that we may as well use permutations as convenient descriptive names for symmetries. (Be careful, however. This discussion is taking place in the context of some fixed but arbitrary numbering of the vertices. While the symmetries do not depend on these numbers, their descriptive names do.)

There are just four symmetries that arise from rotating the square around a vertical axis through its center. They are (1243) , (1342) , $(14)(23)$, and e_4 . Four more symmetries arise from rotations about axes that lie in the plane of the square. The set of all eight symmetries is

$$D_4 = \{e_4, (1243), (1342), (14)(23), (14), (23), (12)(34), (13)(24)\}. \quad (3.5)$$

Evidently, only a third of the 24 permutations in S_4 are symmetries. (Pause for a minute and think about the effect of applying the permutation (12) to the square in Figure 3.2(b).) Observe that D_4 is a group. Indeed, if $\sigma, \tau \in D_4$, the combined symmetry, σ followed by τ , is the symmetry $\sigma\tau$; the function mapping a symmetry to its name is a one-to-one group homomorphism.

DEFINITION 3.37 Let G be a subgroup of S_n . Suppose it is possible to label some geometric figure F in such a way that every element of G is a symmetry of F . Then G is a **symmetry group**.

Among the symmetries of the square are those that satisfy the additional constraint that the transparency must remain flat on top of the original; a *plane symmetry* is one that can be accomplished entirely within the two-dimensional plane. The plane symmetries of the square comprise a symmetry group, namely the cyclic group generated by (1243) . Somewhat ironically, the plane symmetries of the square are described by means of rotations about an axis perpendicular to the plane, while the nonplanar symmetries can be construed as rotations about axes in the plane. The nonplanar symmetries can also be visualized as reflections.

Consider another example, the cube. In the “real world”, it is conventional to number not the vertices, but the faces of cubes. The standard way to number

dice is illustrated in Figure 3.3. How many symmetries does a cube have? Let's begin with an analogy. The square is a two-dimensional figure. It lies in the plane. It seemed natural to divide the symmetries of the square into two types, planar and nonplanar. The cube is a three-dimensional figure; its symmetries can be divided naturally between those that can be accomplished entirely within three-dimensional space, and those that cannot. The three-dimensional symmetries are all rotations (of the kind that take place 24 hours a day in Nevada gambling casinos). While the remaining symmetries can be construed as rotations through the fourth dimension, it may be easier to visualize them as reflections.

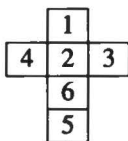


FIGURE 3.3 The numbered faces of a die.

To count the rotations of the cube, observe first that any of the six numbered faces can be rotated to the top. Once the top (and bottom) faces have been determined, any one of the four “equatorial” faces can be rotated to the front. Thus, there are $6 \times 4 = 24$ rotational symmetries of the cube. With respect to the standard numbering of dice, they are listed in Figure 3.4.

(1265)	(12)(34)(56)	(123)(465)	(153)(246)
(1364)	(13)(25)(46)	(124)(365)	(154)(236)
(1463)	(14)(25)(36)	(132)(456)	(16)(25)
(1562)	(15)(26)(34)	(135)(264)	(16)(34)
(2354)	(16)(23)(45)	(142)(356)	(25)(34)
(2453)	(16)(24)(35)	(145)(263)	e_6

FIGURE 3.4 The rotational symmetries of the cube.

Perhaps it is inconsistent to have described the symmetries of a square as permutations of its vertices and the symmetries of a cube as permutations of its faces. Why not look upon the symmetries of a cube as vertex permutations? What difference would it make? The symmetries themselves are independent of whether we *describe* them in terms of faces or vertices, or edges, for that matter. A practical sort of difference is that as permutations of faces, the symmetries of the cube are

elements of S_6 . As permutations of its vertices, they are elements of S_8 , and as permutations of the edges, they form a subgroup of S_{12} . Suppose we number the vertices of the cube as follows:

$$\begin{aligned} 1 &= \{1, 2, 3\}, & 2 &= \{1, 2, 4\}, & 3 &= \{1, 3, 5\}, & 4 &= \{1, 4, 5\}, \\ 5 &= \{2, 3, 6\}, & 6 &= \{2, 4, 6\}, & 7 &= \{3, 5, 6\}, & 8 &= \{4, 5, 6\} \end{aligned} \quad (3.6)$$

Then, for example, “ $6 = \{2, 4, 6\}$ ” means that we are going to assign (vertex number) 6 to the vertex at the intersection of the even numbered faces. Consider the symmetry $\sigma = (1265)$, manifested as a permutation of the faces. Where does σ send vertex 1? The action of σ on the faces *induces* a natural action on the vertices, namely,

$$1 = \{1, 2, 3\} \rightarrow \{\sigma(1), \sigma(2), \sigma(3)\} = \{2, 6, 3\} = 5. \quad (3.7)$$

Let's write $\tilde{\sigma}$ for this induced action of σ , that is,

$$\tilde{\sigma}(\{x, y, z\}) = \{\sigma(x), \sigma(y), \sigma(z)\}.$$

Then, from (3.7), $\tilde{\sigma}(1) = 5$. Okay, what about $\tilde{\sigma}(5)$? Well,

$$\tilde{\sigma}(5) = \tilde{\sigma}(\{2, 3, 6\}) = \{\sigma(2), \sigma(3), \sigma(6)\} = \{6, 3, 5\} = 7.$$

So, $\tilde{\sigma}(5) = 7$. Continuing in this way, we obtain $\tilde{\sigma} = (1573)(2684)$. Figure 3.5 tabulates σ and the corresponding induced permutation, $\tilde{\sigma}$, for all 24 rotational symmetries of the cube. (Observe that $\tilde{\sigma}$ and $\tilde{\tau}$ may have the same cycle structure even when σ and τ do not.)

Whatever its manifestation, the group G exhibited in Figure 3.5 comprises only some of the symmetries of the cube — the 24 rotations. What about reflections? Imagine a die placed on a mirrored table. Suppose face 1 is on top (so face 6 is touching the table), and face 2 is in front. If the reflection could be raised up and superimposed on the die, then faces 1 and 6 would be interchanged. As a permutation of the faces, this reflection is $\tau = (16)$. Given one reflection, it is easy to generate more. If $\sigma \in G$, then $\mu = \sigma\tau$ is a symmetry of the cube. Might it be a rotation? If $\mu \in G$, then $\tau = \sigma^{-1}\mu \in G$, which is a contradiction. Since $\sigma\tau$ cannot be a rotation, it must be another reflection. Indeed, because $\sigma_1\tau = \sigma_2\tau$ if and only if $\sigma_1 = \sigma_2$, the coset $G\tau = \{\sigma\tau : \sigma \in G\}$ contains 24 different reflections. On the other hand, because the die and its reflected image rotate together, $G\tau$ contains all possible reflections. In other words, the group of all 48 symmetries of the cube is $H = G \cup G\tau$.

σ	$\tilde{\sigma}$	σ	$\tilde{\sigma}$
(1265)	(1573)(2684)	(12)(34)(56)	(12)(36)(45)(78)
(1364)	(1562)(3784)	(13)(25)(46)	(13)(27)(45)(68)
(1463)	(1265)(3487)	(14)(25)(36)	(18)(24)(36)(57)
(1562)	(1375)(2486)	(15)(26)(34)	(18)(27)(34)(56)
(2354)	(1342)(5786)	(16)(23)(45)	(15)(27)(36)(48)
(2453)	(1243)(5687)	(16)(24)(35)	(18)(26)(37)(45)
(123)(465)	(253)(467)	(153)(246)	(147)(285)
(124)(365)	(164)(358)	(154)(236)	(176)(238)
(132)(456)	(235)(476)	(16)(25)	(17)(28)(35)(46)
(135)(264)	(174)(258)	(16)(34)	(16)(25)(38)(47)
(142)(356)	(146)(385)	(25)(34)	(14)(23)(58)(67)
(145)(263)	(167)(283)	e_6	e_8

FIGURE 3.5 Rotations of the cube as vertex permutations.

Exercises

- Write down the disjoint cycle factorization of σ^{-1} when
 - $\sigma = (1234)$
 - $\sigma = (12345)$
 - $\sigma = (123)$
 - $\sigma = (12)$
- Find the disjoint cycle factorization of the indicated permutation.
 - $(13)(1234)(13)$
 - $(12)(1234)(12)$
 - $(13)(12345)(13)$
 - $(14)(12345)(14)$
- Prove Lemma 3.2.
- Show that the number of permutations in S_{12} of cycle type
 - $[3^4]$ is 246,400.
 - $[4^3]$ is 1,247,400.
 - $[6^2]$ is 6,652,800.
 - $[2^6]$ is 10,395.
- Show that 1624 of the 5040 permutations in S_7 have disjoint cycle factorizations consisting of exactly three cycles.
- The **Stirling Number of the First Kind**, $s(n, r)$, is the number of permutations in S_n whose disjoint cycle factorizations consist of (exactly) r cycles. From Exercise 5, $s(7, 3) = 1624$.
 - Show that $s(n, n) = 1$. (Hint: Don't forget 1-cycles.)
 - Show that $s(n, 1) = (n-1)!$.
 - If $1 < r \leq n$, prove that $s(n+1, r) = s(n, r-1) + ns(n, r)$.
 - Prove that $s(n, 1)x + s(n, 2)x^2 + \cdots + s(n, n)x^n = x(x+1)(x+2) \cdots (x+n-1)$.
- Prove Theorem 3.10.
- Prove that $o(A_n) = n!/2$, $n > 1$.
- Prove that a k -cycle is even if and only if k is odd.
- Write the disjoint cycle factorizations of all permutations in S_4 and underline those permutations belonging to A_4 .
- Find a permutation $\tau \in S_9$ such that $\tau^{-1}\sigma\tau = \mu$, when
 - $\sigma = (1234)(56789)$ and $\mu = (1324)(58769)$.
 - $\sigma = (1234)(56789)$ and $\mu = (1432)(59876)$.

- c. $\sigma = (12)(3456)(789)$ and $\mu = (123)(45)(6789)$.
12. Prove that every permutation in S_n is conjugate to its inverse.
 13. Let $\sigma = (123)$.
 - a. Show that σ and $\sigma^{-1} = (132)$ are not conjugate in A_3 .
 - b. Find all permutations $\tau \in S_n$ such that $\tau^{-1}(123)\tau = (132)$.
 14. Prove Theorem 3.21.
 15. Let G be a subgroup of S_n . Prove that G is transitive if and only if, for all $y \in \{1, 2, \dots, n\}$, there exists a $\sigma \in G$ such that $\sigma(1) = y$.
 16. Prove that A_n is transitive for all $n \geq 3$.
 17. Let G be a subgroup of S_n and suppose $y \in O_x$. Prove that G_x and G_y are conjugate, that is, prove there exists a permutation $\sigma \in G$ such that $G_y = \sigma^{-1}G_x\sigma$.
 18. Prove that S_n is doubly transitive for all $n \geq 2$.
 19. Let G be a permutation group of degree $n > 2$. Prove that G is doubly transitive if and only if G_y is transitive on $\{1, 2, \dots, n\} \setminus y$ for all $y \in \{1, 2, \dots, n\}$.
 20. Prove that $D_4 = \{e_4, (1243), (14)(23), (1342), (14), (23), (12)(34), (13)(24)\}$ is transitive
 - a. by showing that $O_1 = \{1, 2, 3, 4\}$.
 - b. using Corollary 3.30.
 21. Prove that the group D_4 from Exercise 20 is not doubly transitive
 - a. using Definition 3.31.
 - b. using Theorem 3.32.
 22. The Bell numbers are defined by $B_0 = 1$ and

$$B_{n+1} = \sum_{r=0}^n C(n, r) B_r,$$

where $C(n, r)$ is the binomial coefficient “ n -choose- r ”. Compute B_1 through B_4 . (Hint: $B_5 = 52$.)

23. Let G be a permutation group of degree n . Then G is r -fold transitive if, given any two sequences, $(x_1, x_2, \dots, x_r)$ and $(y_1, y_2, \dots, y_r)$, of distinct integers chosen from $\{1, 2, \dots, n\}$, there exists a permutation $\sigma \in G$ such that $\sigma(x_i) = y_i$, $1 \leq i \leq r$. Prove the following generalization of Theorem 3.32 from [Merris and Pierce (1972)]: If $r \leq n$, then

$$\frac{1}{o(G)} \sum_{\sigma \in G} F(\sigma)^r \geq B_r,$$

the r -th Bell number, with equality if and only if G is r -fold transitive. (Hint: Exercise 22.)

24. Using a direct computation along the lines of Example 3.33, confirm that
 - a. $\frac{1}{6} \sum_{\sigma \in S_3} F(\sigma)^3 = B_3$,
 - b. $\frac{1}{24} \sum_{\sigma \in S_4} F(\sigma)^r = B_r$, $3 \leq r \leq 4$.
 (Hint: Exercises 22 and 23.)
25. Let $G = A_4$, the alternating group of degree 4.
 - a. Find the number of orbits of G using Burnside's Lemma. (Hint: Exercise 10.)
 - b. Prove that G is doubly transitive.
 - c. Is G triply (3-fold) transitive? (Hint: Exercise 23.)
26. A permutation group G of degree n is **semiregular** if $G_x = \{e_n\}$ for all $x \in \{1, 2, \dots, n\}$. Suppose G is semiregular.
 - a. Prove that $o(O_x) = o(G)$, for all $x \in \{1, 2, \dots, n\}$.
 - b. Prove that $n/o(G)$ is an integer, that is, $o(G)$ (exactly) divides n .
 - c. If G is semiregular and transitive, prove that $o(G) = n$.
27. Let $G = S_3$ and $H = A_3$. If $\hat{G}$ is the transitive homomorphic image of G corresponding to H , show that, apart from the objects that it permutes, $\hat{G}$ is identical to S_2 .
28. Let $G = S_3$ and $H = \{e_3, (12)\} \cong S_2$. Let $\hat{G}$ be the transitive homomorphic image of G corresponding to H . Prove or disprove that, apart from the objects that it permutes, $\hat{G}$ is identical to A_3 .
29. Let $G = S_4$ and $H_1 = \{e_4, (12)(34), (13)(24), (14)(23)\}$. Let G' be the transitive homomorphic image of G corresponding to H_1 .
 - a. What is the degree of G' ?
 - b. What is $o(G')$?
 - c. Apart from the objects it permutes, to which familiar permutation group is G' identical?
 - d. Let $H_2 = \{e_4, (12), (34), (12)(34)\}$. Prove that H_2 is isomorphic to H_1 .
 - e. If $\hat{G}$ is the transitive homomorphic image of G corresponding to H_2 , what is the degree of $\hat{G}$?
 - f. What is $o(\hat{G})$?
 - g. Apart from the objects it permutes, to which familiar permutation group is $\hat{G}$ identical?
30. Let H be a subgroup of the finite group G . Let $\hat{G}$ be the transitive homomorphic image of G corresponding to H . If $\hat{G}$ is doubly transitive, prove that

- a. $H = N(H)$, the normalizer of H in G , or $[G : H] = 2$.
- b. H is a maximal subgroup of G .

Application Exercises

31. Denote by D_3 the group of symmetries of the equilateral triangle as permutations of its vertices.
 - a. Show that $D_3 = S_3$.
 - b. Show that the group of plane symmetries of the equilateral triangle is A_3 .
32. Suppose the vertices of a regular pentagon are consecutively numbered 1 through 5 in clockwise order. Use this numbering to express, as a subgroup of S_5 ,
 - a. the group of plane symmetries of the pentagon.
 - b. the group of all symmetries of the pentagon.
33. Denote by D_n the group of all symmetries of the regular n -gon expressed as permutations of its vertices. Show that D_n is generated by symmetries σ and τ that satisfy the relations $\sigma^n = \tau^2 = e_n$ and $\tau^{-1}\sigma\tau = \sigma^{-1}$.
34. A regular tetrahedron is a pyramid with a triangular base in which each of the four triangular faces is equilateral.
 - a. Prove that the regular tetrahedron has 12 rotational symmetries.
 - b. Express the rotational symmetries of the regular tetrahedron as a permutation group of degree 4. (Hint: Number the faces.)
 - c. Express the rotational symmetries of the regular tetrahedron as a permutation group of degree 6. (Hint: Unlike the cube, every pair of faces of the tetrahedron meet to form an edge. Using the face numberings from part b, number the edges lexicographically, that is,

$$\begin{aligned} 1 &= \{1, 2\}, & 2 &= \{1, 3\}, & 3 &= \{1, 4\}, \\ 4 &= \{2, 3\}, & 5 &= \{2, 4\}, & 6 &= \{3, 4\}. \end{aligned}$$

Let G be the group of rotational symmetries as permutations of the four faces. For each $\sigma \in G$, let $\tilde{\sigma}$ be the natural induced action of σ on the edges, that is, $\tilde{\sigma}(\{x, y\}) = \{\sigma(x), \sigma(y)\}$.

35. Prove that the group of 24 rotational symmetries of the cube is transitive but not doubly transitive.
36. Let G be the group of 24 rotational symmetries of the cube. As we are about to see, G is abstractly isomorphic to S_4 . This being the case, it follows from

Exercise 35 and Theorem 3.36 that G corresponds to the action of S_4 on the cosets of one of its subgroups, H .

- Explain why $o(H)$ must be four.
- Nearly everyone's favorite four element subgroup of S_4 is $K = \{e_4, (12)(34), (13)(24), (14)(23)\}$. Explain why G couldn't possibly be identical to the action of S_4 on the cosets of K .
- Show that G is identical to the action of S_4 on $H = \{e_4, (1234), (13)(24), (1432)\}$. (Hint: Number the cosets of H as follows:

$$1 = H;$$

$$2 = \{(12), (1324), (143), (234)\};$$

$$3 = \{(123), (1342), (14), (243)\};$$

$$4 = \{(1243), (134), (142), (23)\};$$

$$5 = \{(124), (132), (1423), (34)\};$$

$$6 = \{(12)(34), (13), (14)(23), (24)\}.$$

- Let G be the group of 24 rotational symmetries of the cube expressed as permutations of its six faces. Then $\tilde{G} = \{\tilde{\sigma} : \sigma \in G\}$ (see Figure 3.5) is an isomorphic image of G as a transitive subgroup of S_8 . Find a subgroup H of G to which this transitive manifestation corresponds.
- Write out the 24 reflections of the cube
 - as permutations of the six faces.
 - as permutations of the eight vertices.
- Let G be the group of all 48 symmetries of the cube expressed as permutations of its six faces. Prove that G is not doubly transitive
 - from geometrical considerations.
 - using Theorem 3.32. (Hint: Exercise 38a.)

CHAPTER 4

Group Representation Theory

Denote by $GL(n, \mathbb{C})$ the multiplicative group consisting of all invertible matrices in $\mathbb{C}_{n,n}$. Among the elements of $GL(n, \mathbb{C})$ are the $n!$ permutation matrices (the $(0,1)$ -matrices with exactly one 1 in each row and column).

For each $\sigma \in S_n$, let

$$A(\sigma) = (\delta_{i,\sigma(j)}), \quad (4.1)$$

the n -by- n matrix whose (i, j) -entry is 1 if $\sigma(j) = i$, and zero otherwise. If $\sigma, \tau \in S_n$, then the (i, j) -entry of $A(\sigma)A(\tau)$ is

$$\sum_{k=1}^n \delta_{i,\sigma(k)} \delta_{k,\tau(j)} = \delta_{i,\sigma\tau(j)},$$

precisely the (i, j) -entry of $A(\sigma\tau)$. Evidently, $A: S_n \rightarrow GL(n, \mathbb{C})$ is a homomorphism. In fact, more is true.

THEOREM 4.1 *The multiplicative group of n -by- n permutation matrices is isomorphic to S_n .*

Proof Let A be the homomorphism defined in Equation (4.1). Then $A(\sigma) = A(\tau)$ if and only if $\sigma(j) = \tau(j)$, $1 \leq j \leq n$, if and only if $\sigma = \tau$. Therefore, A is one-to-one. Because there are exactly $o(S_n) = n!$ permutation matrices in $GL(n, \mathbb{C})$, A must be onto. In other words, A is an isomorphism. $\square$

DEFINITION 4.2 Let G be a group. A **representation** of G of **degree** n is a homomorphism $A: G \rightarrow GL(n, \mathbb{C})$. If the homomorphism is one-to-one, so that G is isomorphic to the image of A , then the representation is **faithful**.

EXAMPLE 4.3 The mapping $\sigma \rightarrow (\delta_{i,\sigma(j)})$ is a faithful representation of S_n of degree n . In this case, the “degree” of the permutations is equal to the “degree” of the representation. On the other hand, the mapping $\sigma \rightarrow A(\sigma)$ given by

$$\begin{aligned} A(e_3) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, & A((12)) &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ A((23)) &= \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix}, & A((123)) &= \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}, \\ A((132)) &= \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}, & \text{and } A((13)) &= \begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix}, \end{aligned}$$

is a faithful representation of S_3 of degree 2. The restriction of A to $H = \{e_3, (13)\}$, namely,

$$A(e_3) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad A((13)) = \begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix},$$

is what we will eventually call a “reducible” representation of H . $\square$

EXAMPLE 4.4 The mapping $\sigma \rightarrow I_n, \sigma \in G$, is a representation of G of degree n . When $n = 1$, it is called the **principal representation** of G . $\square$

EXAMPLE 4.5 Let $U \in GL(n, \mathbb{C})$ be fixed but arbitrary. Then the mappings $A \rightarrow U^{-1}AU$ and $A \rightarrow \det(A)$ are representations of $GL(n, \mathbb{C})$ of degrees n and 1, respectively. $\square$

EXAMPLE 4.6 If A is a representation of G of degree n , and Θ is a representation of $GL(n, \mathbb{C})$ of degree k , then $\sigma \rightarrow \Theta(A(\sigma))$ is a representation of G of degree k . In particular, $\sigma \rightarrow \det(A(\sigma))$ is a representation of G of degree 1. If $A(\sigma) = (\delta_{i,\sigma(j)})$, $\sigma \in S_n$, then $\det(A(\sigma)) = \varepsilon(\sigma)$, the alternating character (Definition 3.14). $\square$

EXAMPLE 4.7 Suppose A and B are representations of G of degrees m and n , respectively. Let C be their direct sum, that is,

$$\begin{aligned} C(\sigma) &= \text{diag } (A(\sigma), B(\sigma)) \\ &= \begin{pmatrix} A(\sigma) & 0 \\ 0 & B(\sigma) \end{pmatrix}. \end{aligned}$$

Then C is a representation of G of degree $m + n$. If $U \in GL(m + n, \mathbb{C})$, then $D(\sigma) = U^{-1}C(\sigma)U, \sigma \in G$, is another representation. The uniform similarity typically obscures the fact that D has been constructed from representations of smaller degree. One might think of it as a “cover-up”, designed to mask the relative simplicity of the underlying structure. One of our goals is to expose such cover-ups. $\square$

The first example of a representation, Equation (4.1), is restricted to permutation groups. On the other hand, by Cayley's Theorem, every finite group is isomorphic to a (transitive) permutation group. Let's have another look at this classical result from the perspective of representations.

CAYLEY'S THEOREM 4.8 *Let G be a finite group. Then G is isomorphic to a group of permutation matrices.*

Proof Suppose $G = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$. For each $\tau \in G$, define $A(\tau) = (a_{ij}(\tau))$ by

$$a_{ij}(\tau) = \begin{cases} 1, & \text{if } \tau\sigma_j = \sigma_i \\ 0, & \text{otherwise.} \end{cases} \quad (4.2)$$

Then $A(\tau)$ is a permutation matrix. The (i, j) -entry of $A(\tau)A(\mu)$, namely,

$$\sum_{k=1}^n a_{ik}(\tau)a_{kj}(\mu),$$

is zero unless there is a k such that $a_{ik}(\tau) = a_{kj}(\mu) = 1$. This is equivalent to the existence of a k such that $\mu\sigma_j = \sigma_k = \tau^{-1}\sigma_i$, because there can be at most one such k ,

$$\sum_{k=1}^n a_{ik}(\tau)a_{kj}(\mu) = \begin{cases} 1 & \text{if } (\tau\mu)\sigma_j = \sigma_i, \\ 0 & \text{otherwise.} \end{cases} \quad (4.3)$$

The right-hand side of Equation (4.3) is the (i, j) -entry of $A(\tau\mu)$. Therefore, $A(\tau\mu) = A(\tau)A(\mu)$, and $A: G \rightarrow GL(n, \mathbb{C})$ is a representation of G . Because the mapping $\tau \rightarrow A(\tau)$ is one-to-one, it is a faithful representation. Therefore, G is isomorphic to a subgroup of the n -by- n permutation matrices. $\square$

DEFINITION 4.9 Let G be a finite group. The representation defined by Equation (4.2) is a **regular representation** of G .

Strictly speaking, the representation defined by Equation (4.2) is a "left" regular representation. If $G = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$ then, mimicking the proof of Theorem 4.8, it can be shown that $\tau \rightarrow B(\tau) = (b_{ij}(\tau))$, $\tau \in G$, where

$$b_{ij}(\tau) = \begin{cases} 1, & \text{if } \sigma_i\tau = \sigma_j \\ 0, & \text{otherwise,} \end{cases}$$

defines a "right" regular representation. Moreover, both $A(\tau)$ and $B(\tau)$ depend, not only on G , but on the ordering of its elements. It is easy to see, however, that if A_1 is a left regular representation corresponding to one ordering of the elements of G and A_2 is a left regular representation corresponding to another, then there exists a fixed permutation matrix P such that

$$A_2(\sigma) = P^{-1}A_1(\sigma)P, \quad \sigma \in G.$$

DEFINITION 4.10 Two representations of G , A_1 and A_2 , are **equivalent** if there exists an invertible matrix U such that

$$A_2(\sigma) = U^{-1}A_1(\sigma)U, \quad \sigma \in G. \quad (4.4)$$

So, any two left regular representations of the same group G are equivalent. What is less obvious, but will follow from subsequent developments, is that every left regular representation of G is equivalent to each of its right regular representations.

The effect of Definition 4.10 is to partition the representations of G into a disjoint union of equivalence classes. Together with Equation (4.4), a system of distinct representatives for these equivalence classes will generate all representations of G . This suggests another perspective from which to view representation theory.

DEFINITION 4.11 Suppose V is an n -dimensional complex vector space. Denote by $GL(V) \subset L(V, V)$ the group of invertible linear operators on V .

If $[T]$ is the matrix representation of $T \in GL(V)$ with respect to a fixed but arbitrary basis of V , then the mapping $T \rightarrow [T]$ is an isomorphism from $GL(V)$ onto $GL(n, \mathbb{C})$. We might just as well define a representation of the abstract group G to be a homomorphism from G into $GL(V)$. From the usual change of basis formula, Definition 4.10 asserts that two matrix representations of G are equivalent if and only if the underlying transformation representations are identical. While concentrating primarily on matrix representations, we will feel free to move back and forth between these two perspectives.

DEFINITION 4.12 Let N be a nonempty set and suppose $S = \{A(\nu) : \nu \in N\}$ is a set of n -by- n matrices indexed by N . Then S is **reducible** if there exists a matrix $U \in GL(n, \mathbb{C})$ and an integer p such that $1 < p < n$ and, for all $\nu \in N$,

$$U^{-1}A(\nu)U = \begin{pmatrix} B(\nu) & 0 \\ C(\nu) & D(\nu) \end{pmatrix}, \quad (4.5)$$

where $B(\nu) \in \mathbb{C}_{p,p}$. The set S is **fully reducible** if U can be chosen so that $C(\nu) = 0$, for all $\nu \in N$. Finally, S is **irreducible** if it is not reducible. The representation A of G is reducible, fully reducible, or irreducible if the set of matrices $\{A(\sigma) : \sigma \in G\}$ has the corresponding property.

Observe that the invertible matrix U in Equation (4.5) does not depend on ν . If a set of matrices is reducible, then the matrices in the set are uniformly similar to (the same) block triangular form.

Consider the matrix

$$A = \begin{pmatrix} 7 & -9 \\ 4 & -5 \end{pmatrix}.$$

Then A is similar to a lower triangular matrix, namely,

$$U^{-1}AU = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad \text{where} \quad U = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}.$$

If A were similar to a diagonal matrix D , then D could only be I_2 . Because I_2 is similar only to itself, A is not diagonalizable. This simple example makes the following result seem even more remarkable.

MASCHKE'S THEOREM 4.13 *Suppose A is a representation of degree n of the finite group G . If A is reducible, then it is fully reducible.*

Proof Suppose $1 < p < n$ and $U \in GL(n, \mathbb{C})$ satisfy

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

where $B(\sigma) \in GL(p, \mathbb{C})$, $\sigma \in G$. Because A is a representation of G , both B and D are representations of G as well. Moreover,

$$C(\sigma\tau) = C(\sigma)B(\tau) + D(\sigma)C(\tau). \quad (4.6)$$

Because $C(e) = 0$, Equation (4.6) implies that

$$0 = C(\sigma)B(\sigma^{-1}) + D(\sigma)C(\sigma^{-1}), \quad \sigma \in G. \quad (4.7)$$

Define an $(n - p)$ -by- p matrix X by

$$X = \frac{1}{o(G)} \sum_{\tau \in G} C(\tau)B(\tau^{-1}).$$

(This is why we need G to be finite.) If

$$T = \begin{pmatrix} I_p & 0 \\ X & I_{n-p} \end{pmatrix},$$

then

$$T^{-1} = \begin{pmatrix} I_p & 0 \\ -X & I_{n-p} \end{pmatrix},$$

and

$$T^{-1}U^{-1}A(\sigma)UT = \begin{pmatrix} B(\sigma) & 0 \\ F(\sigma) & D(\sigma) \end{pmatrix},$$

where $F(\sigma) = C(\sigma) + D(\sigma)X - XB(\sigma)$. The proof will be finished if we can show that

$$XB(\sigma) = C(\sigma) + D(\sigma)X, \quad \sigma \in G. \quad (4.8)$$

For a fixed but arbitrary $\sigma \in G$,

$$\begin{aligned} XB(\sigma) &= \frac{1}{o(G)} \sum_{\tau \in G} C(\tau) B(\tau^{-1}\sigma) \\ &= \frac{1}{o(G)} \sum_{\tau \in G} (C(\sigma) - D(\tau)C(\tau^{-1}\sigma)), \end{aligned}$$

by Equation (4.6). Therefore,

$$\begin{aligned} XB(\sigma) &= C(\sigma) - \frac{1}{o(G)} \sum_{\mu \in G} D(\sigma\mu)C(\mu^{-1}) \\ &= C(\sigma) - D(\sigma) \left(\frac{1}{o(G)} \sum_{\mu \in G} D(\mu)C(\mu^{-1}) \right) \\ &= C(\sigma) + D(\sigma)X, \end{aligned}$$

by Equation (4.7) and the definition of X . This verifies Equation (4.8), completing the proof. $\square$

DEFINITION 4.14 Let A be a representation of G . A **reduction** of A is a similarity

$$\begin{aligned} U^{-1}A(\sigma)U &= \text{diag}(A_1(\sigma), A_2(\sigma), \dots, A_r(\sigma)), \\ &= \begin{pmatrix} A_1(\sigma) & 0 & \dots & 0 \\ 0 & A_2(\sigma) & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & A_r(\sigma) \end{pmatrix}, \end{aligned}$$

$\sigma \in G$, where $A_1, A_2, \dots, A_r$ are irreducible representations of G . The positive integer r is the **length** of the reduction.

THEOREM 4.15 If A is a representation of degree n of the finite group G , then A has a reduction.

Proof The result follows from Maschke's Theorem by induction on n . $\square$

If A is a representation of the finite group G , then, from Theorem 4.15, there exists an invertible matrix U such that

$$A(\sigma) = U[\text{diag}(A_1(\sigma), A_2(\sigma), \dots, A_r(\sigma))]U^{-1}, \quad \sigma \in G,$$

where $A_1, A_2, \dots, A_r$ are irreducible representations of G . Thus, to generate all representations of G it suffices to find a system of distinct representatives for the equivalence classes of irreducible representations. We will refer to such a system as a **complete set of inequivalent irreducible representations** of G .

Suppose $\{A(\nu): \nu \in N\}$ is a set of n -by- n matrices. If B is an ordered basis of an n -dimensional vector space V then, for all $\nu \in N$, there exists a linear operator $T(\nu) \in L(V, V)$ whose matrix representation with respect to B is $A(\nu)$. If $\{A(\nu): \nu \in N\}$ is reducible, then there exists an n -by- n invertible matrix U such that

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

where $B(\nu) \in \mathbb{C}_{p,p}$, $\nu \in N$. Now, $U^{-1}A(\nu)U$ is the matrix representation of $T(\nu)$ with respect to another ordered basis $B' = \{w_1, w_2, \dots, w_n\}$. Moreover, from the block triangular form of $U^{-1}A(\nu)U$, it is evident that $W = \langle w_1, w_2, \dots, w_p \rangle$ is an **invariant subspace** of $T(\nu)$ for all ν , that is, $T(\nu)(w) \in W$ for all $w \in W$ and all $\nu \in N$. Conversely, if W is a proper subspace of V that is invariant under $T(\nu)$, $\nu \in N$, then $\{A(\nu): \nu \in N\}$ is reducible. (Extend any basis of W to a basis B' of V . From this perspective, a reduction of $\{A(\sigma): \sigma \in G\}$ corresponds to a decomposition of V into a direct sum,

$$V = W_1 \oplus W_2 \oplus \dots \oplus W_r,$$

where $W_1, W_2, \dots, W_r$ are minimal invariant subspaces of $T(\sigma)$, $\sigma \in G$.) Finally, $\{A(\nu): \nu \in N\}$ is irreducible if and only if no proper subspace of V is invariant under $T(\nu)$ for every $\nu \in N$. In this case, $\{T(\nu): \nu \in N\}$ is said to be an **irreducible subset** of $L(V, V)$.

SCHUR'S LEMMA 4.16 *Let N be a set. Let $\{S(\nu): \nu \in N\}$ and $\{T(\nu): \nu \in N\}$ be irreducible subsets of $L(V, V)$ and $L(W, W)$, respectively. Let $L \in L(V, W)$ be fixed but arbitrary. If $LS(\nu) = T(\nu)L$ for all $\nu \in N$, then $L = 0$, or $\dim(V) = \dim(W)$ and L is invertible.*

Proof It follows from $LS(\nu) = T(\nu)L$ that the image of L is an invariant subspace of $T(\nu)$, and its kernel is an invariant subspace of $S(\nu)$. $\square$

This deep Lemma has many important and useful implications.

COROLLARY 4.17 *Let $\{A(v) : v \in N\}$ be an irreducible set of n -by- n matrices. If M is a matrix such that $MA(v) = A(v)M$, for all $v \in N$, then M is a multiple of I_n .*

Proof Let λ be an eigenvalue of M . Then $M - \lambda I_n$ is a singular matrix that commutes with each $A(v)$. It follows from Schur's Lemma that $M - \lambda I_n = 0$. $\square$

COROLLARY 4.18 *Suppose that $\{A(\sigma) : \sigma \in G\}$ and $\{B(\sigma) : \sigma \in G\}$ are inequivalent irreducible representations of (the same group) G . If $M \in \mathbb{C}_{n,n}$ satisfies $A(\sigma)M = MB(\sigma)$ for all $\sigma \in G$, then $M = 0$.*

The next application of Schur's Lemma establishes a relation that must hold among the elements of irreducible representations.

THEOREM 4.19 (Schur Relations). *Let $A(\sigma) = (a_{ij}(\sigma))$, $\sigma \in G$, be an irreducible representation of degree n of the finite group G . Then*

$$\sum_{\sigma \in G} a_{is}(\sigma^{-1})a_{tj}(\sigma) = \delta_{i,j}\delta_{s,t}o(G)/n. \quad (4.9)$$

If $B(\sigma) = (b_{ij}(\sigma))$, $\sigma \in G$, is another irreducible representation of G , then

$$\sum_{\sigma \in G} a_{is}(\sigma^{-1})b_{tj}(\sigma) = 0, \quad (4.10)$$

unless B is equivalent to A .

Proof Suppose m is the degree of B . Define a function $f : \mathbb{C}_{n,m} \rightarrow \mathbb{C}_{n,m}$ by

$$f(S) = \sum_{\sigma \in G} A(\sigma^{-1})SB(\sigma).$$

Then $A(\tau)f(S) = f(S)B(\tau)$, for all $\tau \in G$. We will complete the proof by making special choices for S . Denote by E_{st} the n -by- m matrix whose only nonzero entry is a 1 in position (s, t) . The (i, j) -entry of $f(E_{st})$ is

$$\sum_{\sigma \in G} a_{is}(\sigma^{-1})b_{tj}(\sigma) = 0$$

(by Corollary 4.18) when A and B are inequivalent. If $A = B$ then (Corollary 4.17), $f(E_{st}) = c_{st}I_n$. Hence,

$$\begin{aligned} c_{st}\delta_{i,j} &= \sum_{\sigma \in G} a_{is}(\sigma^{-1})a_{tj}(\sigma) \\ &= \sum_{\sigma \in G} a_{is}(\sigma)a_{tj}(\sigma^{-1}) \\ &= c_{ji}\delta_{t,s}. \end{aligned}$$

Thus, $c_{ij} = 0$ unless $i = j$, and $c_{ii} = c$ is independent of i . Therefore,

$$\begin{aligned}
 nc &= \sum_{i=1}^n c_{ii} \\
 &= \sum_{\sigma \in G} \sum_{i=1}^n a_{ii}(\sigma^{-1})a_{ii}(\sigma) \\
 &= \sum_{\sigma \in G} \sum_{i=1}^n a_{ii}(\sigma)a_{ii}(\sigma^{-1}) \\
 &= \sum_{\sigma \in G} a_{ii}(\sigma) \\
 &= o(G),
 \end{aligned}$$

because $A(e) = I_n$. □

EXAMPLE 4.20 Let A be the representation of S_3 of degree 2 given in Example 4.3. As will be shown in Example 4.33, A is irreducible. Observe that

$$\begin{aligned}
 \sum_{\sigma \in S_3} a_{12}(\sigma^{-1})a_{21}(\sigma) &= a_{12}(e_3)a_{21}(e_3) + a_{12}((12))a_{21}((12)) \\
 &\quad + a_{12}((23))a_{21}((23)) + a_{12}((132))a_{21}((123)) \\
 &\quad + a_{12}((123))a_{21}((132)) + a_{12}((13))a_{21}((13)) \\
 &= 0 \times 0 + 1 \times 1 + (-1) \times 0 + (-1) \times (-1) \\
 &\quad + 1 \times 1 + 0 \times (-1) \\
 &= 6/2,
 \end{aligned}$$

and

$$\begin{aligned}
 \sum_{\sigma \in S_3} a_{11}(\sigma^{-1})a_{22}(\sigma) &= 1 \times 1 + 0 \times 0 + (-1) \times 1 + (-1) \times (-1) \\
 &\quad + 0 \times 0 + 1 \times (-1) \\
 &= 0,
 \end{aligned}$$

confirming Equation (4.9). □

The Schur Relations can be generalized as follows:

THEOREM 4.21 *Let $A(\sigma) = (a_{ij}(\sigma))$ and $B = (b_{ij}(\sigma))$ be irreducible representations of the finite group G . Then*

$$\sum_{\sigma \in G} a_{is}(\sigma^{-1}) b_{tj}(\sigma \tau) = \begin{cases} \delta_{s,t} a_{ij}(\tau) o(G)/n, & \text{if } A = B; \\ 0, & \text{if } A \text{ is not equivalent to } B. \end{cases} \quad (4.11)$$

where n is the degree of A .

BURNSIDE'S THEOREM 4.22 *Let $A(\sigma) = (a_{ij}(\sigma))$, $\sigma \in G$, be an irreducible representation of degree n . Then the n^2 functions $a_{ij}: G \rightarrow \mathbb{C}$ are linearly independent.*

Proof¹ Suppose $c_{ij} \in \mathbb{C}$ are constants such that

$$\sum_{i,j=1}^n c_{ij} a_{ij}(\sigma) = 0, \quad \sigma \in G. \quad (4.12)$$

Multiply both sides of Equation (4.12) by $a_{pq}(\sigma^{-1})$ and sum on $\sigma \in G$. By Equation (4.9), the result is $o(G)c_{qp}/n = 0$. $\square$

If $C = (c_{ij}) \in \mathbb{C}_{n,n}$, then

$$\sum_{i,j=1}^n \bar{c}_{ij} a_{ij}(\sigma) = \text{tr}(C^* A(\sigma)).$$

In view of Equation (4.12), this means $\{a_{ij}: 1 \leq i, j \leq n\}$ is linearly independent if and only if, with respect to the inner product $(A, B) = \text{tr}(B^* A)$, the orthogonal complement of $\{A(\sigma): \sigma \in G\}$ is $\{0\}$, if and only if $\{A(\sigma): \sigma \in G\}$ spans $\mathbb{C}_{n,n}$. In particular, if n is the degree of an irreducible representation of G , then $n^2 = \dim(\mathbb{C}_{n,n}) \leq o(G)$.

Frobenius and Schur obtained the following generalization of Burnside's Theorem.

THEOREM 4.23 *Let $A_k(\sigma) = (a_{ij}^k(\sigma))$, $\sigma \in G$, be an irreducible representation of G of degree n_k , $1 \leq k \leq r$. If these r representations are pairwise inequivalent, then the $n_1^2 + n_2^2 + \cdots + n_r^2$ functions $a_{ij}^k: G \rightarrow \mathbb{C}$ are linearly independent.*

¹ Our proof of Burnside's Theorem (not to be confused with Burnside's Lemma) is valid only for finite groups. Proofs of Theorems 4.22 and 4.23 valid for infinite groups can be found, e.g., in [Newman (1968)].

DEFINITION 4.24 Suppose A is a representation of the group G . Let $\chi(\sigma) = \text{tr}(A(\sigma))$, $\sigma \in G$. The function $\chi: G \rightarrow \mathbb{C}$ is the **character** of G afforded by A . An **irreducible character** is a character afforded by an irreducible representation. Denote by $I(G)$ the set of irreducible characters of G .

If the degree of A is 1 then $A(\sigma) = (\chi(\sigma))$, $\sigma \in G$, and χ is called a **linear character**. In particular, every linear character is irreducible. Among the linear characters is $\chi = 1_G$, the character afforded by the principal representation of G .

Because the trace is preserved under similarity, equivalent representations afford the same character. Thus, it makes sense to talk about the character afforded by a representation $T: G \rightarrow GL(V)$.

THEOREM 4.25 (*Orthogonality Relations of the First Kind*). Let G be a finite group. If $\chi, \xi \in I(G)$, then

$$\sum_{\sigma \in G} \chi(\sigma^{-1})\xi(\sigma) = \begin{cases} o(G), & \text{if } \chi = \xi; \\ 0, & \text{otherwise.} \end{cases} \quad (4.13)$$

Proof Let $A(\sigma) = (a_{ij}(\sigma))$ and $B(\sigma) = (b_{ij}(\sigma))$ be representations that afford χ and ξ , respectively. If $n = \chi(e)$ and $m = \xi(e)$ are the degrees of A and B respectively, then

$$\sum_{\sigma \in G} \chi(\sigma^{-1})\xi(\sigma) = \sum_{i=1}^n \sum_{j=1}^m \sum_{\sigma \in G} a_{ii}(\sigma^{-1})b_{jj}(\sigma),$$

which, by Equation (4.10), is zero if $\chi \neq \xi$. If $\chi = \xi$, we may assume $A = B$, in which case the left-hand side of Equation (4.13) becomes

$$\begin{aligned} \sum_{i,j=1}^n \sum_{\sigma \in G} a_{ii}(\sigma^{-1})a_{jj}(\sigma) &= \sum_{i=1}^n o(G)/n \\ &= o(G), \end{aligned}$$

by the Schur Relations (Equation (4.9)). □

Using Theorem 4.21, the following extension of Equation (4.13) can be proved in a similar way.

THEOREM 4.26 Let G be a finite group. If $\chi, \xi \in I(G)$, then

$$\sum_{\sigma \in G} \chi(\sigma^{-1})\xi(\sigma\tau) = \begin{cases} o(G)\chi(\tau)/\chi(e), & \text{if } \chi = \xi; \\ 0, & \text{otherwise.} \end{cases} \quad (4.14)$$

THEOREM 4.27 *Let χ be a character of the finite group G . Then $\chi(\sigma^{-1}) = \overline{\chi(\sigma)}$, $\sigma \in G$.*

Proof Let A be a representation of G that affords χ . Suppose $o(G) = k$. If $\sigma \in G$ is fixed but arbitrary, then $I_n = A(e) = A(\sigma^k) = A(\sigma)^k$, where $n = \chi(e)$ is the degree of A . Hence, the eigenvalues of $A(\sigma)$ are all k -th roots of unity. Therefore, λ is an eigenvalue of $A(\sigma)$, if and only if $\lambda^{-1} = \overline{\lambda}$ is an eigenvalue of $A(\sigma)^{-1} = A(\sigma^{-1})$. $\square$

DEFINITION 4.28 Suppose G is a group. Denote by $\mathbb{C}G$ the set of all functions $f: G \rightarrow \mathbb{C}$.

Observe that $\mathbb{C}G$ is a vector space under the usual definitions of addition and scalar multiplication of complex valued functions, namely,

$$(af + bg)(\sigma) = af(\sigma) + bg(\sigma),$$

for all $a, b \in \mathbb{C}$, and all $f, g: G \rightarrow \mathbb{C}$.

THEOREM 4.29 *If G is a finite group, then*

$$(f, g)_G = \frac{1}{o(G)} \sum_{\sigma \in G} \overline{g(\sigma)} f(\sigma), \quad f, g \in \mathbb{C}G, \quad (4.15)$$

defines an inner product on $\mathbb{C}G$.

When speaking of the inner product space $\mathbb{C}G$, we will always have Equation (4.15) in mind.

COROLLARY 4.30 *Let G be a finite group. Then $I(G)$, the set of irreducible characters of G , is an orthonormal set in the inner product space $\mathbb{C}G$.*

Proof The result is immediate from Theorems 4.25, 4.27, and 4.29. $\square$

It follows from Corollary 4.30 that $o(I(G)) \leq \dim(\mathbb{C}G) = o(G)$.

Because $\text{tr}(AB) = \text{tr}(BA)$ for all $A \in \mathbb{C}_{m,n}$ and $B \in \mathbb{C}_{n,m}$, similar matrices have the same trace. On the other hand,

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

and I_2 have the same trace without being similar.

THEOREM 4.31 *Two representations of the finite group G are equivalent if and only if they afford the same character.*

Proof Let $\sigma \rightarrow A(\sigma) = (a_{ij}(\sigma))$ and $\sigma \rightarrow B(\sigma) = (b_{ij}(\sigma))$ be representations of G . If A and B are equivalent then $A(\sigma)$ and $B(\sigma)$ have the same trace, $\sigma \in G$.

To prove the converse, assume first that A and B are irreducible. If $\text{tr}(A(\sigma)) = \chi(\sigma) = \text{tr}(B(\sigma))$, $\sigma \in G$, then, from Corollary 4.30,

$$\begin{aligned} 1 &= (\chi, \chi)_G \\ &= \frac{1}{o(G)} \sum_{\sigma \in G} \chi(\sigma^{-1}) \chi(\sigma) \\ &= \frac{1}{o(G)} \sum_{i,j=1}^n \sum_{\sigma \in G} a_{ii}(\sigma^{-1}) b_{jj}(\sigma), \end{aligned}$$

where $n = \chi(e)$ is the degree of both A and B . If A were not equivalent to B then, by the Schur Relations (Equation (4.10)) this last expression would be 0.

In the general case, consider reductions of A and B . Let A_i , $1 \leq i \leq t$, be a set of pairwise inequivalent irreducible representations such that every irreducible constituent occurring in either reduction is equivalent to one of them. Let χ_i be the character afforded by A_i . By what we have proved so far, χ_i , $1 \leq i \leq t$, are t different irreducible characters of G . Suppose that A_i is equivalent to r_i constituents in the reduction of A and to s_i constituents in the reduction of B . Then

$$\sum_{i=1}^t r_i \chi_i(\sigma) = \chi(\sigma) = \sum_{i=1}^t s_i \chi_i(\sigma),$$

for all $\sigma \in G$, implying that

$$\sum_{i=1}^t (r_i - s_i) \chi_i = 0, \quad (4.16)$$

the zero function. Because $\{\chi_i : 1 \leq i \leq t\}$ is an orthonormal set in the inner product space $\mathbb{C}G$, it is linearly independent. Therefore, $r_i = s_i$, $1 \leq i \leq t$. Thus, there is a one-to-one correspondence between the irreducible constituents of a reduction of A and the irreducible constituents of a reduction of B in which corresponding constituents are equivalent. Hence, A and B are equivalent. $\square$

Suppose χ is a fixed but arbitrary character of G . Let A be a representation of G affording χ . By Theorem 4.15, A has a reduction. Thus, there exist irreducible characters χ_i and nonnegative integers m_i , $1 \leq i \leq k$, such that $\chi = \sum m_i \chi_i$. Because $I(G)$ is orthonormal,

$$\begin{aligned} m_i &= (\chi, \chi_i)_G \\ &= \frac{1}{o(G)} \sum_{\sigma \in G} \chi(\sigma) \chi_i(\sigma^{-1}). \end{aligned} \quad (4.17)$$

Therefore,

$$(\chi, \chi)_G = \sum_{i=1}^k m_i^2. \quad (4.18)$$

In particular, χ is irreducible if and only if $(\chi, \chi)_G = 1$.

EXAMPLE 4.32 Let A be a permutation representation of degree n of the finite group G . That is, $\{A(\sigma) : \sigma \in G\}$ is a set of permutation matrices. If χ is the character afforded by A , then $\chi(\sigma) = F(\sigma)$, the number of fixed points of the "permutation" $A(\sigma)$. Therefore (Corollary 3.30), $(\chi, 1)_G \geq 1$, with equality if and only if the permutation group $A(G) = \{A(\sigma) : \sigma \in G\}$ is transitive. If $n > 1$, it follows that A is reducible and the principal representation is a constituent of its reduction. Therefore $\xi(\sigma) = \chi(\sigma) - 1$, $\sigma \in G$, defines a character of G . Moreover, from Equation (4.18), ξ is irreducible if and only if $(\chi, \chi)_G = 2$. As we have seen (Theorem 3.32)

$$\begin{aligned} (\chi, \chi)_G &= \frac{1}{o(G)} \sum_{\sigma \in G} F(\sigma)^2 \\ &\geq 2, \end{aligned}$$

with equality if and only if $\{A(\sigma) : \sigma \in G\}$ is doubly transitive. Thus, ξ is irreducible if and only if $A(G)$ is doubly transitive. $\square$

EXAMPLE 4.33 Let ξ be the character afforded by the representation in Example 4.3. Then $\xi(\sigma) = F(\sigma) - 1$, $\sigma \in S_3$, where $F(\sigma)$ is the fixed point character. Because S_3 is doubly transitive, it follows from Example 4.32 that ξ , and hence the representation A , is irreducible. $\square$

EXAMPLE 4.34 Every finite group G has a permutation representation in the sense of Example 4.32: If $G = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$ then, from Equation (4.2), the (i, j) -entry of the corresponding (left) regular representation, $A(\tau)$, is

$$a_{ij}(\tau) = \begin{cases} 1 & \text{if } \tau\sigma_j = \sigma_i, \\ 0 & \text{otherwise.} \end{cases}$$

If ζ is the character afforded by A , then

$$\begin{aligned} \zeta(\tau) &= \sum_{i=1}^n a_{ii}(\tau) \\ &= o(\{i : \tau\sigma_i = \sigma_i\}) \\ &= 0, \end{aligned}$$

unless $\tau = e$, in which case it is $\zeta(e) = n = o(G)$. Therefore,

$$\zeta(\tau) = \begin{cases} o(G) & \text{if } \tau = e, \\ 0 & \text{otherwise.} \end{cases} \quad (4.19)$$

(In particular, any two regular representations of G are equivalent.) Because A has a reduction, we can write

$$\zeta = \sum_{\chi \in I(G)} m_{\chi} \chi,$$

where

$$\begin{aligned} m_{\chi} &= (\zeta, \chi)_G \\ &= \frac{1}{o(G)} \sum_{\sigma \in G} \chi(\sigma^{-1}) \zeta(\sigma) \\ &= \chi(e), \end{aligned} \quad (4.20)$$

by Equation (4.19). In other words, the multiplicity of χ as a constituent of ζ is equal the degree of χ . $\square$

THEOREM 4.35 *If G is a finite group, then*

$$\sum_{\chi \in I(G)} \chi(e)^2 = o(G). \quad (4.21)$$

Proof Let ζ be the character of the regular representation described in Example 4.34. From Equation (4.19), $(\zeta, \zeta)_G = o(G)$. From Equation (4.20), $(\zeta, \zeta)_G = \sum \chi(e)^2$. $\square$

Unless every irreducible character of G is linear, Theorem 4.35 strictly improves the bound $o(I(G)) \leq o(G)$ that emerged from Corollary 4.30. We are about to obtain another improvement.

THEOREM 4.36 *Let G be an arbitrary group. If χ is a character of G , then $\chi(\sigma) = \chi(\tau^{-1}\sigma\tau)$, for all $\sigma, \tau \in G$. That is, χ is constant on the conjugacy classes of G .*

DEFINITION 4.37 Let G be a group. Denote the conjugacy class of G to which σ belongs by $C_G(\sigma) = \{\tau^{-1}\sigma\tau : \tau \in G\}$. The number of different conjugacy classes of G is its **class number**. A function $f \in \mathbb{C}G$ that is constant on conjugacy classes is called a **class function**. Denote by $Z(\mathbb{C}G)$ the subspace of $\mathbb{C}G$ consisting of the class functions.

Suppose $C = C_G(\sigma)$ is a conjugacy class of G . Define $f_C: G \rightarrow \mathbb{C}$ by $f_C(\mu) = 1$ if $\mu \in C$, and 0 otherwise. Then $f_C \in Z(\mathbb{C}G)$ is the **characteristic function** of C . Because the characteristic functions of the conjugacy classes comprise a basis for $Z(\mathbb{C}G)$, its dimension is equal to the class number of G . By Theorem 4.36 and the Orthogonality Relations of the First Kind, the irreducible characters of G comprise an orthonormal set in the inner product space $Z(\mathbb{C}G)$. Therefore, the number of different irreducible characters of G is at most its class number. In fact, these two numbers are equal.

THEOREM 4.38 *Let G be a finite group. Then $I(G)$ is an orthonormal basis of $Z(\mathbb{C}G)$.*

Proof Let $\{A_k: 1 \leq k \leq r\}$ be a complete set of inequivalent irreducible representations of G . Suppose $A_k(\sigma) = (a_{ij}^k(\sigma))$, $\sigma \in G$, is a representation of degree n_k . By Theorems 4.23 and 4.35, the $n_1^2 + n_2^2 + \cdots + n_r^2 = o(G)$ linearly independent functions a_{ij}^k comprise a basis of $\mathbb{C}G$. If $f: G \rightarrow \mathbb{C}$ is a fixed but arbitrary function, there must exist square matrices $B_k = (b_{ij}^k)$, of size n_k , $1 \leq k \leq r$, such that

$$f(\sigma) = \sum_{k=1}^r \sum_{i,j=1}^{n_k} b_{ji}^k a_{ij}^k(\sigma), \quad (4.22)$$

for all $\sigma \in G$. Thus, for any $\sigma, \tau \in G$,

$$\begin{aligned} f(\tau^{-1}\sigma\tau) &= \sum_{k=1}^r \sum_{i,j=1}^{n_k} b_{ji}^k a_{ij}^k(\tau^{-1}\sigma\tau) \\ &= \sum_{k=1}^r \sum_{i,j=1}^{n_k} b_{ji}^k \sum_{s,t=1}^{n_k} a_{is}^k(\tau^{-1}) a_{st}^k(\sigma) a_{ij}^k(\tau) \\ &= \sum_{k=1}^r \sum_{s,t=1}^{n_k} \left(\sum_{i,j=1}^{n_k} a_{ij}^k(\tau) b_{ji}^k a_{is}^k(\tau^{-1}) \right) a_{st}^k(\sigma). \end{aligned} \quad (4.23)$$

If we assume that f is a class function, it follows from Equations (4.22)–(4.23) and the linear independence of the a_{ij}^k that

$$b_{is}^k = \sum_{i,j=1}^{n_k} a_{ij}^k(\tau) b_{ji}^k a_{is}^k(\tau^{-1}),$$

for all s, t, k , and τ . In other words, $B_k = A_k(\tau) B_k A_k(\tau^{-1})$, or $B_k A_k(\tau) = A_k(\tau) B_k$, for all k and τ . Therefore (Corollary 4.17), $B_k = c_k I_{n_k}$. In other words,

$b_{ji}^k = \delta_{i,j} c_k$. Substituting these values into Equation (4.22) yields

$$\begin{aligned} f &= \sum_{k=1}^r \sum_{i=1}^{n_k} c_k a_{ii}^k \\ &= \sum_{k=1}^r c_k \chi_k, \end{aligned}$$

where χ_k is the character afforded by A_k . We have proved that the orthonormal set $I(G)$ spans the inner product space $Z(CG)$. $\square$

COROLLARY 4.39 *Let G be a finite group. Then G is abelian if and only if each of its irreducible representations is linear.*

Proof Because $\sigma\tau = \tau\sigma$ if and only if $\tau^{-1}\sigma\tau = \sigma$, G is abelian if and only if it has class number $o(G)$, if and only if $o(I(G)) = o(G)$. Because $\chi(e)$ is a positive integer for all $\chi \in I(G)$, the result follows from Equation (4.21). $\square$

THEOREM 4.40 (*Orthogonality Relations of the Second Kind*). *Let σ and τ be elements of the finite group G . Then*

$$\sum_{\chi \in I(G)} \chi(\sigma^{-1})\chi(\tau) = \begin{cases} o(G)/o(C_G(\sigma)) & \text{if } \tau \in C_G(\sigma), \\ 0 & \text{otherwise.} \end{cases} \quad (4.24)$$

Proof Let $f_C \in Z(CG)$ be the characteristic function of the conjugacy class $C = C_G(\sigma)$. Because $I(G)$ is an orthonormal basis for $Z(CG)$,

$$\begin{aligned} f_C &= \sum_{\chi \in I(G)} (f, \chi)_G \chi \\ &= \frac{1}{o(G)} \sum_{\chi \in I(G)} o(C) \overline{\chi(C)} \chi, \end{aligned}$$

where $\chi(C)$ is the common value of $\chi(\mu)$, $\mu \in C$. If $D = C_G(\tau)$, then

$$\begin{aligned} \delta_{C,D} &= f_C(D) \\ &= \frac{o(C)}{o(G)} \sum_{\chi \in I(G)} \overline{\chi(C)} \chi(D) \\ &= \frac{o(C)}{o(G)} \sum_{\chi \in I(G)} \chi(\sigma^{-1})\chi(\tau). \end{aligned}$$

$\square$

EXAMPLE 4.41 Let $G = A_3 = \{e_3, (123), (132)\}$, the cyclic group generated by (123) . It follows from Corollary 4.39 that the irreducible characters of G are all linear. (There is very little to distinguish an irreducible representation of degree 1 from its character.)

The three conjugacy classes of A_3 are $C_1 = \{e_3\}$, $C_2 = \{(123)\}$, and $C_3 = \{(132)\}$. Let $I(A_3) = \{\chi_1, \chi_2, \chi_3\}$. Because $1 = \chi_i(e) = \chi_i((123)^3) = \chi_i((123))^3$, $\chi_i((123))$ is a root of $x^3 - 1$, $1 \leq i \leq 3$. The three cube-roots of unity are 1, $\omega = (-1 + i\sqrt{3})/2$, and $\omega^2 = \bar{\omega} = (-1 - i\sqrt{3})/2$. It is, therefore, a simple matter to write down the character table in Figure 4.1. The rows of this table are pairwise orthogonal "vectors" of magnitude $o(A_3)$, confirming the Orthogonality Relations of the First Kind. The columns are pairwise orthogonal and the j -th column has length $o(G)/o(C_j) = 3/1$, confirming the Orthogonality Relations of the Second Kind. $\square$

	C_1	C_2	C_3
χ_1	1	1	1
χ_2	1	ω	$\bar{\omega}$
χ_3	1	$\bar{\omega}$	ω

FIGURE 4.1 The character table for A_3 , where $\omega = (-1 + i\sqrt{3})/2$.

EXAMPLE 4.42 Let $G = D_4 = \{e_4, (1243), (1342), (14)(23), (14), (23), (12)(34), (13)(24)\}$. After some computations, one discovers that G has five conjugacy classes, namely, $C_1 = \{e_4\}$, $C_2 = \{(14)(23)\}$, $C_3 = \{(1243), (1342)\}$, $C_4 = \{(14), (23)\}$, and $C_5 = \{(12)(34), (13)(24)\}$. So, there are five irreducible characters, one of which is the principal character, χ_1 . Suppose the other four irreducible characters have degrees $n_2 \leq n_3 \leq n_4 \leq n_5$. Then (Theorem 4.35)

$$1^2 + n_2^2 + n_3^2 + n_4^2 + n_5^2 = 8.$$

The only possibility is $n_2 = n_3 = n_4 = 1$ and $n_5 = 2$, that is, χ_2, χ_3 , and χ_4 are linear characters, while $\chi_5(e) = 2$. Because the restriction to D_4 of the alternating character $\varepsilon: S_4 \rightarrow \mathbb{C}$ is different from the principal character, we may let $\chi_2 = \varepsilon$. If ξ is either χ_3 or χ_4 then, because it is a homomorphism, $\xi(C_2)^2 = \xi(C_4)^2 = \xi(C_5)^2 = 1 = \xi(C_3)^4$. Because σ is conjugate to σ^{-1} , $\sigma \in D_4$, ξ is real. Therefore, $\xi(C_i) = \pm 1$, $2 \leq i \leq 5$. Together with $\chi_1 = 1$, $\chi_2 = \varepsilon$, and the Orthogonality Relations of the First Kind, this is enough

information to fill in the first four rows of the character table in Figure 4.2 (with the understanding that rows three and four could just as well be interchanged.)

	C_1	C_2	C_3	C_4	C_5
χ_1	1	1	1	1	1
χ_2	1	1	-1	-1	1
χ_3	1	1	-1	1	-1
χ_4	1	1	1	-1	-1
χ_5	2	a	b	c	d

FIGURE 4.2

From the Orthogonality Relations of the Second Kind, column C_1 is orthogonal to column C_j , $j \geq 2$. Thus, $4 + 2a = 0 = b = c = d$, which yields the character table in Figure 4.3. $\square$

	C_1	C_2	C_3	C_4	C_5
χ_1	1	1	1	1	1
χ_2	1	1	-1	-1	1
χ_3	1	1	-1	1	-1
χ_4	1	1	1	-1	-1
χ_5	2	-2	0	0	0

FIGURE 4.3 The character table for D_4 .

EXAMPLE 4.43 Recall (Theorem 3.18) that two permutations are conjugate in S_n if and only if they have the same cycle structure. (Because σ and σ^{-1} have the same cycle structure, $\chi(\sigma) = \chi(\sigma^{-1}) = \chi(\sigma)$, $\sigma \in S_n$, that is, the characters of S_n are real valued. In fact, as we will see momentarily, they are all integer valued.) Denote by C_π the conjugacy class of S_n consisting of all permutations of cycle type $\pi \vdash n$.² A formula for $n(\pi) = o(C_\pi)$ can be found in Theorem 3.7. When

²So, $C_\pi = C_{\pi^{-1}}$ for any/every permutation $\sigma \in S_n$ of cycle type π .

$n = 5$, these numbers are exhibited in Figure 4.4. Because it has seven conjugacy classes, S_5 has 7 irreducible characters.

C_π	$C_{[1^5]}$	$C_{[2,1^3]}$	$C_{[2^2,1]}$	$C_{[3,1^2]}$	$C_{[3,2]}$	$C_{[4,1]}$	$C_{[5]}$
$n(\pi)$	1	10	15	20	20	30	24

FIGURE 4.4 Cardinalities of the conjugacy classes of S_5 .

Among the $7! = 5040$ one-to-one correspondences between partitions of 5 and conjugacy classes of S_5 , $\pi \leftrightarrow C_\pi$ is the only one that might be called *natural*. While it is not so evident, there is also a natural correspondence between partitions of n and irreducible characters of S_n . Anticipating this result, we will write $I(S_5) = \{\chi_\pi : \pi \vdash 5\}$. In this (as yet mysterious) natural correspondence, $\chi_{[5]}$ is the principal character and $\chi_{[1^5]} = \varepsilon$, the alternating character.

Because the equation

$$\sum_{i=1}^7 n_i^2 = 120$$

has many solutions, for example,

$$1 + 1 + 1 + 9 + 36 + 36 + 36 = 120,$$

$$1 + 1 + 4 + 4 + 4 + 25 + 81 = 120,$$

$$1 + 1 + 1 + 4 + 4 + 9 + 100 = 120, \quad \text{and}$$

$$1 + 1 + 16 + 16 + 25 + 25 + 36 = 120,$$

Theorem 4.35 is not so helpful this time. With all the theoretical machinery we have developed so far, it is still no small task just to determine the degrees of the irreducible characters of S_5 , much less construct the character table exhibited in Figure 4.5. However, given Figures 4.4 and 4.5, it is not difficult to confirm that

$$\begin{aligned} (\chi, \chi)_G &= \frac{1}{120} \sum_{\sigma \in S_5} \chi(\sigma)^2 \\ &= \frac{1}{120} \sum_{\pi \vdash 5} n(\pi) \chi(C_\pi)^2. \\ &= 1, \end{aligned}$$

and

$$\sum_{\chi \in I(S_n)} \chi(C_\pi)^2 = 120/n(\pi).$$

□

	$C_{[1^5]}$	$C_{[2,1^3]}$	$C_{[2^2,1]}$	$C_{[3,1^2]}$	$C_{[3,2]}$	$C_{[4,1]}$	$C_{[5]}$
$\chi_{[1^5]}$	1	-1	1	1	-1	-1	1
$\chi_{[2,1^3]}$	4	-2	0	1	1	0	-1
$\chi_{[2^2,1]}$	5	-1	1	-1	-1	1	0
$\chi_{[3,1^2]}$	6	0	-2	0	0	0	1
$\chi_{[3,2]}$	5	1	1	-1	1	-1	0
$\chi_{[4,1]}$	4	2	0	1	-1	0	-1
$\chi_{[5]}$	1	1	1	1	1	1	1

FIGURE 4.5 The character table for S_5 .

While *ad hoc* methods based on the machinery we have developed so far were sufficient to construct the character tables in Examples 4.41 and 4.42, it is clear from Example 4.43 that additional tools are needed before we will be in a position to deal with the symmetric groups. One such tool emerges from the following generalization of a previous notion: Suppose H is a subgroup of the finite group G . Let

$$G = \sigma_1 H \cup \sigma_2 H \cup \cdots \cup \sigma_r H, \quad r = [G : H],$$

be the left coset decomposition of H in G . Recall that G has a homomorphic image, $\hat{G}$, that acts as a transitive permutation group on $\{\sigma_i H : 1 \leq i \leq r\}$. Let's explore the manifestation of $\hat{G}$ as a representation of G . For each $\tau \in G$, define the r -by- r permutation matrix $\hat{A}(\tau) = (\hat{a}_{ij}(\tau))$ by

$$\hat{a}_{ij}(\tau) = \begin{cases} 1, & \text{if } \tau \sigma_j H = \sigma_i H \\ 0, & \text{otherwise.} \end{cases} \quad (4.25)$$

(Compare with Equation (4.2).) Another way to say the same thing is this: Let $1^\#$ be the characteristic function of H in G . That is, $1^\# : G \rightarrow \{0, 1\}$ is defined by $1^\#(\sigma) = 1$, if $\sigma \in H$, and 0 if $\sigma \in G \setminus H$. Then $\hat{a}_{ij}(\tau) = 1^\#(\sigma_i^{-1} \tau \sigma_j)$, that is,

$$\hat{A}(\tau) = \begin{pmatrix} 1^\#(\sigma_1^{-1} \tau \sigma_1) & 1^\#(\sigma_1^{-1} \tau \sigma_2) & \cdots & 1^\#(\sigma_1^{-1} \tau \sigma_r) \\ 1^\#(\sigma_2^{-1} \tau \sigma_1) & 1^\#(\sigma_2^{-1} \tau \sigma_2) & \cdots & 1^\#(\sigma_2^{-1} \tau \sigma_r) \\ \vdots & \vdots & \ddots & \vdots \\ 1^\#(\sigma_r^{-1} \tau \sigma_1) & 1^\#(\sigma_r^{-1} \tau \sigma_2) & \cdots & 1^\#(\sigma_r^{-1} \tau \sigma_r) \end{pmatrix}. \quad (4.26)$$

Just as in the proof of Theorem 4.8, $\hat{A}$ is a permutation representation of G . In fact, the same idea can be used to convert *any* representation of H into a representation of G . If A is a representation of H of degree n , let $A^\#$ be the matrix valued function of G defined by

$$A^\#(\sigma) = \begin{cases} A(\sigma) & \text{if } \sigma \in H \\ 0_n & \text{if } \sigma \in G \setminus H, \end{cases}$$

where 0_n is the n -by- n matrix of zeros. If $\tau \in G$, let $\hat{A}(\tau)$ be the nr -by- nr partitioned matrix whose (i, j) -block is $A^\#(\sigma_i^{-1}\tau\sigma_j)$, $1 \leq i, j \leq r$. (We are still taking $\sigma_1, \sigma_2, \dots, \sigma_r$ to be a fixed but arbitrary system of representatives for the distinct left cosets of H in G .) In other words,

$$\hat{A}(\tau) = \begin{pmatrix} A^\#(\sigma_1^{-1}\tau\sigma_1) & A^\#(\sigma_1^{-1}\tau\sigma_2) & \dots & A^\#(\sigma_1^{-1}\tau\sigma_r) \\ A^\#(\sigma_2^{-1}\tau\sigma_1) & A^\#(\sigma_2^{-1}\tau\sigma_2) & \dots & A^\#(\sigma_2^{-1}\tau\sigma_r) \\ \vdots & \vdots & \ddots & \vdots \\ A^\#(\sigma_r^{-1}\tau\sigma_1) & A^\#(\sigma_r^{-1}\tau\sigma_2) & \dots & A^\#(\sigma_r^{-1}\tau\sigma_r) \end{pmatrix}, \quad (4.27)$$

$\tau \in G$, is the partitioned matrix whose (i, j) -block is $A(\sigma_i^{-1}\tau\sigma_j)$, if $\tau\sigma_j H = \sigma_i H$, and 0_n otherwise. In particular, each row and column of $\hat{A}(\tau)$ contains exactly one nonzero block. (Observe that Equation (4.26) is the special case of Equation (4.27) corresponding to the principal representation of H .)

Using block matrix multiplication, the (i, j) -block of $\hat{A}(\tau)\hat{A}(\mu)$ is

$$\sum_{k=1}^r A^\#(\sigma_i^{-1}\tau\sigma_k)A^\#(\sigma_k^{-1}\mu\sigma_j) = A^\#(\sigma_i^{-1}\tau\mu\sigma_j).$$

In other words, $\hat{A}(\tau)\hat{A}(\mu) = \hat{A}(\tau\mu)$. Because $\hat{A}(e) = I_{nr}$, this proves that $\hat{A}$ is a representation of G .

DEFINITION 4.44 Suppose H is a subgroup of the finite group G . Let $\sigma_1, \sigma_2, \dots, \sigma_r$, $r = [G : H]$, be fixed but arbitrary representatives for the distinct left cosets of H . Suppose A is a representation of H that affords the character χ . The representation $\hat{A}$ of G defined in Equation (4.27) is said to be **induced** by A . The induced character, χ^G , is the character afforded by $\hat{A}$.

The value of the induced character is given by

$$\begin{aligned} \chi^G(\tau) &= \sum_{i=1}^r \text{tr}(A^\#(\sigma_i^{-1}\tau\sigma_i)) \\ &= \sum_{i=1}^r \chi^\#(\sigma_i^{-1}\tau\sigma_i) \\ &= \frac{1}{o(H)} \sum_{\sigma \in G} \chi^\#(\sigma^{-1}\tau\sigma), \end{aligned} \quad (4.28)$$

where, as expected,

$$\chi^\#(\sigma^{-1}\tau\sigma) = \begin{cases} \chi(\sigma^{-1}\tau\sigma), & \text{if } \sigma^{-1}\tau\sigma \in H \\ 0, & \text{otherwise.} \end{cases} \quad (4.29)$$

While an induced representation depends on the coset representatives, the induced character does not. (This is the easy proof that representations of G induced from the same representation A of H , but corresponding to different coset representatives, are equivalent.) Observe that the degree of χ^G is

$$\chi^G(e) = [G : H]\chi(e), \quad (4.30)$$

where, as usual, $[G : H] = o(G)/o(H)$ is the index of H in G .

EXAMPLE 4.45 Let $G = D_4$. From Example 4.42, the conjugacy classes of G are $C_1 = \{e_4\}$, $C_2 = \{(14)(23)\}$, $C_3 = \{(1243), (1342)\}$, $C_4 = \{(14), (23)\}$, and $C_5 = \{(12)(34), (13)(24)\}$. Let $H = \langle (1243) \rangle$ be the cyclic subgroup generated by (1243) . The coset decomposition of G is $G = H \cup (14)H$, that is, we may take $\sigma_1 = e_4$ and $\sigma_2 = (14)$. Let ξ be the character (homomorphism) of H defined by $\xi(e_4) = 1 = \xi((14)(23))$ and $\xi((1243)) = \xi((1342)) = -1$. Then (Equation (4.30)) $\xi^G(C_1) = \xi^G(e_4) = [G : H]\xi(e_4) = 2$. Because $o(C_2) = 1$, $(14)(23) \in Z(G)$, the center of G . From Equation (4.28), $\xi^G(C_2) = 2$,

$$\begin{aligned} \xi^G(C_3) &= \xi^G((1243)) \\ &= \xi^\#(\sigma_1^{-1}(1243)\sigma_1) + \xi^\#(\sigma_2^{-1}(1243)\sigma_2) \\ &= \xi^\#((1243)) + \xi^\#((1342)) \\ &= \xi((1243)) + \xi((1342)) \\ &= -2, \end{aligned}$$

$$\begin{aligned} \xi^G(C_4) &= \xi^G((14)) \\ &= \xi^\#(e_4(14)e_4) + \xi^\#((14)(14)(14)) \\ &= \xi^\#((14)) + \xi^\#((14)) \\ &= 0 + 0 \\ &= 0, \end{aligned}$$

and

$$\begin{aligned} \xi^G(C_5) &= \xi^G((12)(34)) \\ &= \xi^\#((12)(34)) + \xi^\#((13)(24)) \\ &= 0. \end{aligned}$$

Now that we know the values of ξ^G it follows from Equation (4.17) that $\xi^G = \chi_2 + \chi_3$, where χ_2 and χ_3 are among the irreducible characters of D_4 exhibited in Figure 4.3. $\square$

FROBENIUS RECIPROCITY THEOREM 4.46 *Let H be a subgroup of G . Let ξ be a character of H and χ a character of G . Then*

$$(\xi, \chi)_H = (\xi^G, \chi)_G. \quad (4.31)$$

Proof

$$\begin{aligned} (\xi^G, \chi)_G &= \frac{1}{o(G)} \sum_{\tau \in G} \chi(\tau^{-1}) \xi^G(\tau) \\ &= \frac{1}{o(G)o(H)} \sum_{\tau \in G} \sum_{\sigma \in G} \chi(\tau^{-1}) \xi^\#(\sigma^{-1}\tau\sigma) \\ &= \frac{1}{o(G)o(H)} \sum_{\sigma \in G} \sum_{\tau \in G} \chi(\sigma^{-1}\tau^{-1}\sigma) \xi^\#(\sigma^{-1}\tau\sigma) \\ &= \frac{1}{o(G)} \sum_{\sigma \in G} \frac{1}{o(H)} \sum_{\mu \in G} \chi(\mu^{-1}) \xi^\#(\mu) \\ &= \frac{1}{o(G)} \sum_{\sigma \in G} \frac{1}{o(H)} \sum_{\mu \in H} \chi(\mu^{-1}) \xi(\mu) \\ &= \frac{1}{o(G)} \sum_{\sigma \in G} (\xi, \chi)_H \\ &= (\xi, \chi)_H. \end{aligned}$$

$\square$

Returning to the irreducible characters of S_n , suppose $\pi = [\pi_1, \pi_2, \dots, \pi_m]$ is a partition of n of length m . Of the many sub-groups of S_n isomorphic to S_{π_i} consider the one consisting of the permutations that fix every integer *not* contained in the set

$$\{t : \pi_0 + \pi_1 + \dots + \pi_{i-1} < t \leq \pi_0 + \pi_1 + \dots + \pi_i\},$$

where $\pi_0 = 0$. The **Young Subgroup**³ corresponding to $\pi \vdash n$ is the internal direct product,

$$S_\pi = S_{\pi_1} \times S_{\pi_2} \times \dots \times S_{\pi_m}.$$

³Named for Alfred Young (1873–1940).

Denote by

$$1_\pi = 1_{S_\pi}$$

the principal (identically 1) character of S_π and by

$$\varepsilon_\pi = \varepsilon_{S_\pi}$$

the restriction of the alternating character ε to S_π . Then, adopting the abbreviation

$$(\chi, \xi)_{S_\pi} = (\chi, \xi)_\pi, \quad (4.32)$$

it can be shown that $(1_\pi^{S_\pi}, \varepsilon_\rho^{S_\pi})_\pi$ is the number of double cosets,⁴ $S_\pi \sigma S_\rho$, that satisfy $S_\pi \cap \sigma S_\rho \sigma^{-1} = \{e_n\}$. This double coset number also occurs in what would appear to be a totally different context; it is the number of (0,1)-matrices with row sum vector π and column sum vector ρ .⁵

It is easy to see that there is exactly one (0,1)-matrix with row sum vector π and column sum vector π^* . If $\pi = [5, 3, 1, 1]$, for example, the unique matrix is

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix},$$

where the 1's are crowded into the upper left-hand corner in the shape of $F(\pi)$. It follows that

$$(1_\pi^{S_\pi}, \varepsilon_{\pi^*}^{S_\pi})_\pi = 1, \quad \pi \vdash n. \quad (4.33)$$

In other words, there is exactly one irreducible character of S_n that occurs as a constituent of both $1_\pi^{S_\pi}$ and $\varepsilon_{\pi^*}^{S_\pi}$. Let's call it χ_π .⁶

Two special cases are easily identified. Because $S_{[n]} = S_n$, $1_{[n]}^{S_n}$ is the principal character; its only irreducible constituent is itself. Therefore,

$$\chi_{[n]} = 1_{S_n}.$$

If $\pi = [1^n]$, then $1_\pi^{S_n} = \zeta$, the character of the regular representation of S_n . Because $\pi^* = [n]$, $S_{\pi^*} = S_n$ and $\varepsilon_{\pi^*} = \varepsilon$. (Since each irreducible character of S_n occurs in ζ with multiplicity equal to its degree, we know, without reference to Equation (4.33), that $(\zeta, \varepsilon)_n = 1$.) Thus,

$$\chi_{[1^n]} = \varepsilon.$$

⁴See e.g. [James and Kerber (1981), p. 18]. A discussion of double cosets can be found in Exercises 42–44.

⁵By the Gale-Ryser Theorem, this number is nonzero if and only if the conjugate partition, ρ^* , majorizes π . (See Exercise 31, Chapter 1. The theorem is placed in a more general context by Brualdi and Ryser (1991)), [Doubilet, Fox & Rota (1980)], and [James and Kerber (1981)].)

⁶As we discover more about χ_π , it will become clear that $\pi \leftrightarrow \chi_\pi$ fulfills the promise of Example 4.43 regarding a natural correspondence between partitions of n and irreducible characters of S_n . In particular, $I(S_n) = \{\chi_\pi : \pi \vdash n\}$.

THEOREM 4.47 *If $\pi \vdash n$, then $\chi_{\pi^*} = \varepsilon \chi_\pi$, that is, $\chi_{\pi^*}(\tau) = \varepsilon(\tau) \chi_\pi(\tau)$, for all $\tau \in S_n$.*

Proof It is not difficult to show⁷ that $\varepsilon_{\pi^*}^S = \varepsilon(1_\pi^S)$. Therefore, $\varepsilon(\varepsilon_{\pi^*}^S) = 1_\pi^S$. Because, χ_π is an irreducible constituent of both 1_π^S and $\varepsilon_{\pi^*}^S$, $\varepsilon \chi_\pi$ must be an irreducible constituent of $\varepsilon(\varepsilon_{\pi^*}^S) = 1_\pi^S$ and $\varepsilon(1_\pi^S) = \varepsilon_{\pi^*}^S$. By Equation (4.33) (since $\pi^{**} = \pi$), the unique such irreducible constituent is χ_{π^*} . $\square$

COROLLARY 4.48 *The partition $\pi \vdash n$ is self-conjugate if and only if $\chi_\pi(\sigma) = 0$ for every odd permutation $\sigma \in S_n$.*

Proof If π is self-conjugate then, by Theorem 4.47, $\chi_{\pi^*}(\sigma) = \varepsilon(\sigma) \chi_\pi(\sigma)$, $\sigma \in S_n$. If $\chi_\pi(\sigma) = 0$ for every odd $\sigma \in S_n$ then, by another application of Theorem 4.47, $\chi_{\pi^*} = \chi_\pi$. If $\pi^* \neq \pi$, we would eventually find that S_n contained fewer irreducible characters than conjugacy classes. $\square$

Because 1_ρ^S is a character of S_n , there must exist integers, $K_{\pi, \rho}$, such that

$$1_\rho^S = \sum_{\pi \vdash n} K_{\pi, \rho} \chi_\pi, \quad \rho \vdash n. \quad (4.34)$$

DEFINITION 4.49 The numbers $K_{\pi, \rho} = (1_\rho^S, \chi_\pi)_n$ in Equation (4.34) are called **Kostka coefficients**.⁸

By the Frobenius Reciprocity Theorem, $K_{\pi, \rho} = (1_\rho, \chi_\pi)_{S_\rho}$.

EXAMPLE 4.50 If $\rho = [2, 1^2]$, it is not difficult to verify that

$$1_\rho^{S_4} = \chi_{[4]} + 2\chi_{[3,1]} + \chi_{[2^2]} + \chi_{[2,1^2]}.$$

Thus, for example, $K_{[3,1], [2,1^2]} = 2$. $\square$

The next result, a mechanical procedure for computing Kostka coefficients, is commonly known as **Young's Rule**. In order to state the procedure, we need to use a variation of Ferrers diagrams in which the symbols comprising the rows need not be identical boxes.

THEOREM 4.51 *Let $\rho = [\rho_1, \rho_2, \dots, \rho_r] \vdash n$. Beginning with a Ferrers diagram consisting of a single row of length ρ_1 , construct all possible Ferrers diagrams that can be obtained by adjoining ρ_2 additional symbols (of a second kind), subject to the condition that no two of these new symbols are permitted to lie in the same column. From each of the compound diagrams so constructed, form all possible Ferrers diagrams that can be obtained by adjoining ρ_3 additional symbols (of a*

⁷ See Exercise 34.

⁸ See [Kostka (1882)].

third kind) subject to the condition that no two of the new symbols are placed in the same column. Continue in this way until ρ_r symbols (of an r -th kind) have been added in all possible ways to each of the previously constructed Ferrers diagrams, subject to the same condition. Then $K_{\pi, \rho}$ is the number of times $F(\pi)$ occurs among the resulting diagrams.⁹

EXAMPLE 4.52 Suppose $\rho = [3, 3, 1]$. Beginning with a single row of $\rho_1 = 3$ symbols, **XXX**, adjoin $\rho_2 = 3$ new symbols, in all possible ways, subject to the condition that no two of the new symbols may lie in the same column of any of the resulting Ferrers diagrams. The four possibilities are

XXXYYY **XXXYY** **XXXYY** and **XXX**
 Y **YY** **YYY**

Notice, for example, that

XXXYY
Y
Y

is not permitted. Next, construct all possible Ferrers diagrams that can be obtained from the four (permissible) compound diagrams by adding an additional symbol corresponding to $\rho_3 = 1$. From **XXXYYY**, we obtain

XXXYYYZ and **XXXYYY**
 Z

corresponding to $\chi_{[7]}$ and $\chi_{[6,1]}$. Adding **Z** in all possible ways to

XXXYY
Y

yields the three compound diagrams

XXXYYZ **XXXYY** **XXXYY**
Y **YZ** and **Y**
 Z

corresponding to $\chi_{[6,1]}$, $\chi_{[5,2]}$, and $\chi_{[5,1^2]}$. Adding **Z** in all possible ways to

XXXYY
YY

yields

XXXYZ **XXXYY** **XXXYY**
YY **YYZ** and **YY**
 Z

⁹A Proof can be found, e.g., in [James & Kerber (1981), §2.8] or [Sagan (1991), §2.11].

corresponding to $\chi_{[5,2]}$, $\chi_{[4,3]}$, and $\chi_{[4,2,1]}$. Finally, adding Z to

$$\begin{array}{c} \text{XXX} \\ \text{YYY} \end{array}$$

in all possible ways produces

$$\begin{array}{c} \text{XXXZ} \\ \text{YYY} \end{array} \quad \text{and} \quad \begin{array}{c} \text{XXX} \\ \text{YYY} \\ Z \end{array}$$

corresponding to $\chi_{[4,3]}$ and $\chi_{[3^2,1]}$. (Notice that

$$\begin{array}{c} \text{XXX} \\ \text{YYYZ} \end{array}$$

is not permitted as it is not a Ferrers diagram.) Summarizing, we have that

$$\begin{aligned} 1_{[3^2,1]} &= \chi_{[7]} + 2\chi_{[6,1]} + 2\chi_{[5,2]} + \chi_{[5,1^2]} + 2\chi_{[4,3]} \\ &\quad + \chi_{[4,2,1]} + \chi_{[3^2,1]}. \end{aligned}$$

So, $K_{[7],[3^2,1]} = 1$, $K_{[6,1],[3^2,1]} = 2$, $K_{[5,2],[3^2,1]} = 2$, and so on. □

	[4]	[3,1]	[2 ²]	[2,1 ²]	[1 ⁴]
[4]	1	1	1	1	1
[3,1]	0	1	1	2	3
[2 ²]	0	0	1	1	2
[2,1 ²]	0	0	0	1	3
[1 ⁴]	0	0	0	0	1

FIGURE 4.6 Table of Kostka Coefficients.

EXAMPLE 4.53 The table of Kostka coefficients, $K_{\pi,\rho}$, $\pi, \rho \vdash 4$, appears in Figure 4.6. Observe that the fourth column, corresponding to $\rho = [2, 1^2]$, confirms Example 4.50. Using the notation $(\cdot, \cdot)_n$ for the inner product in $\mathbb{C}S_n$, the first row of the table corresponds to

$$\begin{aligned} K_{[4],\rho} &= (1_{\rho}^{S_4}, \chi_{[4]})_4 \\ &= (1_{\rho}^{S_4}, 1)_4 \\ &= (1, 1)_{S_4} \\ &= 1, \end{aligned}$$

by the Frobenius Reciprocity Theorem. Similarly, the last column consists of the Kostka coefficients

$$\begin{aligned} K_{\pi, [1^n]} &= (1_{[1^n]}^{S_n}, \chi_\pi)_n \\ &= (1, \chi_\pi)_{\{e\}} \\ &= \chi_\pi(e), \end{aligned}$$

the degree of χ_π . Observe that the entries in row π are increasing. This is a consequence of the fact that $K_{\pi, \rho} \geq K_{\pi, \nu}$ whenever ν majorizes ρ .¹⁰ $\square$

COROLLARY 4.54 *The Kostka coefficient $K_{\pi, \pi} = 1$, for all $\pi \vdash n$.*

THEOREM 4.55¹¹ *Let $\rho = [\rho_1, \rho_2, \dots, \rho_r]$ and $\pi = [\pi_1, \pi_2, \dots, \pi_s]$ be partitions of n . Then χ_π is an irreducible constituent of $1_\rho^{S_n}$, that is, the Kostka coefficient $K_{\pi, \rho} \neq 0$, if and only if π majorizes ρ . In particular, if the partitions indexing its rows and columns are arranged in reverse lexicographic order, the Kostka matrix $K = (K_{\pi, \rho})$ is upper triangular.*

The proof depends on a technical lemma concerning the relation of *partial dominance*: Let $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_s)$ and $\beta = (\beta_1, \beta_2, \dots, \beta_r)$ be two nonincreasing sequences of positive integers. Then $\alpha \triangleright \beta$ if $s \leq r$,

$$\sum_{i=1}^k \alpha_i \geq \sum_{i=1}^k \beta_i, \quad 1 \leq k \leq s,$$

and $\alpha_1 + \alpha_2 + \dots + \alpha_s \geq \beta_1 + \beta_2 + \dots + \beta_r$. Thus, partial dominance is a weak form of majorization in the sense that $\alpha > \beta$ if and only if $\alpha \triangleright \beta$ and $\alpha_1 + \alpha_2 + \dots + \alpha_s = \beta_1 + \beta_2 + \dots + \beta_r$.

LEMMA 4.56 *Let $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_s)$ and $\beta = (\beta_1, \beta_2, \dots, \beta_r)$, $r > 1$, be two nonincreasing sequences of positive integers. Suppose $\alpha \triangleright \beta$. Then some β_r of the columns of the Ferrers diagram $F(\alpha)$ can be shortened by one box each in such a way as to produce a new Ferrers diagram, $F(\alpha')$, such that $\alpha' \triangleright \beta' = (\beta_1, \beta_2, \dots, \beta_{r-1})$.*

Because β_r boxes are being removed from $F(\alpha)$ to produce $F(\alpha')$ and from $F(\beta)$ to produce $F(\beta')$, if α majorizes β in Lemma 4.56, then α' majorizes β' .

¹⁰ See [Liebler & Vitale (1973)], [White (1980a)], or [James & Kerber (1981), pp. 44].

¹¹ Conjectured in [Snapper (1971)], the sufficiency part of Theorem 4.55 was first proved in [Lam (1977)] and [Merris (1977)].

Proof The proof is by induction on s . If $s = 1$, simply remove the last β_r boxes from row 1 of $F(\alpha)$. Suppose $s > 1$. If $\alpha_s \geq \beta_r$, then we may take $\alpha' = (\alpha_1, \alpha_2, \dots, \alpha_{s-1}, \alpha_s - \beta_r)$. If $\alpha_s < \beta_r$, the situation is more complicated and we need an intermediate step.

Temporarily remove the first α_s columns from both $F(\alpha)$ and $F(\beta)$. This produces $\alpha^{(1)} = (\alpha_1 - \alpha_s, \alpha_2 - \alpha_s, \dots, \alpha_t - \alpha_s)$, where t is maximal so that $\alpha_t > \alpha_s$, and $\beta^{(1)} = (\beta_1 - \alpha_s, \beta_2 - \alpha_s, \dots, \beta_r - \alpha_s)$. Because $s \leq r$, no more boxes have been deleted from $F(\alpha)$ than from $F(\beta)$. In particular, $\alpha^{(1)} \triangleright \beta^{(1)}$. Therefore, by the induction hypothesis, we can remove $\beta_r - \alpha_s$ boxes from the column ends of $F(\alpha^{(1)})$, producing $F(\alpha^{(2)})$, where

$$\alpha^{(2)} = (\gamma_1, \gamma_2, \dots, \gamma_q) \triangleright \beta^{(2)} = (\beta_1 - \alpha_s, \beta_2 - \alpha_s, \dots, \beta_{r-1} - \alpha_s).$$

We now restore the detached columns. To the left-hand side of $F(\alpha^{(2)})$, attach a rectangular array of s rows and α_s columns obtaining $F(\alpha^{(3)})$. When a rectangular array of r rows and α_s columns is attached to the left-hand side of $F(\beta^{(2)})$, we obtain $\beta^{(3)} = (\beta_1, \beta_2, \dots, \beta_{r-1}, \alpha_s)$. More importantly,

$$\alpha^{(3)} = (\gamma_1 + \alpha_s, \gamma_2 + \alpha_s, \dots, \gamma_q + \alpha_s, \alpha_s, \dots, \alpha_s) \triangleright (\beta_1, \beta_2, \dots, \beta_{r-1}, \alpha_s).$$

Finally, removing α_s boxes from the last row of $F(\alpha^{(3)})$ and $F(\beta^{(3)})$, we obtain $F(\alpha')$ and $F(\beta')$, respectively, where

$$\alpha' \triangleright \beta' = (\beta_1, \beta_2, \dots, \beta_{r-1}).$$

□

Proof (of Theorem 4.55) Suppose χ_π occurs among the irreducible constituents of $1_{\rho}^{S_n}$. Because the construction in Theorem 4.51 begins with a row of ρ_1 symbols, every partition that emerges in the end satisfies $\pi_1 \geq \rho_1$. None of the next ρ_2 symbols can be put in row 3 without putting two of them in column 1, violating the condition that no two of them may occur in the same column. This requires that all of these ρ_2 symbols be placed in the first two rows. (Of course, rows 1 and 2 may grow longer when subsequent symbols are added.) Therefore, $\pi_1 + \pi_2 \geq \rho_1 + \rho_2$. When the next ρ_3 symbols are added, the condition that no two of them can be placed in the same column means that all of them must be placed in the first three rows, so $\pi_1 + \pi_2 + \pi_3 \geq \rho_1 + \rho_2 + \rho_3$. And so on. Because both π and ρ are partitions of (the same integer) n , we conclude that $\pi \succ \rho$.

Conversely, assume that π majorizes $\rho = [\rho_1, \rho_2, \dots, \rho_r]$. The proof that π occurs among the partitions built up from ρ in Theorem 4.51 is by induction on r . If $r = 1$, then $\pi = \rho$ and the result is clear. If $r > 1$ then, by Lemma 4.56, ρ_r symbols can be removed from the π_1 symbols constituting the column ends of the Ferrers diagram $F(\pi)$ in such a way that the result is the Ferrers diagram of a

new partition π' , and $\pi' > \rho' = [\rho_1, \rho_2, \dots, \rho_{r-1}]$. By the induction hypothesis, $F(\pi')$ occurs among the Ferrers diagrams constructed from ρ' via Theorem 4.51. Now, add ρ_r symbols to $F(\pi')$ to produce $F(\pi)$. Because $F(\pi')$ was obtained from $F(\pi)$ by detaching only the *ends* of columns, no two of these ρ_r symbols lie in the same column of $F(\pi)$. Thus, $K_{\pi, \rho} \neq 0$.

To prove the upper triangularity of the Kostka matrix, it suffices to show that π majorizes ρ , only if ρ comes before π in lexicographic order. If $\pi > \rho$, suppose k is the smallest positive integer such that $\pi_k \neq \rho_k$. Because

$$\pi_1 + \pi_2 + \dots + \pi_k \geq \rho_1 + \rho_2 + \dots + \rho_k,$$

it can only be that $\rho_k < \pi_k$. □

EXAMPLE 4.57 Because $K = (K_{\pi, \rho})$ is an integer upper triangular matrix with 1's on its diagonal, K^{-1} is an integer upper triangular matrix with 1's on its diagonal. If $n = 4$ then (Figure 4.6)

$$K = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 & 3 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

and

$$K^{-1} = \begin{pmatrix} 1 & -1 & 0 & 1 & -1 \\ 0 & 1 & -1 & -1 & 2 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

It follows from Equation (4.34) that

$$\chi_{\pi} = \sum_{\rho \vdash n} (K^{-1})_{\rho, \pi} 1_{\rho}^{S_n}.$$

From the fourth and fifth columns of K^{-1} we see, for example, that

$$\chi_{[2, 1^2]} = 1_{[4]}^{S_4} - 1_{[3, 1]}^{S_4} - 1_{[2^2]}^{S_4} + 1_{[2, 1^2]}^{S_4},$$

and

$$\varepsilon = \chi_{[1^4]} = -1_{[4]}^{S_4} + 2 \times 1_{[3, 1]}^{S_4} + 1_{[2^2]}^{S_4} - 3 \times 1_{[2, 1^2]}^{S_4} + 1_{[1^4]}^{S_4}.$$

□

In principle, it is now a straight-forward mechanical process to compute irreducible characters of S_n . The job is typically done with the help of so-called Young tableaux.¹² We will omit the details and merely describe the results.

¹²See, for example, [James & Kerber (1981)].

DEFINITION 4.58 Let $\pi = [\pi_1, \pi_2, \dots, \pi_r]$ be a partition of n . To each ordered pair (i, j) , $1 \leq i \leq r$, $1 \leq j \leq \pi_i$, there corresponds a box, B_{ij} , in the Ferrers diagram $F(\pi)$. Box B_{ij} determines a unique hook in $F(\pi)$ consisting of B_{ij} itself, all the boxes in row i of $F(\pi)$ to the right of B_{ij} , and all boxes in column j of $F(\pi)$ below B_{ij} . The number of boxes in the hook determined by B_{ij} is its length, $h_{ij} = (\pi_i - i) + (\pi_j^* - j) + 1$.

5	4	1
3	2	
2	1	

FIGURE 4.7 Hook lengths in $F([3, 2^2])$.

EXAMPLE 4.59 The numbers in the boxes of the Ferrers diagram $F([3, 2^2])$, in Figure 4.7, are the corresponding hook lengths. For example, $h_{12} = 4$ and $h_{21} = 3$. (A figure obtained by placing numbers in the boxes of a Ferrers diagram is a **Young tableau**.) $\square$

THEOREM 4.60 (*Frame-Robinson-Thrall Hook Length Formula*). If π is a partition of n , then the degree of the irreducible character of S_n corresponding to π is

$$\chi_\pi(e) = \frac{n!}{\prod_{i=1}^{L(\pi)} \prod_{j=1}^{\pi_i} h_{ij}},$$

where, recall, $L(\pi)$ is the length of π .

Beyond [Frame, Robinson & Thrall (1954)], proofs of Theorem 4.60 can be found in [Boerner (1970)], [James & Kerber (1981), p. 56], and [Sagan (1990) & (1991)].

EXAMPLE 4.61 Let $\pi = [3, 2^2]$. Then, from Theorem 4.60 and Figure 4.7, $\chi_\pi(e) = 7!/(5 \times 4 \times 1 \times 3 \times 2 \times 2 \times 1) = 21$.

If $\pi = [r, 1^{n-r}]$, then $F(\pi)$ consists of a single-hook. We will frequently write χ_r in place of $\chi_{[r, 1^{n-r}]}$, referring to it as a **single-hook** character. Because the product of the hook lengths of $F([r, 1^{n-r}])$ is

$$\prod_{i=1}^{n-r+1} \prod_{j=1}^{\pi_i} h_{ij} = n(r-1)!(n-r)!,$$

$\chi_r(e) = C(n-1, r-1)$. If $n = 5$ then, for example, $\chi_3(e) = C(4, 2) = 6$, confirming the value for $\chi_{[3,1^2]}(C_{[1^3]})$ in Figure 4.5. $\square$

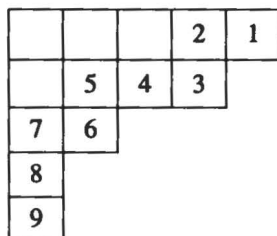


FIGURE 4.8 The border of $F([5, 4, 2, 1^2])$.

DEFINITION 4.62 Suppose $\pi \vdash n$. The **border** of $F(\pi)$ consists of those boxes whose right edge, bottom edge, or bottom right vertex belong to the geometric boundary of the diagram. A **border strip** (or **regular boundary part**) of $F(\pi)$ is a connected set of border boxes the deletion of which would result in another Ferrers diagram. If the boxes of a border strip overlap r rows of $F(\pi)$, then the **height** of the strip, $h = r - 1$, is the number of “vertical steps” that the border strip contains.

EXAMPLE 4.63 The numbered boxes in Figure 4.8 comprise the border of $F([5, 4, 2, 1^2])$; its 13 border strips are listed in Figure 4.9. $\square$

THEOREM 4.64 (Murnaghan-Nakayama Rule). Suppose $\pi, \rho \vdash n$. Let $k_1, k_2, \dots, k_s$ be the parts of ρ arranged in some fixed but arbitrary order. Consider all possible ways the Ferrers Diagram $F(\pi)$ can be reduced to nothing by successively deleting border strips of cardinalities $k_1, k_2, \dots, k_s$. If, altogether, the border strips occurring in the t -th way contain a total of v_t vertical steps, then

$$\chi_\pi(C_\rho) = \sum_t (-1)^{v_t}.$$

Apart from [Murnaghan (1937)] and [Nakayama (1940)], proofs of Theorem 4.64 can be found in [Boerner (1970)], [James & Kerber (1981)], and [Sagan (1991)].

EXAMPLE 4.65 In actual computations, one takes advantage of the fact that the parts of ρ may be listed in any (fixed) order. If, for example, $\pi = [5, 4, 2, 1^2]$, then (from Example 4.63 and Figure 4.9), no border strip of $F(\pi)$ consists of five boxes. It follows that $\chi_\pi(\sigma) = 0$ for any permutation $\sigma \in S_{13}$ whose disjoint cycle factorization contains a 5-cycle. $\square$

border strip	height
{1}	0
{1,2,3}	1
{1,2,3,4}	1
{1,2,3,4,5,6}	2
{1,2,3,4,5,6,7,8,9}	4
{3}	0
{3,4}	0
{3,4,5,6}	1
{3,4,5,6,7,8,9}	3
{6}	0
{6,7,8,9}	2
{8,9}	1
{9}	0

FIGURE 4.9

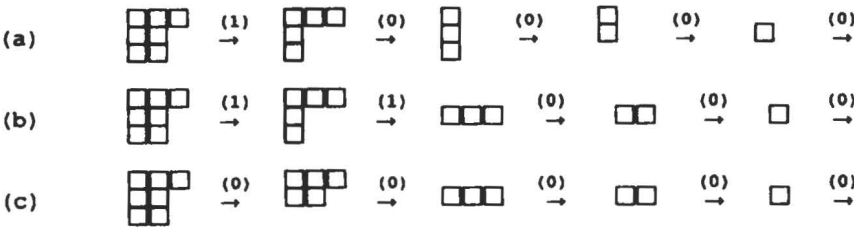


FIGURE 4.10

EXAMPLE 4.66 Let's use the Murnaghan-Nakayama rule to compute $\chi_{[3,2^2]}(C_{[2^2,1^3]})$, the value of the irreducible character $\chi_{[3,2^2]}$ on the permutations in S_7 of cycle type $[2^2, 1^3]$. Set $k_1 = k_2 = 2$ and $k_3 = k_4 = k_5 = 1$. There are three ways to annihilate $F([3, 2^2])$ by successively deleting border strips of cardinalities 2, 2, 1, 1, and 1. They are illustrated in Figure 4.10. The border strips in the first way (Figure 4.10(a)) contain $1 + 0 + 0 + 0 + 0 = 1$ vertical steps altogether. The (total) number of vertical steps in the second way (Figure 4.10(b)) is $1 + 1 + 0 + 0 + 0 = 2$. The third way involves no vertical steps at all. Thus, for any permutation $\sigma \in C_{[2^2,1^3]}$,

$$\chi_{[3,2^2]}(\sigma) = (-1)^1 + (-1)^2 + (-1)^0 = 1.$$

What about $\chi_{[3,2^2]}(C_{[2^3,1]})$? Let $k_1 = k_2 = k_3 = 2$ and $k_4 = 1$. Coincidentally, there are again three ways of reducing $F([3, 2^2])$ to nothing, this time by successively deleting border strips of cardinalities 2, 2, 2, and 1. They are illustrated in Figure 4.11. The deleted border strips in Figure 4.11(a) contain $1+0+1+0 = 2$ vertical steps altogether; the total number in Figure 4.11(b) is $1+1+0+0 = 2$; and in Figure 4.11(c), there are $0+0+0+0 = 0$ vertical steps altogether. Thus, $\chi_{[3,2^2]}(C_{[2^3,1]}) = (-1)^2 + (-1)^2 + (-1)^0 = 3$. $\square$

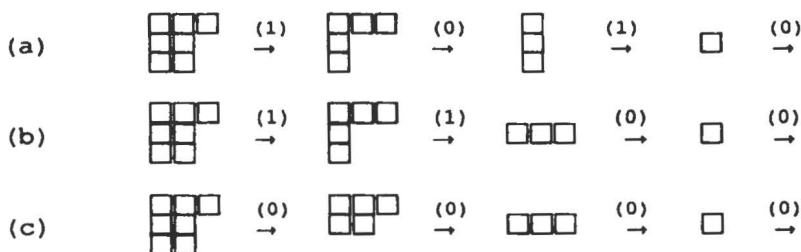


FIGURE 4.11

Suppose χ is an irreducible character of G . If H is a subgroup of G , then the restriction of χ to H may no longer be irreducible. When $G = S_n$ and $H = S_{[n-1,1]} \cong S_{n-1}$, the situation is described by the so-called **branching theorem**:

COROLLARY 4.67 Suppose $\pi = [\pi_1, \pi_2, \dots, \pi_r]$ is a (nonincreasing) partition of $n > 1$. Upon restriction to S_{n-1} ,

$$\chi_\pi = \chi_1 + \chi_2 + \dots + \chi_r,$$

where $\chi_i = 0$ if $\pi_{i+1} = \pi_i$, and χ_i is the irreducible character of S_{n-1} corresponding to the partition

$$[\pi_1, \pi_2, \dots, \pi_{i-1}, \pi_i - 1, \pi_{i+1}, \dots, \pi_r], \quad (4.35)$$

otherwise.

Proof Because we are viewing S_{n-1} as $\{\sigma \in S_n : \sigma(n) = n\}$, we may begin the Murnaghan-Nakayama calculation of $\chi(\sigma)$ by deleting from $F(\pi)$, in all possible ways, a border strip of cardinality 1 (a single box). The results are Ferrers diagrams of partitions of $n - 1$ of the type exhibited in (4.35), where $\pi_i > \pi_{i+1}$. This completes the proof because the vertical step contribution from deleting a single box is 0. $\square$

COROLLARY 4.68 *If $\sigma \in S_n$ and $\chi \in I(S_n)$, then $\chi(\sigma)$ is an integer.*

Proof The result is immediate either from the Murnaghan-Nakayama Rule or the fact that the inverse Kostka matrix is integral. $\square$

Corollary 4.68 does not extend to all finite groups.¹³ (If χ is a linear character of G , it is a homomorphism into the nonzero complex numbers. Thus, χ needn't be real.)

THEOREM 4.69 *Let G be a finite group. Suppose $\chi \in I(G)$. If $\chi(e) > 1$, there exists a $\sigma \in G$ such that $\chi(\sigma) = 0$.*

Proof If $G = S_n$, then $1 = (\chi, \chi)_G$ is equivalent to

$$\begin{aligned} o(G) &= \sum_{\sigma \in G} \chi(\sigma^{-1})\chi(\sigma) \\ &= \sum_{\sigma \in G} \chi(\sigma)^2. \end{aligned}$$

By Corollary 4.68, the right-hand side of this equation is a sum of squares of $o(G)$ integers. Because one of them (namely $\chi(e)$) is greater than one, they cannot all be nonzero. For arbitrary finite groups, the analogous statement is

$$o(G) = \sum_{\sigma \in G} |\chi(\sigma)|^2. \quad (4.36)$$

While $\chi(\sigma) = \text{tr}(A(\sigma))$ need not be an integer, because it is a sum of roots of unity (the eigenvalues of $A(\sigma)$), it is an "algebraic integer". This, together with Equation (4.36) is enough to establish the result. (Details can be found in [Feit (1967), p. 36] or [Isaacs (1976), p. 40].) $\square$

The final theorems of the chapter are useful technical results.

THEOREM 4.70 *Let χ be a character of the finite group G . Then χ is afforded by a unitary representation.*

¹³The character table of G is integral if and only if the following condition is satisfied: For all $\sigma, \tau \in G$, the cyclic subgroup $\langle \sigma \rangle$ is conjugate to $\langle \tau \rangle$ only if σ is conjugate to τ .

Proof Suppose $\chi(e) = n$. Let $\sigma \rightarrow A(\sigma)$, $\sigma \in G$, be a representation of G affording χ . Let

$$X = \sum_{\tau \in G} A(\tau)A(\tau)^*.$$

By Lemma 2.43, X is positive definite hermitian, so there is a unitary matrix U such that $X = U^*DU$, where D is a diagonal matrix with positive diagonal entries. Let $D^{1/2}$ be the positive definite square-root of D . If $H = (D^{-1/2})U$, then $(H^{-1})^* = D^{1/2}U$. Let $B(\sigma) = HA(\sigma)H^{-1}$, $\sigma \in G$. Then

$$\begin{aligned} B(\sigma)B(\sigma)^* &= (HA(\sigma)H^{-1})(HA(\sigma)H^{-1})^* \\ &= (HA(\sigma)U^*D^{1/2})(D^{1/2}UA(\sigma)^*H^*) \\ &= HA(\sigma)XA(\sigma)^*H^* \\ &= HXH^* \\ &= D^{-1/2}UXU^*D^{-1/2} \\ &= D^{-1/2}DD^{-1/2} \\ &= I_n, \end{aligned}$$

because $A(\sigma)XA(\sigma)^* = X$, $\sigma \in G$. □

THEOREM 4.71 *Let χ be an irreducible character of the finite group G . Then $\chi(e)$ exactly divides $o(G)$.*

This theorem is a nice illustration of the vitality to be found at the boundary between group representation theory and algebraic number theory. Proofs can be found in any of the standard books on representation theory.

There are many applications of group representation theory in physics and quantum chemistry. One of them will emerge in Chapter 6. Others can be found, for example, in [Hollas (1967)], [James & Liebeck (1993)], [Lomont (1993)], and [Schensted (1976)].

Exercises

- Write out the permutation matrix $A(\sigma) = (\delta_{i, \sigma(j)})$ for each $\sigma \in S_3$.
- Let $S = \{A(v) : v \in N\} \subset \mathbf{C}_{n,n}$. Suppose S spans $\mathbf{C}_{n,n}$.
 - Prove that S is irreducible.
 - Prove that $GL(n, \mathbf{C})$ is irreducible.
- Use the approach of Exercise 2 to show that the representation of S_3 in Example 4.3 is irreducible.
- Let A be the representation of S_3 in Example 4.3. Find an invertible 2-by-2 matrix U such that $U^{-1}A(\sigma)U$ is diagonal for all $\sigma \in A_3 = \{e_3, (123), (132)\}$.
- Let G be the multiplicative group of nonzero complex numbers. Then the typical element of G is $z = x + iy$, where x and y are real numbers, not both zero. Prove that

$$z \rightarrow \begin{pmatrix} x & -y \\ y & x \end{pmatrix}$$

is a faithful irreducible representation of G .

- Let $\{A(v) : v \in N\}$ be a set of pairwise commuting n -by- n matrices.
 - Prove that there exists a matrix $U \in GL(n, \mathbf{C})$ such that $U^{-1}A(v)U$ is upper triangular for all $v \in N$. (Hint: Corollary 4.17.)
 - Prove that the matrix U in part (a) can be chosen to be unitary. (Compare with Exercise 34 in Chapter 2.)
- In the manner of Example 4.20,
 - use the representation A in Example 4.3 to confirm Equation (4.9) when $i = t = 1$ and $s = j = 2$.
 - use the representation A in Example 4.3 and the representation $B(\sigma) = (\varepsilon(\sigma))$, $\sigma \in S_3$ to confirm Equation (4.10) when $t = j = 1$, for all four choices of i and s .
- One way to describe finite groups is by means of generators and relations. The group of 24 rotational symmetries of the cube, for example, is generated by elements x and y with defining relations $x^2 = y^3 = (xy)^4 = e$.
 - Show that these are the generators and relations for S_4 . (This gives another proof that the group of rotational symmetries of the cube is isomorphic to S_4 .)
 - Show that $X = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ and $Y = \frac{1}{4} \begin{pmatrix} -2 & -3 \\ 4 & -2 \end{pmatrix}$ satisfy the relations $X^2 = Y^3 = (XY)^2 = I_2$.
 - Show that $x \rightarrow X$ and $y \rightarrow Y$ extends to a representation of S_4 .

- d. Is the representation defined in part (c) faithful?
- e. Is the representation defined in part (c) irreducible?
- 9. Prove Theorem 4.21. (Hint: Because B is a representation, $b_{ij}(\sigma\tau) = \sum_k b_{ik}(\sigma)b_{kj}(\tau)$.)
- 10. Prove the converse of Theorem 4.22: If $\sigma \rightarrow A(\sigma) = (a_{ij}(\sigma))$ is a representation in which the n^2 functions $a_{ij}: G \rightarrow \mathbb{C}$ are linearly independent, then A is irreducible. (Compare with Exercise 2.)
- 11. Prove Theorem 4.23 for finite groups. (Hint: Mimic the proof of Burnside's Theorem.)
- 12. Prove Theorem 4.26.
- 13. Prove Theorem 4.29.
- 14. Let χ be a character of the finite group G . If $\sigma, \tau \in G$, prove that $\chi(\sigma\tau) = \chi(\tau\sigma)$.
- 15. Prove Theorem 4.36. (Hint: If A is any representation of G , then $A(\tau^{-1}) = A(\tau)^{-1}$, $\tau \in G$.)
- 16. Let A be the (left) regular representation of $G = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$. Prove that

$$\sum_{i=1}^n A(\sigma_i) = J,$$

the n -by- n matrix each of whose entries is 1.

- 17. Let G be the subset of $C_{n,n}$, $n > 1$, consisting of the invertible lower triangular matrices. (A lower triangular matrix is invertible if and only if there are no zeros on its main diagonal.)
 - a. Show that G is a multiplicative group.
 - b. Show that the representation $A \rightarrow A$, $A \in G$, is reducible but not fully reducible. (Explain why this does not violate Maschke's Theorem.)
 - c. Show that $\dim(\langle A: A \in G \rangle) < n^2$, where n is the degree of the representation $A \rightarrow A$, $A \in G$.
- 18. Let ξ be the character afforded by the representation A of Example 4.3. An indirect transitivity argument was used in Example 4.33 to prove that ξ is irreducible.
 - a. Show by a direct computation that $(\xi, \xi)_{S_3} = 1$.
 - b. Prove that $I(S_3) = \{1, \varepsilon, \xi\}$, where ε is the alternating character and ξ is the character in part a.
- 19. Let A be a representation of the finite group G . Define $B(\sigma) = A(\sigma^{-1})^*$, $\sigma \in G$; $C(\sigma) = A(\sigma^{-1})^t$, the transpose of $A(\sigma^{-1})$, $\sigma \in G$; and $D(\sigma) = \overline{A(\sigma)}$, the complex conjugate of $A(\sigma)$, $\sigma \in G$.

- a. Prove that B , C , and D are representations of G .
 - b. Prove that B is equivalent to A .
 - c. Prove that C is equivalent to D . (Hint: Theorem 4.70.) Representations A and C are "contragredient".
20. Let G be a finite group. If $\sigma \rightarrow A(\sigma) = (a_{ij}(\sigma))$ is a representation of G , define $\Sigma_A: G \rightarrow \mathbb{C}$ by

$$\Sigma_A(\sigma) = \sum_{i,j=1}^n a_{ij}(\sigma).$$

If A and B are irreducible unitary representations of G , prove that

$$(\Sigma_A, \Sigma_B)_G = \begin{cases} n & \text{if } A = B \text{ has degree } n, \\ 0 & \text{if } A \text{ and } B \text{ are inequivalent.} \end{cases}$$

21. Prove that $GL(n, \mathbb{C})$ contains a finite irreducible subgroup, $n \geq 1$.
22. Let $\sigma \rightarrow A(\sigma) = (a_{ij}(\sigma))$ be an irreducible representation of the finite group G . Prove that

$$\sum_{\tau \in C_G(\sigma)} a_{ij}(\tau) = \delta_{i,j} \chi(\sigma) o(C_G(\sigma)) / \chi(e),$$

where χ is the character afforded by A .

23. The idea behind this Exercise is to prove Corollary 4.39 from first principles. So, let G be a finite group.
 - a. If G is abelian, use Corollary 4.17 to prove that $\chi(e) = 1$ for every $\chi \in I(G)$.
 - b. If $\chi(e) = 1$ for every $\chi \in I(G)$, use the fact that the regular representation of G is faithful to prove that G is abelian.
24. Construct the character table for
 - a. the cyclic group of order 4.
 - b. the noncyclic group of order 4.
25. Let $G = D_4$ (from Example 4.42).
 - a. Show that $G = \langle (1243), (12)(34) \rangle$, the group generated by (1243) and $(12)(34)$.
 - b. Show that

$$A((1243)) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \quad \text{and} \quad A((12)(34)) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

determines an irreducible representation of G that affords the character χ_5 of Figure 4.3.

26. The “quaternion” group of order 8 is generated by elements x and y that satisfy $x^4 = e$, $x^2 = y^2$, and $y^{-1}xy = x^3$. Show that,
- the quaternion group is not isomorphic to D_4 . (See Exercise 25.)
 - (by a judicious choice of names for their conjugacy classes and characters) the character tables for the two groups are identical. (Hint: Figure 4.3.)
27. Find the character table for
- the cyclic group of order 8.
 - the group of order 8 determined by the generators and relations $x^2 = y^4 = e$ and $xyx = y$.
 - the group G of order 8 that satisfies $\sigma^2 = e$, $\sigma \in G$.
28. Let χ be an irreducible character of degree 5 of a finite group G . Prove that

$$\sum_{\sigma \in G} \chi(\sigma) = 0.$$

29. Let χ be an irreducible character of the finite group G .
- Prove that $|\chi(\sigma)| \leq \chi(e)$ for all $\sigma \in G$.
 - Let $Z_\chi(G) = \{\sigma \in G : |\chi(\sigma)| = \chi(e)\}$. Prove that $Z_\chi(G)$ is a subgroup of G .
 - Prove that $\xi(\sigma) = \chi(\sigma)/\chi(e)$, $\sigma \in Z_\chi(G)$, defines a character of $Z_\chi(G)$.
 - Prove that $Z(G) \subset Z_\chi(G)$, where $Z(G)$ is the center of G .
 - Prove that

$$Z(G) = \bigcap_{\chi \in I(G)} Z_\chi(G).$$

(Hint: Consider the regular representation of G .)

30. Let H be a normal subgroup of the finite group G . Suppose $\{\sigma_1, \sigma_2, \dots, \sigma_r\}$, $r = [G : H]$, is a system of representatives for the distinct cosets of H in G . Let A be an irreducible representation of the quotient group G/H . If $\sigma \in G$, define $B(\sigma) = A(\sigma_i H)$, where $\sigma_i H$ is the coset of G containing σ .
- Prove that B is a representation of G .
 - Prove that every proper normal subgroup of G is contained in the kernel of some nonprincipal irreducible representation of G .
31. Let H be a subgroup of the finite group G . Suppose $\{\sigma_i : 1 \leq i \leq r\}$ is a system of representatives for the distinct left cosets of H in G . Let $\mu_1, \mu_2, \dots, \mu_r$ be fixed but arbitrary elements of H . If A is a representation of H , let $\hat{A}$ be the representation of G induced by A using $\{\sigma_i : 1 \leq i \leq r\}$, and $\bar{A}$ the representation of G induced by A using coset representatives $\{\sigma_i \mu_i : 1 \leq i \leq r\}$. Find an invertible matrix U such that $U^{-1} \hat{A}(\tau) U = \bar{A}(\tau)$, $\tau \in G$.

32. Let G be a finite group. If $\sigma, \tau \in G$, their commutator is $\sigma\tau\sigma^{-1}\tau^{-1}$. The commutator subgroup of G , denoted G' , is the subgroup of G generated by its commutators.
- Prove that G is abelian if and only if $G' = \{e\}$.
 - Prove that G' is a normal subgroup of G .
 - Prove that G/G' is abelian.
 - If χ_1 and χ_2 are different linear characters of G/G' , prove that they extend (as in Exercise 30a) to different characters of G .
 - If G has m different linear characters, prove that $m \geq [G : G']$.
 - Suppose χ is a linear character of G . Prove that $\chi(\sigma) = 1, \sigma \in G'$.
 - Prove that the number of linear characters of G is equal to $[G : G']$.
33. Let K be a subgroup of H and H a subgroup of G . If χ is a character of K , prove that $(\chi^H)^G = \chi^G$.
34. Prove that $\varepsilon_{S_n}^{\varepsilon} = \varepsilon(1_{S_n})$.
35. Prove that the principal and alternating characters are the only characters of S_n of degree 1
- using the result in Exercise 32g.
 - using the Frame-Robinson-Thrall hook length formula.
36. Use the Murnaghan-Nakayama Rule to
- show that $\chi_{[n]}$ is the principle character of S_n .
 - show that $\chi_{[1^n]} = \varepsilon$, the alternating character of S_n .
 - confirm that $\chi_{[5,2]}([2, 1^5]) = 6$.
 - confirm that $\chi_{[5,1^2]}([2, 1^5]) = \chi_{[4,2,1]}([2, 1^5])$.
 - confirm that $\chi_{[n-1,1]}(\sigma) = F(\sigma) - 1$, where F is the fixed point character of S_n .
 - confirm the values in the character table of S_4 given in Figure 4.12.
 - compute the values of $\chi_{[3,2]}(\sigma), \sigma \in S_5$. (Hint: Use Figure 4.5 to confirm your answers.)

	$C_{[1^4]}$	$C_{[2,1^2]}$	$C_{[2^2]}$	$C_{[3,1]}$	$C_{[4]}$
$\chi_{[1^4]}$	1	-1	1	1	-1
$\chi_{[2,1^2]}$	3	-1	-1	0	1
$\chi_{[2^2]}$	2	0	2	-1	0
$\chi_{[3,1]}$	3	1	-1	0	-1
$\chi_{[4]}$	1	1	1	1	1

FIGURE 4.12 The Character Table of S_4 .

37. Let ξ be the restriction of $\chi_{[3,2]}$ to S_4 .
- Use your results from Exercise 36g and Equation (4.18) to prove that ξ is reducible.
 - Use your results from Exercises 36f and 36g to express ξ as a nonnegative integral linear combination of the irreducible characters of S_4 .
 - Use the branching theorem to express ξ as a nonnegative integral linear combination of the irreducible characters of S_4 .
38. Use the Frame-Robinson-Thrall hook length formula to confirm the values for the character degrees given in Figure 4.5. (Hint: $C_{[1^5]} = \{e\}$.)
39. Suppose χ is an irreducible character of the finite group G . Let $Z_\chi(G)$ be the group defined in Exercise 29b and ξ the character of $Z_\chi(G)$ defined in Exercise 29c.
- Prove that $\chi(e)^2 \leq [G : Z_\chi(G)]$.
 - Prove that equality holds in part (a) if and only if χ is the only irreducible character of G whose restriction to $Z_\chi(G)$ contains ξ as a component. (Hint: Consider ξ^G .)
 - Prove that equality holds in part (a) if and only if $Z_\chi(G) = \{\sigma \in G : \chi(\sigma) \neq 0\}$.
40. Let H be a subgroup of G . Let χ be a linear character of H . For each $\sigma \in G$, define χ^σ be the character of $\sigma H \sigma^{-1}$ defined by $\chi^\sigma(\sigma \tau \sigma^{-1}) = \chi(\tau)$, $\tau \in H$. Prove that χ^G is an irreducible character of G if and only if, for all $\sigma \in G \setminus H$, χ^σ and χ restrict to different characters of $H \cap \sigma H \sigma^{-1}$.
41. Let G be a finite group. Let V be the complex vector space of all formal complex linear combinations of the elements of G , that is,

$$V = \left\{ \sum_{\sigma \in G} c_\sigma \sigma : c_\sigma \in \mathbb{C}, \sigma \in G \right\}.$$

- Prove that V is isomorphic to $\mathbb{C}G$.
- Prove that V is a ring under the multiplication

$$\left(\sum_{\sigma \in G} c_\sigma \sigma \right) \left(\sum_{\tau \in G} d_\tau \tau \right) = \sum_{\mu \in G} \left(\sum_{\tau \in G} c_{\mu\tau^{-1}} d_\tau \right) \mu.$$

- Define

$$t(G, \chi) = \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma) \sigma, \quad \chi \in I(G).$$

If $\chi, \xi \in I(G)$, prove that $t(G, \chi)t(G, \xi) = \delta_{\chi, \xi} t(G, \chi)$.

- d. Prove that $t(G, \chi) \in Z(V)$, the center of the ring V . (In other words, prove that $vt(G, \chi) = t(G, \chi)v$ for all $v \in V$.)
 - e. Prove that the vector space $Z(V)$ is isomorphic to $Z(CG)$.
 - f. Prove that $\{t(G, \chi) : \chi \in I(G)\}$ is a basis for the vector space $Z(V)$.
 - g. Let ξ be the character of $Z_\chi(G)$ defined in Exercise 29c. Prove that $t(Z_\chi(G), \xi) = t(G, \chi)$ if and only if $[G : Z_\chi(G)] = \chi(e)^2$.
42. Suppose H and K are subgroups of the finite group G . Define a relation among the elements of G as follows: If $\sigma_1, \sigma_2 \in G$, then $\sigma_1 \sim \sigma_2$ means there exist $\tau_1 \in H$ and $\tau_2 \in K$ such that $\tau_1\sigma_1 = \sigma_2\tau_2$.
- a. Prove that $\sim$ is an equivalence relation.
 - b. Prove that the equivalence class containing σ is the double coset $H\sigma K = \{\tau_1\sigma\tau_2 : \tau_1 \in H \text{ and } \tau_2 \in K\}$.
43. Let $H = \{e_4, (1243), (14)(23), (1342)\}$, $K = \{e_4, (14), (23), (14)(23)\}$, and $G = S_4$.
- a. Show that the double coset $He_4K = H \cup H(14) = D_4$, the group of order 8 in Example 4.42. (See Exercise 42 for the definition of "double coset".)
 - b. Show that $H(12)K = H(12) \cup H(13) \cup H(24) \cup H(34)$.
 - c. Give an example to show that $o(H\sigma_1K)$ need not equal $o(H\sigma_2K)$.
44. Let G be a doubly transitive subgroup of S_n . Suppose $\tau \in G \setminus G_x$ for some fixed but arbitrary $x \in \{1, 2, \dots, n\}$. Prove that G is the disjoint union of the double cosets $G_x e_n G_x$ and $G_x \tau G_x$. (See Exercise 42 for the definition of "double coset".)
45. Suppose $C_1, C_2, \dots, C_r$ are the different conjugacy classes and $\chi_1, \chi_2, \dots, \chi_r$ the different irreducible characters of the finite group G . Let $U = (u_{ij})$ be the r -by- r matrix whose (i, j) -entry is

$$u_{ij} = \chi_i(C_j)(o(C_j)/o(G))^{1/2}.$$

- a. Prove that $UU^* = I_r$ is equivalent to the Orthogonality Relations of the First Kind.
 - b. Prove that $U^*U = I_r$ is equivalent to the Orthogonality Relations of the Second Kind.
46. Let $G = S_3$ and $H = \{e_3, (13)\}$.
- a. Show that $\sigma_1 = e_3$, $\sigma_2 = (12)$, and $\sigma_3 = (23)$ are representatives for the different left cosets of H in G .
 - b. If $\tau = (123)$, show that $\tau\sigma_1H = \sigma_3H$, $\tau\sigma_2H = \sigma_1H$, and $\tau\sigma_3H = \sigma_2H$.
 - c. If $\hat{A}$ is the representation of G induced by the principal representation of H , show that

$$\hat{A}((123)) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

- d. Express the character 1_H^G as a sum of irreducible characters of S_3 .
47. Let $G = S_4$ and $H = G_4 \cong S_3$ be the stabilizer subgroup of 4. Let χ be the irreducible character of H of degree 2. Express χ^G as a sum of irreducible characters of S_4 .
48. Let σ and τ be elements of the finite group G . Prove that $\chi(\sigma) = \chi(\tau)$, for all $\chi \in I(G)$, if and only if σ and τ are conjugate in G .
49. Let G be a finite group. Suppose $\tau \in G$. Prove that $\chi(\tau)$ is real, for all $\chi \in I(G)$, if and only if τ is conjugate in G to τ^{-1} .
50. Prove Corollary 4.54.
51. Prove that the Kostka coefficients satisfy

$$(1_{\pi}^{S_n}, 1_{\rho}^{S_n})_n = \sum_{\nu \vdash n} K_{\nu, \pi} K_{\nu, \rho}.$$

52. Knowing that the Kostka matrix is upper-triangular, show how the result of Exercise 51 can be used recursively to obtain the table in Figure 4.6.
53. Let n be fixed but arbitrary. For $1 \leq r \leq n$, denote by χ_r the single-hook character of S_n corresponding to the partition $[r, 1^{n-r}]$. Suppose $\tau \in S_n$ has cycle type $[n]$, that is τ is a full n -cycle.
- a. If $\xi \in I(S_n)$, prove that

$$\xi(\tau) = \begin{cases} 0, & \text{if } \xi \text{ is not a single-hook character.} \\ (-1)^{n-r}, & \text{if } \xi = \chi_r \end{cases}$$

- b. Let ρ be a partition of n having $s = L(\rho)$ parts. If $\pi = [r, 1^{n-r}]$, prove that the Kostka coefficient $K_{\pi, \rho} = C(s-1, n-r)$.
54. Let $C_{[n]}$ be the conjugacy class of S_n comprised of those permutations whose disjoint cycle factorizations consist of a single cycle of length n . Let f_n be the characteristic function of $C_{[n]}$, that is, $f_n(\sigma) = 1$, if $\sigma \in C_{[n]}$, and 0, otherwise.
- a. Show that

$$f_n = \frac{1}{n} \sum_{\pi \vdash n} \chi_{\pi}(C_{[n]}) \chi_{\pi}.$$

- b. Show that $f_5 = (1 - \chi_{[4,1]} + \chi_{[3,1^2]} - \chi_{[2,1^3]} + \epsilon)/5$. (Hint: Exercise 53.)

55. Suppose π and ν are different partitions of n . Let $\tau \in S_n$ be a transposition. Suppose π majorizes ν .
- Prove that $\chi_\pi(\tau)/\chi_\pi(e) > \chi_\nu(\tau)/\chi_\nu(e)$.
 - If $\nu = \pi^*$, prove that $\chi_\pi(\tau) > 0$.
56. Show how Theorem 4.71 can be used to eliminate the spurious possibilities for the character degrees of S_7 in Example 4.43 (page 94).
57. According to Theorem 4.71, $o(S_n)/\chi_\pi(e)$ is an integer, $\pi \vdash n$. Describe a formula for this integer in terms of π .

CHAPTER 5

Tensor Spaces

Multilinear algebra, proper, begins with the study of tensor spaces. In the most general setting, this involves dealing with m vector spaces, each having its own basis and its own inner product. Keeping track of it all requires what Elie Cartan called “une débauche d’indices”, an intimidating proliferation of superscripts and subscripts. The peak of this mountain of notation occurs when matrix representations of linear transformations on the various vector spaces are assembled to produce the matrix representation of a linear transformation on the tensor space. Following the introduction of the induced inner product, the going will be enormously simplified by setting all m vector spaces equal. (The reader may find it useful to introduce this simplification earlier and rewrite difficult passages setting $V_1 = V_2 = \cdots = V_m = V$.)

Let $V_1, V_2, \dots, V_m$ be finite dimensional complex vector spaces. Their cartesian product is the set

$$V_1 \times V_2 \times \cdots \times V_m = \{(v_1, v_2, \dots, v_m) : v_i \in V_i, \ 1 \leq i \leq m\}.$$

Under componentwise addition and scalar multiplication defined by

$$\begin{aligned} c(v_1, v_2, \dots, v_m) + d(w_1, w_2, \dots, w_m) \\ = (cv_1 + dw_1, cv_2 + dw_2, \dots, cv_m + dw_m), \end{aligned}$$

$V_1 \times V_2 \times \cdots \times V_m$ is a vector space.

LEMMA 5.1 *Suppose vector space V_i has dimension n_i , $1 \leq i \leq m$. Then the dimension of $V_1 \times V_2 \times \cdots \times V_m$ is $n_1 + n_2 + \cdots + n_m$.*

Proof If $\{e_{ij} : 1 \leq j \leq n_i\}$ is a basis of V_i , $1 \leq i \leq m$, then

$$\begin{array}{c} i\text{-th component} \\ \downarrow \\ \{(0, 0, \dots, 0, e_{ij}, 0, \dots, 0) : 1 \leq j \leq n_i, 1 \leq i \leq m\} \end{array}$$

is a basis of $V_1 \times V_2 \times \dots \times V_m$. □

DEFINITION 5.2 Let $V_1, V_2, \dots, V_m$, and W be vector spaces. A function $f : V_1 \times V_2 \times \dots \times V_m \rightarrow W$ is **m -linear** (or **multilinear**) if f is linear separately in each component of $V_1 \times V_2 \times \dots \times V_m$. Thus, f is m -linear if

$$\begin{array}{c} i\text{-th component} \\ \downarrow \\ f(v_1, v_2, \dots, cu_i + dw_i, \dots, v_m) \\ = cf(v_1, v_2, \dots, u_i, \dots, v_m) + df(v_1, v_2, \dots, w_i, \dots, v_m), \end{array}$$

$c, d \in \mathbb{C}$, $1 \leq i \leq m$.

EXAMPLES 5.3 (i) Let $A \in \mathbb{C}_{m,n}$ be fixed but arbitrary. Define $f : \mathbb{C}_{1,m} \times \mathbb{C}_{n,1} \rightarrow \mathbb{C}$ by $f(x, y) = xAy$. Then f is 2-linear, (or **bilinear**). (ii) The determinant of an n -by- n matrix is an n -linear function of its rows (or columns). (iii) Let $f_i : V_i \rightarrow \mathbb{C}$ be a fixed but arbitrary linear functional on V_i , $1 \leq i \leq m$. (Then $f_i \in V'_i = L(V_i, \mathbb{C})$, the dual space of V_i .) The function

$$\Theta : V_1 \times V_2 \times \dots \times V_m \rightarrow \mathbb{C},$$

defined by

$$\Theta(v_1, v_2, \dots, v_m) = \prod_{i=1}^m f_i(v_i),$$

is m -linear. The notation $\prod f_i$ will be used to denote Θ . (iv) Let $v_i \in V_i$ be fixed but arbitrary vectors, $1 \leq i \leq m$. Define

$$\Psi : L(V_1, \mathbb{C}) \times L(V_2, \mathbb{C}) \times \dots \times L(V_m, \mathbb{C}) \rightarrow \mathbb{C}$$

by

$$\Psi(f_1, f_2, \dots, f_m) = \prod_{i=1}^m f_i(v_i).$$

Then Ψ is m -linear. (v) Suppose $f : W_1 \times W_2 \times \dots \times W_m \rightarrow W$ is m -linear. If $T_i \in L(V_i, W_i)$, is a fixed but arbitrary linear transformation, $1 \leq i \leq m$, then $g : V_1 \times V_2 \times \dots \times V_m \rightarrow W$, defined by

$$g(v_1, v_2, \dots, v_m) = f(T_1(v_1), T_2(v_2), \dots, T_m(v_m)),$$

is m -linear. (vi) Let $f: V_1 \times V_2 \times \cdots \times V_m \rightarrow V$ be m -linear. If $T \in L(V, W)$, then $Tf: V_1 \times V_2 \times \cdots \times V_m \rightarrow W$ is m -linear. $\square$

If $T \in L(V, W)$, then T is completely and uniquely determined by its action on a basis of V . Indeed, it is common to define a linear transformation by describing its action on a basis and saying the magic words, "linear extension". As we now see, multilinear functions behave analogously.

THEOREM 5.4 (Multilinear Extension). *Let $\{e_{ij}: 1 \leq j \leq n_i\}$ be a basis of vector space V_i , $1 \leq i \leq m$. Then there is precisely one multilinear function $f: V_1 \times V_2 \times \cdots \times V_m \rightarrow W$ that takes prescribed values on the elements of*

$$\{(e_{1j_1}, e_{2j_2}, \dots, e_{mj_m}) : 1 \leq j_i \leq n_i, 1 \leq i \leq m\}. \quad (5.1)$$

Note that the $n_1 \times n_2 \times \cdots \times n_m$ vectors in (5.1) typically comprise more than a basis of the $n_1 + n_2 + \cdots + n_m$ dimensional vector space $V_1 \times V_2 \times \cdots \times V_m$.

Proof Let $f: V_1 \times V_2 \times \cdots \times V_m \rightarrow W$ be m -linear. Suppose

$$v_i = \sum_{j=1}^{n_i} a_{ij} e_{ij}, \quad 1 \leq i \leq m,$$

are fixed but arbitrary vectors. Then

$$\begin{aligned} & f(v_1, v_2, \dots, v_m) \\ &= f\left(\sum_{j=1}^{n_1} a_{1j} e_{1j}, \sum_{j=1}^{n_2} a_{2j} e_{2j}, \dots, \sum_{j=1}^{n_m} a_{mj} e_{mj}\right) \\ &= \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} \cdots \sum_{j_m=1}^{n_m} a_{1j_1} a_{2j_2} \cdots a_{mj_m} f(e_{1j_1}, e_{2j_2}, \dots, e_{mj_m}), \end{aligned} \quad (5.2)$$

by m -linearity. Because the coefficients, $\prod a_{ij_i}$, depend only on $v_1, v_2, \dots, v_m$, we see that f is completely determined by the values

$$f(e_{1j_1}, e_{2j_2}, \dots, e_{mj_m}), \quad 1 \leq j_i \leq n_i, \quad 1 \leq i \leq m.$$

In particular, if another m -linear function, $g: V_1 \times V_2 \times \cdots \times V_m \rightarrow W$, agreed with f on the vectors in (5.1), then g would necessarily be identical to f . On the other hand, if f is an arbitrary function from the set of vectors in (5.1) into a vector space W , then f can be extended to a (unique) m -linear function from $V_1 \times V_2 \times \cdots \times V_m$ into W . The value of the extended function on $(v_1, v_2, \dots, v_m)$ is obtained from Equation (5.2). $\square$

The process of extending a function defined on the elements of (5.1) to an m -linear function of $V_1 \times V_2 \times \cdots \times V_m$ is called **multilinear extension**.

Having found an analogy between linear and multilinear functions, we now discuss a dissimilarity. If $T \in L(V, W)$, then the image of T (namely, $T(V) = \{T(v) : v \in V\}$) is a subspace of W . The image of a multilinear function, on the other hand, need not be a subspace.

EXAMPLE 5.5 Let $V_1 = V_2 = V$ be a two-dimensional vector space with basis $\{e_1, e_2\}$. Let E_{ij} be the 2-by-2 matrix whose only nonzero entry is a 1 in position (i, j) . Define $f: V \times V \rightarrow \mathbf{C}_{2,2}$ by $f(e_i, e_j) = E_{ij}$, $1 \leq i, j \leq 2$, and multilinear extension. Let $u = c_1e_1 + c_2e_2$ and $v = d_1e_1 + d_2e_2$ be fixed but arbitrary vectors in V . Then, from Equation (5.2) and the definition of f ,

$$f(u, v) = \begin{pmatrix} c_1d_1 & c_1d_2 \\ c_2d_1 & c_2d_2 \end{pmatrix}.$$

Observe that the image of f contains $\{E_{ij} : 1 \leq i, j \leq 2\}$, a basis of $\mathbf{C}_{2,2}$. Thus, if the image were a subspace, it would be all of $\mathbf{C}_{2,2}$. However, $\det(f(u, v)) = 0$ for all $u, v \in V$. $\square$

DEFINITION 5.6 Let f be a multilinear function. The **reach** of f is the linear closure of its image.

It is natural to wonder how we might use our extensive knowledge of linear functions to study multilinear functions. The answer is anticipated by the following.

DEFINITION 5.7 Let $V_1, V_2, \dots, V_m$ be vector spaces. A vector space $\mathbb{T}$ and an m -linear function $\Phi: V_1 \times V_2 \times \cdots \times V_m \rightarrow \mathbb{T}$ are said to satisfy the **Universal Factorization Property** if, for every vector space W , and every m -linear function $f: V_1 \times V_2 \times \cdots \times V_m \rightarrow W$, there exists a linear function $h: \mathbb{T} \rightarrow W$ such that $f = h\Phi$.

If the pair $(\mathbb{T}, \Phi)$ satisfies the Universal Factorization Property for $V_1, V_2, \dots, V_m$, then any m -linear function of $V_1 \times V_2 \times \cdots \times V_m$ can be “factored” as the composition of a linear transformation h (that depends on f and W), and a fixed m -linear function Φ (that depends on neither f nor W). The existence of such a universal pair effectively reduces the study of m -linear functions to the study of the single m -linear function, Φ , and the vector space $\mathbb{T}$. Definition 5.7 is perhaps best illustrated by means of the commutative diagram in Figure 5.1.

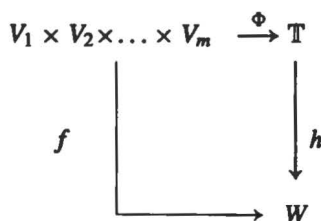


FIGURE 5.1

The existence of universal pairs is easily settled by Theorem 5.4: Let $\{e_{ij} : 1 \leq j \leq n_i\}$ be a basis of V_i , $1 \leq i \leq m$. Consider the set

$$E = \{[e_{1j_1}, e_{2j_2}, \dots, e_{mj_m}] : 1 \leq j_i \leq n_i, 1 \leq i \leq m\}. \quad (5.3)$$

Let T be the vector space consisting of all formal linear combinations of elements of E , that is, the **free vector space** generated by E . By definition, E is a basis of T . Observe that T is not a subspace of $V_1 \times V_2 \times \dots \times V_m$; the addition is different. Whereas the dimension of $V_1 \times V_2 \times \dots \times V_m$ is $n_1 + n_2 + \dots + n_m$,

$$\dim(T) = \prod_{i=1}^m n_i. \quad (5.4)$$

(This explains the use of square brackets in Equation (5.3); while the elements of E are m -tuples of vectors, E is not a subset of the Cartesian product space.)

Define $\Phi: V_1 \times V_2 \times \dots \times V_m \rightarrow T$ by

$$\Phi(e_{1j_1}, e_{2j_2}, \dots, e_{mj_m}) = [e_{1j_1}, e_{2j_2}, \dots, e_{mj_m}], \quad (5.5)$$

$1 \leq j_i \leq n_i$, $1 \leq i \leq m$, and multilinear extension. If W is an arbitrary vector space, and $f: V_1 \times V_2 \times \dots \times V_m \rightarrow W$ is m -linear, let $h \in L(T, W)$ be the unique linear transformation defined by

$$h([e_{1j_1}, e_{2j_2}, \dots, e_{mj_m}]) = f(e_{1j_1}, e_{2j_2}, \dots, e_{mj_m}), \quad (5.6)$$

$1 \leq j_i \leq n_i$, $1 \leq i \leq m$, and linear extension. By Example 5.3(vi), $h\Phi$ is m -linear and, by the uniqueness part of Theorem 5.4, $h\Phi = f$. Thus, the pair (T, Φ) satisfies the Universal Factorization Property for $V_1, V_2, \dots, V_m$. Moreover, (by Equation (5.5)) the reach of Φ is all of T .

DEFINITION 5.8 Suppose (S, Ψ) and (T, Φ) satisfy the Universal Factorization Property for $V_1, V_2, \dots, V_m$. The pairs are **isomorphic** if there exists an invertible linear transformation T from S onto T such that $\Phi = T\Psi$.

THEOREM 5.9 Suppose (S, Ψ) and (T, Φ) satisfy the Universal Factorization Property for $V_1, V_2, \dots, V_m$. If the reach of Ψ (respectively Φ) is all of S (respectively T), then (S, Ψ) and (T, Φ) are isomorphic.

Proof Because (S, Ψ) and (T, Φ) satisfy the Universal Factorization Property, there exist $k \in L(S, T)$ and $h \in L(T, S)$ such that $\Psi = h\Phi$ and $\Phi = k\Psi$. (See Figure 5.2.) Therefore, $\Psi = hk\Psi$. In particular (because the reach of Ψ is S), hk is the identity on S . Similarly, kh is the identity on T . It follows that $k = h^{-1}$, and the proof is complete. $\square$

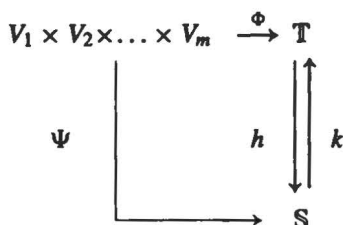


FIGURE 5.2

Let $V_1, V_2, \dots, V_m$ be vector spaces. It follows from Theorem 5.9 that, up to isomorphism, there is a unique pair (T, Φ) that satisfies the Universal Factorization Property where the reach of Φ is all of T .

DEFINITION 5.10 Let T be a vector space and $\Phi: V_1 \times V_2 \times \dots \times V_m \rightarrow T$ a multilinear map whose reach is all of T . If (T, Φ) satisfies the Universal Factorization Property for $V_1, V_2, \dots, V_m$ then T is the **tensor product** of $V_1, V_2, \dots, V_m$, written

$$T = V_1 \otimes V_2 \otimes \dots \otimes V_m, \quad (5.7)$$

and $\Phi(v_1, v_2, \dots, v_m)$ is a **decomposable tensor**.

Let V_i be a vector space of dimension n_i , $1 \leq i \leq m$. Suppose U is an arbitrary vector space of dimension $n_1 \times n_2 \times \dots \times n_m$. Let T be a fixed but arbitrary invertible linear transformations from $V_1 \otimes V_2 \otimes \dots \otimes V_m$ onto U , and define $\Psi = T\Phi$. If W is a vector space, and $f: V_1 \times V_2 \times \dots \times V_m \rightarrow W$ an m -linear function, then there exists a linear transformation $k: U \rightarrow W$ such that $f = k\Psi$,

namely, $k = hT^{-1}$, where $h\Phi = f$. (In other words, the diagram in Figure 5.3 is commutative.) In particular, (U, Ψ) satisfies the Universal Factorization Property. Because U is the reach of Ψ , U is a model for the tensor product. Evidently, $V_1 \otimes V_2 \otimes \cdots \otimes V_m$ is just a vector space of the right dimension. What makes it special are the decomposable tensors, an additional structure afforded by the multilinear map.

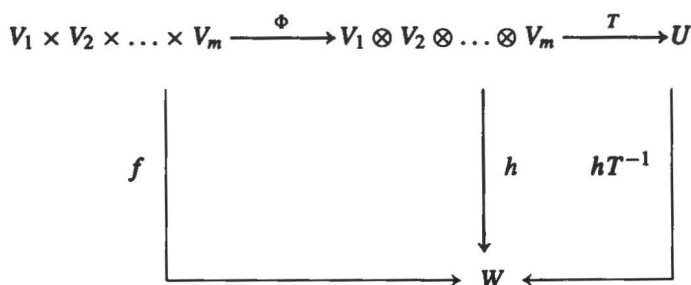


FIGURE 5.3

This approach affords a rapid, direct proof that the tensor product of vector spaces exists and that, up to isomorphism, it is unique. Unfortunately, this way of looking at things produces a somewhat sterile abstraction. Let us construct a more concrete model.

DEFINITION 5.11 Suppose $V_1, V_2, \dots, V_m$ are complex vector spaces. Denote by $M(V_1, V_2, \dots, V_m)$ the set of m -linear functions $f: V_1 \times V_2 \times \cdots \times V_m \rightarrow \mathbb{C}$.

Under the usual "pointwise" addition and scalar multiplication of functions, namely,

$$\begin{aligned} (cf + dg)(v_1, v_2, \dots, v_m) \\ = cf(v_1, v_2, \dots, v_m) + dg(v_1, v_2, \dots, v_m), \end{aligned} \quad (5.8)$$

$M = M(V_1, V_2, \dots, V_m)$ is a vector space. We are interested in certain distinguished elements of its dual space, M' .

Let $v_i \in V_i$, $1 \leq i \leq m$. In anticipation of the next few results, denote by $v_1 \otimes v_2 \otimes \cdots \otimes v_m$ the mapping from M into $\mathbb{C}$ defined by

$$(v_1 \otimes v_2 \otimes \cdots \otimes v_m)(f) = f(v_1, v_2, \dots, v_m). \quad (5.9)$$

(Compare with Equation (2.3).) Because

$$\begin{aligned}
 & (v_1 \otimes v_2 \otimes \cdots \otimes v_m)(cf + dg) \\
 &= (cf + dg)(v_1, v_2, \dots, v_m) \\
 &= cf(v_1, v_2, \dots, v_m) + dg(v_1, v_2, \dots, v_m) \\
 &= c(v_1 \otimes v_2 \otimes \cdots \otimes v_m)(f) + d(v_1 \otimes v_2 \otimes \cdots \otimes v_m)(g),
 \end{aligned}$$

$v_1 \otimes v_2 \otimes \cdots \otimes v_m$ is linear. It turns out that these special functionals span M' . In fact, more is true.

THEOREM 5.12 Let $\mathcal{B}_i = \{e_{ij} : 1 \leq j \leq n_i\}$ be a basis of V_i , $1 \leq i \leq m$. Then

$$\mathcal{B} = \{e_{1j_1} \otimes e_{2j_2} \otimes \cdots \otimes e_{mj_m} : 1 \leq j_i \leq n_i, 1 \leq i \leq m\} \quad (5.10)$$

is a basis of the dual space of $M(V_1, V_2, \dots, V_m)$.

Proof Let $\{f_{ij} : 1 \leq j \leq n_i\}$ be the basis of $L(V_i, \mathbb{C})$ dual to $\mathcal{B}_i$. Then f_{ij} is defined by $f_{ij}(e_{ik}) = \delta_{j,k}$, and linear extension. Let

$$\prod_{t=1}^m f_{tj_t} \in M(V_1, V_2, \dots, V_m)$$

be the m -linear functional defined by

$$\left(\prod_{t=1}^m f_{tj_t} \right) (v_1, v_2, \dots, v_m) = \prod_{t=1}^m f_{tj_t}(v_t).$$

We claim that

$$\left\{ \prod_{t=1}^m f_{tj_t} : 1 \leq j_t \leq n_t, 1 \leq t \leq m \right\} \quad (5.11)$$

is a basis of $M = M(V_1, V_2, \dots, V_m)$. If $g \in M$ then

$$g = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \cdots \sum_{i_m=1}^{n_m} g(e_{1i_1}, e_{2i_2}, \dots, e_{mi_m}) \prod_{t=1}^m f_{ti_t}. \quad (5.12)$$

(Compare with Equation (2.2).) To verify this identity, observe that its right-hand side is a linear combination of m -linear functions and hence is m -linear. Therefore (Theorem 5.4), it suffices to evaluate both sides on $(e_{1j_1}, e_{2j_2}, \dots, e_{mj_m})$:

$$\begin{aligned}
 & g(e_{1j_1}, e_{2j_2}, \dots, e_{mj_m}) \\
 &= \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \cdots \sum_{i_m=1}^{n_m} g(e_{1i_1}, e_{2i_2}, \dots, e_{mi_m}) \prod_{t=1}^m f_{ti_t}(e_{tj_t}).
 \end{aligned}$$

Because $\prod f_{it_i}(e_{t_i}) = 0$ unless $i_t = j_t$, $1 \leq t \leq m$, the right-hand side of this equation collapses to the left-hand side. Because (5.10) is the dual basis of (5.11), the proof is complete. $\square$

Set $\mathbb{T} = M'$, the dual space of $M = M(V_1, V_2, \dots, V_m)$, and define $\Phi : V_1 \times V_2 \times \dots \times V_m \rightarrow \mathbb{T}$ by $\Phi(v_1, v_2, \dots, v_m) = v_1 \otimes v_2 \otimes \dots \otimes v_m$. If $f \in M$, then

$$\begin{aligned}
 & \begin{array}{c} i\text{-th position} \\ \downarrow \end{array} \\
 & (v_1 \otimes v_2 \otimes \dots \otimes [cu + dw] \otimes \dots \otimes v_m)(f) \\
 & = f(v_1, v_2, \dots, [cu + dw], \dots, v_m) \\
 & = cf(v_1, v_2, \dots, u, \dots, v_m) + df(v_1, v_2, \dots, w, \dots, v_m) \\
 & = c(v_1 \otimes v_2 \otimes \dots \otimes u \otimes \dots \otimes v_m)(f) \\
 & \quad + d(v_1 \otimes v_2 \otimes \dots \otimes w \otimes \dots \otimes v_m)(f),
 \end{aligned}$$

by the multilinearity of f . Because f was arbitrary,

$$\begin{aligned}
 & (v_1 \otimes v_2 \otimes \dots \otimes [cu + dw] \otimes \dots \otimes v_m) \\
 & = c(v_1 \otimes v_2 \otimes \dots \otimes u \otimes \dots \otimes v_m) \\
 & \quad + d(v_1 \otimes v_2 \otimes \dots \otimes w \otimes \dots \otimes v_m).
 \end{aligned}$$

Therefore, Φ is m -linear. To prove that $(\mathbb{T}, \Phi)$ satisfies the Universal Factorization Property, let W be a vector space. Suppose that $g : V_1 \times V_2 \times \dots \times V_m \rightarrow W$ is m -linear. Let $\mathcal{B}$ be the basis of $\mathbb{T}$ given in Theorem 5.12 and define $h : \mathbb{T} \rightarrow W$ by

$$h(e_{1j_1} \otimes e_{2j_2} \otimes \dots \otimes e_{mj_m}) = g(e_{1j_1}, e_{2j_2}, \dots, e_{mj_m}),$$

$1 \leq j_i \leq n_i$, $1 \leq i \leq m$, and linear extension. Then the m -linear functions $h\Phi$ and g agree on

$$\{(e_{1j_1}, e_{2j_2}, \dots, e_{mj_m}) : 1 \leq j_i \leq n_i, 1 \leq i \leq m\}.$$

Therefore (Theorem 5.4), $g = h\Phi$, and $(\mathbb{T}, \Phi)$ is a universal pair. Moreover, by Theorem 5.12, the reach of Φ is all of $\mathbb{T}$. In other words, $\mathbb{T} = V_1 \otimes V_2 \otimes \dots \otimes V_m$.

We are now going to adopt the notation, $v_1 \otimes v_2 \otimes \dots \otimes v_m$, from the multilinear functional model, for the abstract definition. The notation

$$[e_{1j_1}, e_{2j_2}, \dots, e_{mj_m}]$$

used in the first construction is to be replaced with

$$e_{1j_1} \otimes e_{2j_2} \otimes \cdots \otimes e_{mj_m}.$$

Because $\Phi(v_1, v_2, \dots, v_m) = v_1 \otimes v_2 \otimes \cdots \otimes v_m$, we no longer have any special need for the symbol Φ . These notational choices are formalized in the following:

DEFINITION 5.13 Henceforth, the generic, abstract, decomposable tensor is denoted $v_1 \otimes v_2 \otimes \cdots \otimes v_m$ with the understanding that it may be viewed as a linear functional on $M(V_1, V_2, \dots, V_m)$ whenever it is useful to do so.

It is worth emphasizing that

$$V_1 \otimes V_2 \otimes \cdots \otimes V_m \neq \{v_1 \otimes v_2 \otimes \cdots \otimes v_m : v_i \in V_i, 1 \leq i \leq m\}.$$

THEOREM 5.14 Let $\{e_{ij} : 1 \leq j \leq n_i\}$ be a basis of V_i , $1 \leq i \leq m$. If

$$v_i = \sum_{j=1}^{n_i} a_{ij} e_{ij}, \quad 1 \leq i \leq m,$$

then

$$\begin{aligned} & v_1 \otimes v_2 \otimes \cdots \otimes v_m \\ &= \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} \cdots \sum_{j_m=1}^{n_m} \left(\prod_{i=1}^m a_{ij_i} \right) e_{1j_1} \otimes e_{2j_2} \otimes \cdots \otimes e_{mj_m}. \end{aligned}$$

THEOREM 5.15 Let $v_i \in V_i$, $1 \leq i \leq m$. Then $v_1 \otimes v_2 \otimes \cdots \otimes v_m = 0$ if and only if $v_i = 0$ for some i .

Proof If $v_i \neq 0$ for all i , there exist $f_i \in L(V_i, \mathbb{C})$ such that $f_i(v_i) = 1$, $1 \leq i \leq m$. Let $f = \prod f_i \in M(V_1, V_2, \dots, V_m)$. Then

$$\begin{aligned} (v_1 \otimes v_2 \otimes \cdots \otimes v_m)(f) &= f(v_1, v_2, \dots, v_m) \\ &= \prod_{i=1}^m f_i(v_i) \\ &= 1, \end{aligned}$$

so $v_1 \otimes v_2 \otimes \cdots \otimes v_m$ is not the zero functional. The converse is left to the exercises. $\square$

THEOREM 5.16 Suppose $v_i, w_i \in V_i$ where $w_i \neq 0$, $1 \leq i \leq m$. Then

$$v_1 \otimes v_2 \otimes \cdots \otimes v_m = w_1 \otimes w_2 \otimes \cdots \otimes w_m$$

if and only if there exist m complex numbers $c_1, c_2, \dots, c_m$ such that $v_i = c_i w_i$, $1 \leq i \leq m$, and $c_1 \times c_2 \times \cdots \times c_m = 1$.

Proof Suppose $v_1 \otimes v_2 \otimes \cdots \otimes v_m = w_1 \otimes w_2 \otimes \cdots \otimes w_m \neq 0$. If v_k is not a multiple of w_k , for some fixed but arbitrary k , then $\{v_k, w_k\}$ is linearly independent, in which case there exists a linear functional $f_k \in L(V_k, \mathbb{C})$ such that $f_k(v_k) = 0$ and $f_k(w_k) = 1$. Choose $f_i \in L(V_i, \mathbb{C})$, $i \neq k$, such that $f_i(w_i) = 1$ and set $f = \prod f_i$. Then

$$\begin{aligned} (v_1 \otimes v_2 \otimes \cdots \otimes v_m)(f) &= \prod_{i=1}^m f_i(v_i) \\ &= 0 \\ &\neq 1 \\ &= \prod_{i=1}^m f_i(w_i) = (w_1 \otimes w_2 \otimes \cdots \otimes w_m)(f), \end{aligned}$$

contradicting the hypothesis. Therefore, there exist $c_i \in \mathbb{C}$ such that $v_i = c_i w_i$, $1 \leq i \leq m$. By multilinearity,

$$\begin{aligned} w_1 \otimes w_2 \otimes \cdots \otimes w_m &= v_1 \otimes v_2 \otimes \cdots \otimes v_m \\ &= c_1 w_1 \otimes c_2 w_2 \otimes \cdots \otimes c_m w_m \\ &= \left(\prod_{i=1}^m c_i \right) w_1 \otimes w_2 \otimes \cdots \otimes w_m. \end{aligned}$$

Because $w_1 \otimes w_2 \otimes \cdots \otimes w_m \neq 0$, $c_1 \times c_2 \times \cdots \times c_m = 1$. The proof of sufficiency is left to the exercises. $\square$

We now consider yet another model for the tensor product space.

THEOREM 5.17 Suppose $1 \leq k < m$. Then $V_1 \otimes V_2 \otimes \cdots \otimes V_m$ is a model for $(V_1 \otimes V_2 \otimes \cdots \otimes V_k) \otimes (V_{k+1} \otimes V_{k+2} \otimes \cdots \otimes V_m)$ in which

$$\begin{aligned} (u_1 \otimes u_2 \otimes \cdots \otimes u_k) \otimes (w_1 \otimes w_2 \otimes \cdots \otimes w_{m-k}) \\ = u_1 \otimes \cdots \otimes u_k \otimes w_1 \otimes \cdots \otimes w_{m-k}. \end{aligned}$$

Proof We need to show the existence of a linear transformation

$$T : ((V_1 \otimes V_2 \otimes \cdots \otimes V_k) \otimes (V_{k+1} \otimes V_{k+2} \otimes \cdots \otimes V_m)) \rightarrow V_1 \otimes V_2 \otimes \cdots \otimes V_m$$

such that

$$T((v_1 \otimes \cdots \otimes v_k) \otimes (v_{k+1} \otimes \cdots \otimes v_m)) = v_1 \otimes v_2 \otimes \cdots \otimes v_m,$$

for all $v_i \in V_i$, $1 \leq i \leq m$. Because $\{v_1 \otimes v_2 \otimes \cdots \otimes v_m : v_i \in V_i\}$ spans $V_1 \otimes V_2 \otimes \cdots \otimes V_m$, any such transformation is onto and hence, by a dimension argument, one-to-one.

Let Ψ be the unique bilinear function¹ that satisfies

$$\Psi(v_1 \otimes \cdots \otimes v_k, v_{k+1} \otimes \cdots \otimes v_m) = v_1 \otimes v_2 \otimes \cdots \otimes v_m,$$

for all $v_i \in V_i$, $1 \leq i \leq m$. Then the existence of T is established by the Universal Factorization Property (illustrated in Figure 5.4). $\square$

$$(V_1 \otimes \cdots \otimes V_k) \times (V_{k+1} \otimes \cdots \otimes V_m) \rightarrow (V_1 \otimes \cdots \otimes V_k) \otimes (V_{k+1} \otimes \cdots \otimes V_m)$$

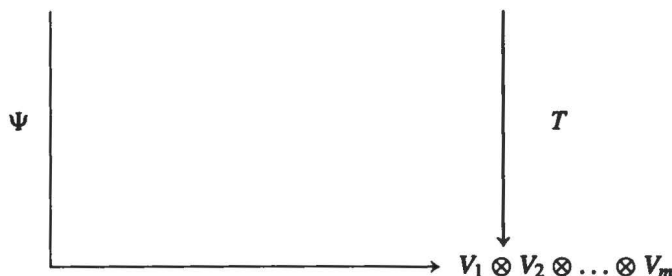


FIGURE 5.4

If $m = 3$ and $k = 2$, Theorem 5.17 becomes $(V_1 \otimes V_2) \otimes V_3 \cong V_1 \otimes V_2 \otimes V_3$. Evidently,

$$(V_1 \otimes V_2) \otimes V_3 \cong V_1 \otimes (V_2 \otimes V_3). \quad (5.13)$$

When $V_1 = V_2 = \cdots = V_m$, Theorem 5.17 is the basis for the theory of tensor algebras, a fascinating, but well treated subject about which we will say no more. (See [Bourbaki (1948)], [Greub (1967)] or [Marcus (1973)].)

¹ See Exercise 10.

EXAMPLE 5.18 Suppose $V_1 = \mathbf{C}_{1,n}$ and $V_2 = \mathbf{C}_{1,k}$. Let $\Phi: V_1 \times V_2 \rightarrow \mathbf{C}_{n,k}$ be the bilinear function defined by $\Phi(X, Y) = X'Y$. Let E^j be the 1-by- k matrix whose only nonzero entry is a 1 in column j , and E_i the 1-by- n matrix whose only nonzero entry is a 1 in column i . Then $\Phi(E_i, E^j) = E_{ij}$, the n -by- k matrix whose only nonzero entry is a 1 in position (i, j) . (Compare with Example 5.5.) Because $\{E_{ij}: 1 \leq i \leq n, 1 \leq j \leq k\}$ is a basis of $\mathbf{C}_{n,k}$, the reach of Φ is all of $\mathbf{C}_{n,k}$. If W is a vector space and $f: V_1 \times V_2 \rightarrow W$ is bilinear, define $h: \mathbf{C}_{n,k} \rightarrow W$ by $h(E_{ij}) = f(E_i, E^j)$, $1 \leq i \leq n, 1 \leq j \leq k$, and linear extension. From Theorem 5.4, $f = h\Phi$. Therefore, $\mathbf{C}_{n,k}$ is yet another model for $V_1 \otimes V_2$, one in which the decomposable tensor $X \otimes Y = X'Y$. (Because $(V_1 \otimes V_2) \otimes V_3 \cong V_1 \otimes V_2 \otimes V_3$, this model could be used for an inductive construction of $V_1 \otimes V_2 \otimes \cdots \otimes V_m$.) $\square$

Now that the study of multilinear functions has been reduced to the study of tensor products (including the decomposable tensors), we are going to shift our perspective and think of multilinear functions as a means of illuminating tensor products. Put another way, we are not so much interested in multilinear functions for their own sake as we are in the light they shed on tensor spaces.

Let $V_1, V_2, \dots, V_m$ and $W_1, W_2, \dots, W_m$ be vector spaces, and suppose $T_i \in L(V_i, W_i)$, $1 \leq i \leq m$. Then the function $\Psi: V_1 \times V_2 \times \cdots \times V_m \rightarrow W_1 \otimes W_2 \otimes \cdots \otimes W_m$ defined by

$$\Psi(v_1, v_2, \dots, v_m) = T_1(v_1) \otimes T_2(v_2) \otimes \cdots \otimes T_m(v_m)$$

is m -linear. (Observe that $v_1 \times v_2 \times \cdots \times v_m \rightarrow v_1 \otimes v_2 \otimes \cdots \otimes v_m$ is m -linear and apply Example 5.3(v).) Thus, there is a unique linear transformation $h: V_1 \otimes V_2 \otimes \cdots \otimes V_m \rightarrow W_1 \otimes W_2 \otimes \cdots \otimes W_m$ such that

$$h(v_1 \otimes v_2 \otimes \cdots \otimes v_m) = T_1(v_1) \otimes T_2(v_2) \otimes \cdots \otimes T_m(v_m), \quad (5.14)$$

for all $v_i \in V_i$, $1 \leq i \leq m$.

DEFINITION 5.19 Let $T_i \in L(V_i, W_i)$, $1 \leq i \leq m$. The unique $h \in L(V_1 \otimes V_2 \otimes \cdots \otimes V_m, W_1 \otimes W_2 \otimes \cdots \otimes W_m)$ determined by Equation (5.14) is said to be **induced** by $T_1, T_2, \dots, T_m$. The notation for this induced linear transformation is $h = T_1 \otimes T_2 \otimes \cdots \otimes T_m$.

If I_k is the identity operator on V_k , $1 \leq k \leq m$, then it follows from Equation (5.14), and the fact that the decomposable tensors span, that $I_1 \otimes I_2 \otimes \cdots \otimes I_m$ is the identity operator on $V_1 \otimes V_2 \otimes \cdots \otimes V_m$.

THEOREM 5.20 If $S_i \in L(U_i, V_i)$ and $T_i \in L(V_i, W_i)$, $1 \leq i \leq m$, then

$$(T_1 \otimes T_2 \otimes \cdots \otimes T_m)(S_1 \otimes S_2 \otimes \cdots \otimes S_m) = (T_1 S_1) \otimes (T_2 S_2) \otimes \cdots \otimes (T_m S_m). \quad (5.15)$$

COROLLARY 5.21 *Suppose $T_i \in L(V_i, W_i)$, $1 \leq i \leq m$. Then $T_1 \otimes T_2 \otimes \cdots \otimes T_m$ is invertible if and only if T_i is invertible, $1 \leq i \leq m$.*

Proof If each T_i is invertible then, letting $U_i = W_i$ and $S_i = T_i^{-1}$ in Equation (5.15) establishes the identity

$$(T_1^{-1}) \otimes (T_2^{-1}) \otimes \cdots \otimes (T_m^{-1}) = (T_1 \otimes T_2 \otimes \cdots \otimes T_m)^{-1}. \quad (5.16)$$

Conversely, suppose

$$T_1 \otimes T_2 \otimes \cdots \otimes T_m \in L(V_1 \otimes V_2 \otimes \cdots \otimes V_m, W_1 \otimes W_2 \otimes \cdots \otimes W_m)$$

is one-to-one and onto. If $v_i \neq 0$, $1 \leq i \leq m$, then, by Theorem 5.15,

$$\begin{aligned} & T_1(v_1) \otimes T_2(v_2) \otimes \cdots \otimes T_m(v_m) \\ &= (T_1 \otimes T_2 \otimes \cdots \otimes T_m)(v_1 \otimes v_2 \otimes \cdots \otimes v_m) \\ &\neq 0. \end{aligned}$$

Therefore, $T_i(v_i) \neq 0$, $1 \leq i \leq m$, and each T_i is nonsingular. It follows from a dimension argument that each T_i is onto. $\square$

By now the reader has probably observed a notational ambiguity. How do we distinguish

$$T_1 \otimes T_2 \otimes \cdots \otimes T_m \in L(V_1 \otimes V_2 \otimes \cdots \otimes V_m, W_1 \otimes W_2 \otimes \cdots \otimes W_m)$$

from

$$T_1 \otimes T_2 \otimes \cdots \otimes T_m \in L(V_1, W_1) \otimes L(V_2, W_2) \otimes \cdots \otimes L(V_m, W_m)?$$

THEOREM 5.22 *The vector space*

$$L(V_1 \otimes V_2 \otimes \cdots \otimes V_m, W_1 \otimes W_2 \otimes \cdots \otimes W_m)$$

is a model for the tensor product

$$L(V_1, W_1) \otimes L(V_2, W_2) \otimes \cdots \otimes L(V_m, W_m)$$

in which $\{T_1 \otimes T_2 \otimes \cdots \otimes T_m : T_i \in L(V_i, W_i)\}$ is the set of decomposable tensors.

$$L_1 \times L_2 \times \dots \times L_m \rightarrow L_1 \otimes L_2 \otimes \dots \otimes L_m$$



FIGURE 5.5

Proof Let $T_1 \otimes T_2 \otimes \dots \otimes T_m$ be the induced transformation and, temporarily, let $T_1 \odot T_2 \odot \dots \odot T_m$ denote the decomposable tensor in $L(V_1, W_1) \otimes L(V_2, W_2) \otimes \dots \otimes L(V_m, W_m)$. Consider the diagram illustrated in Figure 5.5, where $L_i = L(V_i, W_i)$, $1 \leq i \leq m$,

$$\mathbb{L} = L(V_1 \otimes V_2 \otimes \dots \otimes V_m, W_1 \otimes W_2 \otimes \dots \otimes W_m),$$

and $\Psi(T_1, T_2, \dots, T_m) = T_1 \otimes T_2 \otimes \dots \otimes T_m$. Because we cannot assume that $T_1 \otimes T_2 \otimes \dots \otimes T_m$ is a decomposable tensor, we are obliged to prove that Ψ is multilinear. That, however, is an easy consequence of the identity

$$\begin{aligned} (T_1 \otimes \dots \otimes (cT_i + dT'_i) \otimes \dots \otimes T_m)(v_1 \otimes v_2 \otimes \dots \otimes v_m) \\ = T_1(v_1) \otimes \dots \otimes (cT_i(v_i) + dT'_i(v_i)) \otimes \dots \otimes T_m(v_m), \end{aligned}$$

because the tensor product of vectors on the right-hand side of this equation is multilinear.

Now that we know Ψ to be multilinear, we can use the Universal Factorization Property to deduce the existence of a (unique) linear transformation

$$h: L_1 \otimes L_2 \otimes \dots \otimes L_m \rightarrow \mathbb{L}$$

such that $h(T_1 \odot T_2 \odot \dots \odot T_m) = T_1 \otimes T_2 \otimes \dots \otimes T_m$. It remains to prove that h is invertible. Suppose $n_i = \dim(V_i)$ and $k_i = \dim(W_i)$, $1 \leq i \leq m$. Then $\dim(L_i) = n_i k_i$, so $\dim(L_1 \otimes L_2 \otimes \dots \otimes L_m) = \prod n_i k_i$. Because $\dim(\mathbb{L}) = (\prod n_i) (\prod k_i)$ is the same, it suffices to show that h is onto.

Suppose $\{v_{ij}: 1 \leq i \leq n_i\}$ and $\{w_{ij}: 1 \leq i \leq k_i\}$ are bases of V_i and W_i , respectively, $1 \leq i \leq m$. If S is a fixed but arbitrary linear transformation in $\mathbb{L}$

then, by Theorem 5.12, there exist complex coefficients such that

$$\begin{aligned} S(v_{1j_1} \otimes v_{2j_2} \otimes \cdots \otimes v_{mj_m}) \\ = \sum_{(i)} c_{(i_1, \dots, i_m), (j_1, \dots, j_m)} w_{1i_1} \otimes w_{2i_2} \otimes \cdots \otimes w_{mi_m}, \end{aligned}$$

$1 \leq j_i \leq n_i$, $1 \leq i \leq m$, where

$$\sum_{(i)} = \sum_{i_1=1}^{k_1} \sum_{i_2=1}^{k_2} \cdots \sum_{i_m=1}^{k_m}.$$

Define $T_{jr}^i: V_i \rightarrow W_i$ by $T_{jr}^i(v_{it}) = \delta_{j,t} w_{ir}$, $1 \leq j, t \leq n_i$, $1 \leq r \leq k_i$, $1 \leq i \leq m$, and linear extension. Then

$$S = \sum_{(i)} \sum_{(t)} c_{(i_1, \dots, i_m), (t_1, \dots, t_m)} T_{t_1 i_1}^1 \otimes T_{t_2 i_2}^2 \otimes \cdots \otimes T_{t_m i_m}^m, \quad (5.17)$$

where

$$\sum_{(t)} = \sum_{t_1=1}^{n_1} \sum_{t_2=1}^{n_2} \cdots \sum_{t_m=1}^{n_m}.$$

To confirm Equation (5.17), apply both sides to $v_{1j_1} \otimes v_{2j_2} \otimes \cdots \otimes v_{mj_m}$. Because it is a linear combination of things of the form $T_1 \otimes T_2 \otimes \cdots \otimes T_m$, S belongs to the image of h . Therefore, h is onto. $\square$

We turn now to matrix representations of induced linear transformations.

DEFINITION 5.23 Let $\mathcal{B}_i = \{e_{ij}: 1 \leq j \leq n_i\}$ be an ordered basis of V_i , $1 \leq i \leq m$. The basis

$$\mathcal{B} = \{e_{1j_1} \otimes e_{2j_2} \otimes \cdots \otimes e_{mj_m}: 1 \leq j_i \leq n_i, 1 \leq i \leq m\} \quad (5.18)$$

of $V_1 \otimes V_2 \otimes \cdots \otimes V_m$ is said to be **induced** by $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_m$. The induced basis $\mathcal{B}$ is ordered lexicographically by the subscripts. That is,

$$e_{1i_1} \otimes e_{2i_2} \otimes \cdots \otimes e_{mi_m} \text{ precedes } e_{1j_1} \otimes e_{2j_2} \otimes \cdots \otimes e_{mj_m}$$

provided the first nonzero difference, $j_t - i_t$, is positive.

EXAMPLE 5.24 $e_{1r} \otimes e_{2s} \otimes e_{3t}$ comes before $e_{1i} \otimes e_{2j} \otimes e_{3k}$ in lexicographic order if

- (1) $r < i$; or if
- (2) $r = i$ and $s < j$; or if
- (3) $r = i$, $s = j$, and $t < k$.

$\square$

THEOREM 5.25 Let $\{v_{ij} : 1 \leq j \leq n_i\}$ be an ordered basis of V_i , and $\{w_{ij} : 1 \leq j \leq k_i\}$ be an ordered basis of W_i , $1 \leq i \leq m$. Let

$$E = \{v_{1j_1} \otimes v_{2j_2} \otimes \cdots \otimes v_{mj_m} : 1 \leq j_i \leq n_i, 1 \leq i \leq m\}$$

and

$$F = \{w_{1j_1} \otimes w_{2j_2} \otimes \cdots \otimes w_{mj_m} : 1 \leq j_i \leq k_i, 1 \leq i \leq m\}$$

be the lexicographically ordered, induced bases of $V_1 \otimes V_2 \otimes \cdots \otimes V_m$ and $W_1 \otimes W_2 \otimes \cdots \otimes W_m$, respectively. Let $T_p \in L(V_p, W_p)$ be defined by

$$T_p(v_{pj}) = \sum_{i=1}^{k_p} a_{ij}^p w_{pi}, \quad 1 \leq j \leq n_p,$$

and linear extension (so the matrix representation of T_p , with respect to $\{v_{pr} : 1 \leq r \leq n_p\}$ and $\{w_{pr} : 1 \leq r \leq k_p\}$, is $A_p = (a_{ij}^p)$.) Then, the

$$((i_1, i_2, \dots, i_m), (j_1, j_2, \dots, j_m))\text{-entry}$$

of the matrix representation of $T_1 \otimes T_2 \otimes \cdots \otimes T_m$ with respect to E and F is

$$\prod_{p=1}^m a_{i_p, j_p}^p. \quad (5.19)$$

Proof

$$\begin{aligned} & (T_1 \otimes T_2 \otimes \cdots \otimes T_m)(v_{1j_1} \otimes v_{2j_2} \otimes \cdots \otimes v_{mj_m}) \\ &= T_1(v_{1j_1}) \otimes T_2(v_{2j_2}) \otimes \cdots \otimes T_m(v_{mj_m}) \\ &= \sum_{(i)} a_{i_1 j_1}^1 a_{i_2 j_2}^2 \cdots a_{i_m j_m}^m w_{1i_1} \otimes w_{2i_2} \otimes \cdots \otimes w_{mi_m}. \end{aligned}$$

□

DEFINITION 5.26 Let $A_p = (a_{ij}^p)$ be a k_p -by- n_p matrix, $1 \leq p \leq m$. The **Kronecker product**, $A_1 \otimes A_2 \otimes \cdots \otimes A_m$, is a $\prod k_p$ -by- $\prod n_p$ matrix whose rows are indexed by the set $\{(i_1, i_2, \dots, i_m) : 1 \leq i_p \leq k_p\}$ and whose columns are indexed by $\{(j_1, j_2, \dots, j_m) : 1 \leq j_p \leq n_p\}$, both ordered lexicographically. The $((i_1, i_2, \dots, i_m), (j_1, j_2, \dots, j_m))$ -entry of this big matrix is

$$\prod_{p=1}^m a_{i_p, j_p}^p.$$

The isomorphism between $L(V_i, W_i)$ and C_{k_i, n_i} , together with Theorem 5.22, yields the expected isomorphism

$$C_{\prod k_i, \prod n_i} \cong C_{k_1, n_1} \otimes C_{k_2, n_2} \otimes \cdots \otimes C_{k_m, n_m}. \quad (5.20)$$

COROLLARY 5.27 *If the matrix representation of T_i is A_i then, with respect to appropriate induced bases, the matrix representation of $T_1 \otimes T_2 \otimes \cdots \otimes T_m$ is $A_1 \otimes A_2 \otimes \cdots \otimes A_m$.*

Proof In view of Definition 5.26, this is just a restatement of Theorem 5.25. $\square$

EXAMPLE 5.28 Let $A_1 = A = (a_{ij}) \in \mathbb{C}_{p,q}$ and $A_2 = B = (b_{rs}) \in \mathbb{C}_{m,n}$. By Definition 5.26, the $((i, r), (j, s))$ -entry of $A \otimes B \in \mathbb{C}_{pm, qn}$ is $a_{ij}b_{rs}$. In lexicographic order, (i_1, r_1) comes before (i_2, r_2) if $i_1 < i_2$ or if $i_1 = i_2$ and $r_1 < r_2$. Thus, the first m rows of $A \otimes B$ are the ones indexed $(1, 1), (1, 2), \dots, (1, m)$. Similarly, the first n columns are those indexed $(1, 1), (1, 2), \dots, (1, n)$. Denote by L the submatrix of $A \otimes B$ lying in its first m rows and first n columns. Then the (r, s) -entry of L is the $((1, r), (1, s))$ -entry of $A \otimes B$, namely, $a_{11}b_{rs}$. Evidently, $L = a_{11}B$. What about M , the submatrix of $A \otimes B$ lying in its first m rows, and columns $n + 1$ through $2n$? The (r, s) -entry of M is the $((1, r), (2, s))$ -entry of $A \otimes B$, namely, $a_{12}b_{rs}$. That is, $M = a_{12}B$. More generally, $A \otimes B$ is the block partitioned matrix

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1q}B \\ a_{21}B & a_{22}B & \cdots & a_{2q}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1}B & a_{p2}B & \cdots & a_{pq}B \end{pmatrix}. \quad (5.21)$$

EXAMPLE 5.29 Let $m = p$ and $n = q$ in Example 5.28, so that $A, B \in \mathbb{C}_{p,q}$. The **Hadamard (or Schur) product** of A and B is $A \cdot B = (a_{ij}b_{ij})$, that is, the p -by- q matrix whose (i, j) -entry is $a_{ij}b_{ij}$. Observe that $A \cdot B$ is the principal submatrix of $A \otimes B$ lying in rows $(1, 1), (2, 2), \dots, (p, p)$ and columns $(1, 1), (2, 2), \dots, (q, q)$. $\square$

EXAMPLE 5.30 Suppose $P = (p_{ij})$ and $Q = (q_{ij})$ are n -by- n complex matrices. Define $T \in L(\mathbb{C}_{n,n}, \mathbb{C}_{n,n})$ by $T(A) = PAQ$. Let's compute the matrix representation of T with respect to the basis $B = \{E_{ij} : 1 \leq i, j \leq n\}$, where E_{ij} is the n -by- n matrix whose only nonzero entry is a 1 in position (i, j) . By definition,

$$\begin{aligned} T(E_{rs}) &= P E_{rs} Q \\ &= \left(\sum_{i=1}^n p_{ir} E_{is} \right) Q \\ &= \sum_{i=1}^n p_{ir} \left(\sum_{j=1}^n q_{sj} E_{ij} \right) \\ &= \sum_{i,j=1}^n (p_{ir} q_{sj}) E_{ij}. \end{aligned}$$

So, the $((i, j), (r, s))$ -entry of the matrix representation of T with respect to $\mathcal{B}$ is $p_{ir}q_{sj}$, precisely the $((i, j), (r, s))$ -entry of $P \otimes Q'$. $\square$

Let $\mathcal{B}_i = \{e_{ij} : 1 \leq j \leq n_i\}$ be an orthonormal basis of the inner product space V_i , $1 \leq i \leq m$. Then there exists a unique inner product on $V_1 \otimes V_2 \otimes \cdots \otimes V_m$ with respect to which the induced basis

$$\mathcal{B} = \{e_{1j_1} \otimes e_{2j_2} \otimes \cdots \otimes e_{mj_m} : 1 \leq j_i \leq n_i, 1 \leq i \leq m\}$$

is orthonormal. Namely, if

$$v = \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} \cdots \sum_{j_m=1}^{n_m} a_{(j_1, j_2, \dots, j_m)} e_{1j_1} \otimes e_{2j_2} \otimes \cdots \otimes e_{mj_m}$$

and

$$w = \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} \cdots \sum_{j_m=1}^{n_m} b_{(j_1, j_2, \dots, j_m)} e_{1j_1} \otimes e_{2j_2} \otimes \cdots \otimes e_{mj_m},$$

then their inner product is

$$(u, v) = \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} \cdots \sum_{j_m=1}^{n_m} a_{(j_1, j_2, \dots, j_m)} \bar{b}_{(j_1, j_2, \dots, j_m)}. \quad (5.22)$$

From its derivation, it would seem that the inner product defined in Equation (5.22) depends on the orthonormal bases $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_m$ that were used in its construction. In fact, this turns out not to be the case.

THEOREM 5.31 *Let $(\cdot, \cdot)_i$ be an inner product on V_i , $1 \leq i \leq m$. If $(\cdot, \cdot)$ is the inner product defined by Equation (5.22), then*

$$(v_1 \otimes v_2 \otimes \cdots \otimes v_m, w_1 \otimes w_2 \otimes \cdots \otimes w_m) = \prod_{i=1}^m (v_i, w_i)_i, \quad (5.23)$$

$v_i, w_i \in V_i$, $1 \leq i \leq m$.

Proof Let $\mathcal{B}_i = \{e_{ij} : 1 \leq j \leq n_i\}$ be the orthonormal basis of V_i used in the derivation of Equation (5.22). Let

$$v_i = \sum_{j=1}^{n_i} a_{ij} e_{ij} \quad \text{and} \quad w_i = \sum_{j=1}^{n_i} b_{ij} e_{ij},$$

$1 \leq i \leq m$. Then (Theorem 5.14)

$$\begin{aligned} & v_1 \otimes v_2 \otimes \cdots \otimes v_m \\ &= \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} \cdots \sum_{j_m=1}^{n_m} \left(\prod_{t=1}^m a_{tj_t} \right) e_{1j_1} \otimes e_{2j_2} \otimes \cdots \otimes e_{mj_m}, \end{aligned}$$

and

$$\begin{aligned} & w_1 \otimes w_2 \otimes \cdots \otimes w_m \\ &= \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} \cdots \sum_{j_m=1}^{n_m} \left(\prod_{t=1}^m b_{tj_t} \right) e_{1j_1} \otimes e_{2j_2} \otimes \cdots \otimes e_{mj_m}. \end{aligned}$$

Therefore, by Equation (5.22),

$$\begin{aligned} & (v_1 \otimes v_2 \otimes \cdots \otimes v_m, w_1 \otimes w_2 \otimes \cdots \otimes w_m) \\ &= \sum_{(j)} \left(\prod_{t=1}^m a_{tj_t} \right) \left(\prod_{t=1}^m \bar{b}_{tj_t} \right) \\ &= \sum_{(j)} \left(\prod_{t=1}^m a_{tj_t} \bar{b}_{tj_t} \right) \\ &= \prod_{i=1}^m \left(\sum_{j=1}^{n_i} a_{ij} \bar{b}_{ij} \right) \\ &= \prod_{i=1}^m \left(\sum_{j=1}^{n_i} a_{ij} \bar{b}_{ij} (e_{ij}, e_{ij})_i \right) \\ &= \prod_{i=1}^m \left(\sum_{j=1}^{n_i} a_{ij} e_{ij}, \sum_{k=1}^{n_i} b_{ik} e_{ik} \right)_i \\ &= \prod_{i=1}^m (v_i, w_i)_i. \end{aligned}$$

□

Because the decomposable tensors span $V_1 \otimes V_2 \otimes \cdots \otimes V_m$, the inner product defined by Equation (5.22) is completely and uniquely determined by Equation (5.23). In particular, it is basis independent. Because Equation (5.23) is so much simpler and more appealing than Equation (5.22), we may as well use it in the formal definition.

DEFINITION 5.32 Let $V_1, V_2, \dots, V_m$ be inner product spaces. (As it will cause no confusion, let $(\cdot, \cdot)$, without a subscript, denote the inner product in V_i , $1 \leq i \leq m$.) The unique inner product on $V_1 \otimes V_2 \otimes \dots \otimes V_m$ that satisfies

$$(v_1 \otimes v_2 \otimes \dots \otimes v_m, w_1 \otimes w_2 \otimes \dots \otimes w_m) = \prod_{i=1}^m (v_i, w_i),$$

$v_i, w_i \in V_i$, $1 \leq i \leq m$, is called the **induced inner product**.

COROLLARY 5.33 Let $V_1, V_2, \dots, V_m$ be inner product spaces. Suppose $T_i \in L(V_i, V_i)$, $1 \leq i \leq m$. Then (with respect to the induced inner product)

$$(T_1 \otimes T_2 \otimes \dots \otimes T_m)^* = T_1^* \otimes T_2^* \otimes \dots \otimes T_m^*. \quad (5.24)$$

Proof Observe that

$$\begin{aligned} & ((T_1 \otimes T_2 \otimes \dots \otimes T_m)(v_1 \otimes v_2 \otimes \dots \otimes v_m), w_1 \otimes w_2 \otimes \dots \otimes w_m) \\ &= (T_1(v_1) \otimes T_2(v_2) \otimes \dots \otimes T_m(v_m), w_1 \otimes w_2 \otimes \dots \otimes w_m) \\ &= \prod_{i=1}^m (T_i(v_i), w_i) \\ &= \prod_{i=1}^m (v_i, T_i^*(w_i)) \\ &= (v_1 \otimes v_2 \otimes \dots \otimes v_m, T_1^*(w_1) \otimes T_2^*(w_2) \otimes \dots \otimes T_m^*(w_m)) \\ &= (v_1 \otimes v_2 \otimes \dots \otimes v_m, (T_1^* \otimes T_2^* \otimes \dots \otimes T_m^*)(w_1 \otimes w_2 \otimes \dots \otimes w_m)). \end{aligned}$$

This completes the proof because the adjoint is unique and the decomposable tensors span $V_1 \otimes V_2 \otimes \dots \otimes V_m$. $\square$

We come at last to the “enormous simplifications” promised at the beginning of the chapter. Although it is not always necessary, from now on we are going to assume that $V_1 = V_2 = \dots = V_m = V$. This identification makes the notation $V \otimes V \otimes \dots \otimes V$ obsolete.

DEFINITION 5.34 The m -th tensor power of V , denoted $V^{\otimes m}$, is the tensor product of m copies of V , that is, $V^{\otimes 0} = \mathbb{C}$, $V^{\otimes 1} = V$, and $V^{\otimes m} = V_1 \otimes V_2 \otimes \dots \otimes V_m$, $m > 1$, where $V_1 = V_2 = \dots = V_m = V$.

With the vector spaces all equal, it is no longer necessary to deal with m bases; one will do. We can get by with one less subscript. If $\mathcal{B} = \{e_1, e_2, \dots, e_n\}$ is a

basis of V , then

$$\{e_{j_1} \otimes e_{j_2} \otimes \cdots \otimes e_{j_m} : 1 \leq j_i \leq n, 1 \leq i \leq m\}$$

is the induced basis of $V^{\otimes m}$. Unfortunately, we are still left with such monstrosities as

$$v = \sum_{j_1=1}^n \sum_{j_2=1}^n \cdots \sum_{j_m=1}^n c_{(j_1, j_2, \dots, j_m)} e_{j_1} \otimes e_{j_2} \otimes \cdots \otimes e_{j_m}. \quad (5.25)$$

Recall (Definition 1.20) that $\Gamma_{m,n}$ is the set of all functions from $\{1, 2, \dots, m\}$ into $\{1, 2, \dots, n\}$. Identifying $(j_1, j_2, \dots, j_m)$ with the function $\alpha \in \Gamma_{m,n}$ defined by $\alpha(i) = j_i$, $1 \leq i \leq m$, allows us to rewrite Equation (5.25) more attractively as

$$v = \sum_{\alpha \in \Gamma_{m,n}} c_{\alpha} e_{\alpha(1)} \otimes e_{\alpha(2)} \otimes \cdots \otimes e_{\alpha(m)}. \quad (5.26)$$

DEFINITION 5.35 If $v_1, v_2, \dots, v_m \in V$ and $\alpha \in \Gamma_{m,n}$, let

$$v_{\alpha}^{\otimes} = v_{\alpha(1)} \otimes v_{\alpha(2)} \otimes \cdots \otimes v_{\alpha(m)}.$$

If $\mathcal{B} = \{e_1, e_2, \dots, e_n\}$ is an ordered basis of V then, in the notation of Definition 5.35, the induced basis of $V^{\otimes m}$ is

$$\{e_{\alpha}^{\otimes} : \alpha \in \Gamma_{m,n}\}, \quad (5.27)$$

which we take to be ordered lexicographically by the subscripts $\alpha \in \Gamma_{m,n}$. We can now simplify Equation (5.26) even further and write the generic tensor in $V^{\otimes m}$ as

$$v = \sum_{\alpha \in \Gamma_{m,n}} c_{\alpha} e_{\alpha}^{\otimes}. \quad (5.28)$$

EXAMPLE 5.36 Let $A_r = (a_{ij}^r)$ and $B_r = (b_{ij}^r)$ be n -by- n complex matrices, $1 \leq i \leq m$. Then the equation

$$\begin{aligned} & (A_1 \otimes A_2 \otimes \cdots \otimes A_m)(B_1 \otimes B_2 \otimes \cdots \otimes B_m) \\ &= A_1 B_1 \otimes A_2 B_2 \otimes \cdots \otimes A_m B_m \end{aligned}$$

follows from Theorem 5.20 and Corollary 5.27. However, a direct proof may be instructive. From the definition of matrix multiplication, the (α, β) -entry of

$(A_1 \otimes A_2 \otimes \cdots \otimes A_m)(B_1 \otimes B_2 \otimes \cdots \otimes B_m)$ is

$$\begin{aligned}
 & \sum_{\gamma \in \Gamma_{m,n}} \left(\prod_{r=1}^m a_{\alpha(r)\gamma(r)}^r \right) \left(\prod_{r=1}^m b_{\gamma(r)\beta(r)}^r \right) \\
 &= \sum_{\gamma \in \Gamma_{m,n}} \prod_{r=1}^m a_{\alpha(r)\gamma(r)}^r b_{\gamma(r)\beta(r)}^r \\
 &= \prod_{r=1}^m \left(\sum_{j=1}^n a_{\alpha(r)j}^r b_{j\beta(r)}^r \right) \\
 &= \prod_{r=1}^m (A_r B_r)_{\alpha(r)\beta(r)},
 \end{aligned}$$

which is the (α, β) -entry of $A_1 B_1 \otimes A_2 B_2 \otimes \cdots \otimes A_m B_m$. $\square$

THEOREM 5.37 Suppose $A_i \in \mathbb{C}_{n,n}$, $1 \leq i \leq m$. If $A_i \geq 0$, $1 \leq i \leq m$, then $A_1 \otimes A_2 \otimes \cdots \otimes A_m \geq 0$.

Proof Suppose $A_i = B_i^* B_i$, $1 \leq i \leq m$. Then

$$\begin{aligned}
 A_1 \otimes A_2 \otimes \cdots \otimes A_m &= (B_1^* B_1) \otimes (B_2^* B_2) \otimes \cdots \otimes (B_m^* B_m) \\
 &= (B_1^* \otimes B_2^* \otimes \cdots \otimes B_m^*) (B_1 \otimes B_2 \otimes \cdots \otimes B_m) \\
 &= (B_1 \otimes B_2 \otimes \cdots \otimes B_m)^* (B_1 \otimes B_2 \otimes \cdots \otimes B_m).
 \end{aligned}$$

$\square$

COROLLARY 5.38 Let B_i and C_i be n -by- n positive semidefinite hermitian matrices. Suppose $A_i = B_i + C_i$, $1 \leq i \leq m$. Then

$$A_1 \otimes A_2 \otimes \cdots \otimes A_m \geq B_1 \otimes B_2 \otimes \cdots \otimes B_m + C_1 \otimes C_2 \otimes \cdots \otimes C_m.$$

Proof From the multilinearity of the Kronecker product,

$$\begin{aligned}
 A_1 \otimes A_2 \otimes \cdots \otimes A_m &= (B_1 + C_1) \otimes (B_2 + C_2) \otimes \cdots \otimes (B_m + C_m) \\
 &= B_1 \otimes B_2 \otimes \cdots \otimes B_m + \sum C_1 \otimes C_2 \otimes \cdots \otimes C_m,
 \end{aligned}$$

where “ $\sum$ ” represents the sum of all $2^m - 2$ terms of the form $X_1 \otimes X_2 \otimes \cdots \otimes X_m$ where each X_i is either B_i or C_i , and at least one of each occurs. By Theorem 5.37, each of these $2^m - 2$ terms is positive semidefinite hermitian. $\square$

COROLLARY 5.39 *Let V be an inner product space. If $S_i, T_i \in L(V, V)$, are positive semidefinite, $1 \leq i \leq m$, then $S_1 \otimes S_2 \otimes \cdots \otimes S_m$ and $T_1 \otimes T_2 \otimes \cdots \otimes T_m$ are positive semidefinite. Moreover,*

$$(S_1 + T_1) \otimes (S_2 + T_2) \otimes \cdots \otimes (S_m + T_m) \geq S_1 \otimes S_2 \otimes \cdots \otimes S_m + T_1 \otimes T_2 \otimes \cdots \otimes T_m.$$

Corollary 5.39 is the operator version of Corollary 5.38. Another variation on this theme is the following: Let V be an inner product space of dimension n . Suppose $T \in L(V, V)$ is positive semidefinite. By the Spectral Theorem, there is an orthonormal basis $\{u_1, u_2, \dots, u_n\}$ of V consisting of eigenvectors of T , that is, $T(u_i) = \lambda_i u_i$ and $\lambda_i \geq 0$, $1 \leq i \leq n$. If r is a positive real number, define $T^r \geq 0$ by

$$T^r(u_i) = \lambda_i^r u_i, \quad 1 \leq i \leq n, \quad (5.29)$$

and linear extension. (Compare with Equation (2.16).) The following is a special case of a much more general result proved in [Lieb (1973)] and [Ando (1979)]:

THEOREM 5.40 *Let V be an n -dimensional inner product space. Suppose S_i and T_i are positive semidefinite hermitian operators on V , $1 \leq i \leq n$. If $0 \leq \theta \leq 1$, then*

$$\begin{aligned} & (\theta S_1 + (1 - \theta)T_1)^{1/n} \otimes (\theta S_2 + (1 - \theta)T_2)^{1/n} \otimes \cdots \otimes (\theta S_n + (1 - \theta)T_n)^{1/n} \\ & \geq \theta \left(S_1^{1/n} \otimes S_2^{1/n} \otimes \cdots \otimes S_n^{1/n} \right) + (1 - \theta) \left(T_1^{1/n} \otimes T_2^{1/n} \otimes \cdots \otimes T_n^{1/n} \right). \end{aligned}$$

DEFINITION 5.41 If $T \in L(V, V)$, denote by $T^{\otimes m} \in L(V^{\otimes m}, V^{\otimes m})$ the operator $T \otimes T \otimes \cdots \otimes T$ (m -times). Similarly, if $A = (a_{ij}) \in \mathbb{C}_{n,n}$, its m -th Kronecker power is the matrix $A^{\otimes m} = A \otimes A \otimes \cdots \otimes A$ (m -times). Thus, $A^{\otimes m}$ is an n^m -by- n^m matrix whose rows and columns are indexed by $\Gamma_{m,n}$. The (α, β) -entry of $A^{\otimes m}$ is

$$(A^{\otimes m})_{\alpha, \beta} = \prod_{i=1}^m a_{\alpha(i)\beta(i)}. \quad (5.30)$$

Combining Theorem 5.40 and Definition 5.41, we obtain the following: If A and B are positive semidefinite hermitian n -by- n matrices and $0 \leq \theta \leq 1$, then

$$\left((\theta A + (1 - \theta)B)^{1/n} \right)^{\otimes n} \geq \theta (A^{1/n})^{\otimes n} + (1 - \theta) (B^{1/n})^{\otimes n}.$$

Exercises

1. Suppose $V_1, V_2, \dots, V_m$ are one-dimensional vector spaces over $\mathbb{C}$, that is, $V_1 = V_2 = \dots = V_m = \mathbb{C}$. Define $f: V_1 \times V_2 \times \dots \times V_m \rightarrow \mathbb{C}$ by

$$f(c_1, c_2, \dots, c_m) = \prod_{i=1}^m c_i.$$

Prove that f is m -linear. (This is the prototype for all multilinear functions.)

2. Let $V_1 = V_2 = \dots = V_m = \mathbb{C}_{n,n}$. Define $f: V_1 \times V_2 \times \dots \times V_m \rightarrow \mathbb{C}_{n,n}$ by

$$f(C_1, C_2, \dots, C_m) = C_1 C_2 \dots C_m.$$

Prove that f is m -linear.

3. Explicitly write out all the elements of the basis $\mathcal{B}$ in Equation (5.10) when $m = 2 = n_1$, and $n_2 = 3$.
4. Let $\{u, v, x, y\}$ be a linear independent set of vectors in a vector space V . Show that the tensor $u \otimes v + x \otimes y \in V \otimes V$ is not decomposable.
5. Finish the proof of Theorem 5.15. (Hint: $0 + 0 = 0$.)
6. Finish the proof of Theorem 5.16.
7. Let $\{v_1, v_2, \dots, v_k\}$ be a linearly independent set of vectors.
- a. Prove that

$$\sum_{i=1}^k v_i \otimes u_i = 0$$

if and only if $u_1 = u_2 = \dots = u_k = 0$.

- b. Suppose $A = (a_{ij}) \in \mathbb{C}_{k,k}$ satisfies $AA^t = I_k$. If

$$w_j = \sum_{i=1}^k a_{ij} v_i, \quad 1 \leq j \leq k,$$

prove that

$$\sum_{i=1}^k v_i \otimes v_i = \sum_{i=1}^k w_i \otimes w_i.$$

8. Prove Theorem 5.14.
9. Suppose $(\mathbb{T}, \Phi)$ is a universal pair for $V_1, V_2, \dots, V_m$. Prove that the following conditions are equivalent.

- (i) The reach of Φ is all of T .
- (ii) For every m -linear function $f: V_1 \times V_2 \times \cdots \times V_m \rightarrow W$, there exists a *unique* linear transformation $h: T \rightarrow W$ such that $f = h\Phi$.
10. Let $\mathcal{B}_i = \{e_{i1}, e_{i2}, \dots, e_{in_i}\}$ be a basis of V_i , $1 \leq i \leq m$. Suppose $1 \leq k < m$. Define

$$\Psi: (V_1 \otimes V_2 \otimes \cdots \otimes V_k) \times (V_{k+1} \otimes V_{k+2} \otimes \cdots \otimes V_m) \rightarrow V_1 \otimes V_2 \otimes \cdots \otimes V_m$$

by

$$\begin{aligned} \Psi((e_{1i_1} \otimes e_{2i_2} \otimes \cdots \otimes e_{ki_k}), (e_{(k+1)i_{k+1}} \otimes e_{(k+2)i_{k+2}} \otimes \cdots \otimes e_{mi_m})) \\ = e_{1i_1} \otimes e_{2i_2} \otimes \cdots \otimes e_{mi_m}, \end{aligned}$$

and bilinear extension. Prove that

$$\Psi((v_1 \otimes v_2 \otimes \cdots \otimes v_k), (v_{k+1} \otimes v_{k+2} \otimes \cdots \otimes v_m)) = v_1 \otimes v_2 \otimes \cdots \otimes v_m,$$

for all $v_i \in V_i$, $1 \leq i \leq m$.

11. Let $V^\otimes$ be the direct sum of the vector spaces $V^{\otimes m}$, $m \geq 0$. Define a “product” on the elements of $V^\otimes$ by $uv = u \otimes v$, $u, v \in V^\otimes$. Prove that
- $(au)v = u(av) = a(uv)$ for all $a \in \mathbb{C}$ and all $u, v \in V^\otimes$.
 - $u(v + w) = uv + uw$ for all $u, v, w \in V^\otimes$.
 - $uv = 0$ implies $u = 0$ or $v = 0$.
12. Let $\mathcal{B} = \{e_1, e_2, \dots, e_n\}$ be a basis of $V_1 = V_2 = \cdots = V_m = V$. Suppose $T_i \in L(V, V)$, $1 \leq i \leq m$. Define $T_1 \otimes T_2 \otimes \cdots \otimes T_m$ by

$$\begin{aligned} (T_1 \otimes T_2 \otimes \cdots \otimes T_m)(e_{\alpha(1)} \otimes e_{\alpha(2)} \otimes \cdots \otimes e_{\alpha(m)}) \\ = T_1(e_{\alpha(1)}) \otimes T_2(e_{\alpha(2)}) \otimes \cdots \otimes T_m(e_{\alpha(m)}), \end{aligned}$$

$\alpha \in \Gamma_{m,n}$, and linear extension. Prove directly (using the multilinearity of the tensor product, but not the Universal Factorization Property) that

$$(T_1 \otimes T_2 \otimes \cdots \otimes T_m)(v_1 \otimes v_2 \otimes \cdots \otimes v_m) = T_1(v_1) \otimes T_2(v_2) \otimes \cdots \otimes T_m(v_m).$$

14. Prove Theorem 5.20.
15. Using Theorem 5.15 and Definition 5.19 (but not Theorem 5.22), prove directly that $T_1 \otimes T_2 \otimes \cdots \otimes T_m = 0$, the zero transformation, if and only if $T_i = 0$ for some i .

16. Suppose $\sigma \in S_m$. Prove that there exists a unique invertible linear operator $P(\sigma^{-1}) \in L(V^{\otimes m}, V^{\otimes m})$ satisfying

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

for all $v_1, v_2, \dots, v_m \in V$.

17. Show that the operator $P(\sigma^{-1})$ from Exercise 16 is not of the form $T_1 \otimes T_2 \otimes \cdots \otimes T_m$ unless $\sigma = e_m$, or $\dim(V) = 1$.
18. Prove that $\sigma \rightarrow P(\sigma)$ is a representation of S_m , where $P(\sigma^{-1})$ is defined in Exercise 16. (Hint: Show that $P(\sigma\tau) = P(\sigma)P(\tau)$ and that $P(e_m)$ is the identity operator on $V^{\otimes m}$.)
19. Explicitly write out the Kronecker product
- a. $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \otimes I_2$ b. $I_2 \otimes \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ c. $\begin{pmatrix} 1 \\ 2 \end{pmatrix} \otimes (3 \ 4)$
20. Let $A, B, C \in \mathbb{C}_{n,n}$.
- a. Show that $A \otimes (B + C) = A \otimes B + A \otimes C$. (Hint: What about $A \otimes (B + C) \otimes D \otimes E$?)
- b. Show that $(A \oplus B) \otimes C = (A \otimes C) \oplus (B \otimes C)$.
- c. Show that, in general, $A \otimes (B \oplus C) \neq (A \otimes B) \oplus (A \otimes C)$.
21. Let $n = pq$. Show that there exists a permutation matrix $P \in \mathbb{C}_{n,n}$ such that for all $A \in \mathbb{C}_{p,p}$ and $B \in \mathbb{C}_{q,q}$, $P^{-1}(A \otimes B)P = B \otimes A$. (See [Hartwig & Morris (1975)] and [Lewis (1996)] for interesting perspectives on P .)
22. If A, B , and C are matrices, prove or disprove that $A \otimes (B \otimes C) = (A \otimes B) \otimes C$.
23. Show that the Kronecker product of two permutation matrices is a permutation matrix.
24. Suppose $A \in \mathbb{C}_{p,p}$ and $B \in \mathbb{C}_{q,q}$.
- a. Prove that $(A \otimes I_q)(I_p \otimes B) = A \otimes B$.
- b. Prove that $\det(I_p \otimes B) = \det(B)^p$. (Hint: Example 5.28.)
- c. Prove that $\det(A \otimes I_q) = \det(A)^q$. (Hint: Exercise 21.)
- d. Prove that $\det(A \otimes B) = \det(A)^q \det(B)^p$.
25. Let $A_1, A_2, \dots, A_m$ be square matrices. If $A_1 \otimes A_2 \otimes \cdots \otimes A_m$ is invertible, prove that A_i is invertible $1 \leq i \leq m$.
26. If A_i is similar to B_i , $1 \leq i \leq m$, prove that $A_1 \otimes A_2 \otimes \cdots \otimes A_m$ is similar to $B_1 \otimes B_2 \otimes \cdots \otimes B_m$.
27. Suppose $A \in \mathbb{C}_{p,p}$ and $B \in \mathbb{C}_{q,q}$. Prove that $\operatorname{tr}(A \otimes B) = \operatorname{tr}(A) \operatorname{tr}(B)$.

28. Let G be a group. Suppose that χ_i is a character of G afforded by the representation A_i , $1 \leq i \leq m$. If

$$B(\sigma) = A_1(\sigma) \otimes A_2(\sigma) \otimes \cdots \otimes A_m(\sigma), \quad \sigma \in G,$$

Prove that

- a. B is a representation of G .
- b. the character afforded by B is given by

$$\chi(\sigma) = \prod_{i=1}^m \chi_i(\sigma), \quad \sigma \in G.$$

29. Suppose G_1 is a subgroup of S_r and G_2 a subgroup of S_t . Set $m = r + t$ and let G be the direct product, $G_1 G_2$. Suppose χ_i is a character of G_i afforded by the representation A_i , $i = 1, 2$.
 - a. Show that $B(\sigma, \tau) = A_1(\sigma) \otimes A_2(\tau)$, $(\sigma, \tau) \in G$, defines a representation of G .
 - b. Let ξ be the character of G afforded by B . Show that $\xi(\sigma, \tau) = \chi_1(\sigma)\chi_2(\tau)$, $(\sigma, \tau) \in G$.
30. Prove that Equation (5.22) defines an inner product on $V_1 \otimes V_2 \otimes \cdots \otimes V_m$. (Hint: Explain why it may be regarded as the dot product.)
31. Let $P(\sigma) \in L(V^{\otimes m}, V^{\otimes m})$ be the linear operator from Exercises 16–18. Show that, with respect to the induced inner product, $P(\sigma)^* = P(\sigma^{-1}) = P(\sigma)^{-1}$.
32. Suppose $T_1 \otimes T_2 \otimes \cdots \otimes T_m$ is hermitian with respect to the induced inner product. Prove or disprove that T_i is hermitian, $1 \leq i \leq m$.
33. Suppose $T_i \in L(V_i, V_i)$ is normal, $1 \leq i \leq m$. Prove that $T_1 \otimes T_2 \otimes \cdots \otimes T_m$ is normal with respect to the induced inner product. Discuss the converse.
34. Suppose $T_i \in L(V_i, V_i)$ is unitary, $1 \leq i \leq m$. Prove that $T_1 \otimes T_2 \otimes \cdots \otimes T_m$ is unitary with respect to the induced inner product. Discuss the converse.
35. Suppose A and B are positive semidefinite hermitian n -by- n matrices. Prove that $A \cdot B \geq 0$. (Hint: Example 5.29 and Theorem 5.37.)
36. In the proof of Corollary 5.38, let $m = 3$ and explicitly write out all 6 (mixed) terms of the form $X_1 \otimes X_2 \otimes X_3$, where $X_i = B_i$ or C_i , $1 \leq i \leq 3$.
37. Suppose $A_i \geq B_i$, $1 \leq i \leq m$. Prove that $A_1 \otimes A_2 \otimes \cdots \otimes A_m \geq B_1 \otimes B_2 \otimes \cdots \otimes B_m$.
38. Let $\lambda_1, \lambda_2, \dots, \lambda_p$ and $\eta_1, \eta_2, \dots, \eta_q$ be the eigenvalues of $A \in \mathbf{C}_{p,p}$ and $B \in \mathbf{C}_{q,q}$, respectively.

- a. Prove that the eigenvalues of $A \otimes B$ are $\lambda_i \eta_j$, $1 \leq i \leq p$, $1 \leq j \leq q$. (Hint: Let u_i be an eigenvector of A corresponding to λ_i and v_j an eigenvector of B corresponding to η_j . Prove that $u_i \otimes v_j$ is an eigenvector of $A \otimes B$ corresponding to $\lambda_i \eta_j$.)
- b. Use part (a) to prove that $\det(A \otimes B) = \det(A)^q \det(B)^p$.
39. Let $\lambda_1, \lambda_2, \dots, \lambda_p$ and $\eta_1, \eta_2, \dots, \eta_q$ be the eigenvalues of $A \in \mathbb{C}_{p,p}$ and $B \in \mathbb{C}_{q,q}$, respectively. Prove that the eigenvalues of $A \otimes I_q + I_p \otimes B$ are $\lambda_i + \eta_j$, $1 \leq i \leq p$, $1 \leq j \leq q$.
40. Let $X = (x_{ij})$ be an n -by- n matrix whose entries are n^2 independent indeterminates (variables) over the complex numbers. If

$$f = f(x_{11}, x_{12}, x_{13}, \dots, x_{ij}, \dots, x_{nn})$$

is a monomial of (total) degree m in the n^2 variables comprising the entries of X , show that f is an entry of $X^{\otimes m}$.

41. Suppose G is a subgroup of S_n . Let $\sigma \rightarrow Q(\sigma) = (\delta_{i, \sigma(j)})$, $\sigma \in G$, be the natural representation of G by permutation matrices. Then Q affords the fixed point character of G . Let r be a fixed positive integer. Define $K_r(\sigma) = Q(\sigma)^{\otimes r}$, the r -th Kronecker power of $Q(\sigma)$.
- a. Show that $\sigma \rightarrow K_r(\sigma)$, $\sigma \in G$, is a representation of G .
- b. Show that the character of K_r is

$$F(\sigma)^r = \sum_{\gamma \in \Gamma_{r,n}} \delta_{\gamma, \sigma\gamma}, \quad \sigma \in G.$$

- c. Define $G(\gamma) = \{\sigma \in G : \sigma\gamma = \gamma\}$, $\gamma \in \Gamma_{r,n}$. If $\chi \in I(G)$, show that the multiplicity of χ as an irreducible constituent of F^r is

$$(\chi, F^r)_G = \sum_{\gamma \in \Gamma_{r,n}} (\chi, 1)_{G(\gamma)} / [G : G(\gamma)],$$

where $[G : G(\gamma)] = o(G)/o(G(\gamma))$ is the index of $G(\gamma)$ in G .

- d. Show that $(\chi, F^r)_G \leq (\chi, F^{r+1})_G$.
- e. Let χ_π be the irreducible character of S_n corresponding to $\pi \vdash n$. If $1 \leq r < n$, show that

$$(\chi_\pi, F^r)_n = \sum_{t=1}^r S(r, t) (\chi_\pi, 1)_{n-t},$$

where $S(r, t)$ is the number of ways to partition an r -element set into the disjoint union of t nonempty subsets. (These numbers are known as Stirling Numbers of the Second Kind.)

- f. Let χ_π be the irreducible character of S_n corresponding to $\pi \vdash n$. Let k be minimal so that χ_π is an irreducible constituent of F^k . Prove that $(\chi_\pi, F^k)_n = (\chi_\pi, 1)_{n-k}$.
- g. Let χ_π be the irreducible character of S_n corresponding to the partition $\pi = [\pi_1, \pi_2, \dots, \pi_s]$ of n . If $s > 1$, prove that the smallest positive integer k such that χ_π is an irreducible constituent of F^k is $k = n - \pi_1$. (Hint: Use part (f) along with several applications of the Branching Theorem.)
42. Suppose $A, B \in \mathbb{C}_{n,n}$. Prove that $(A \cdot B)^*(A \cdot B) \leq (A^*A) \cdot (B^*B)$. (Hint: Exercise 51, Chapter 2.)
43. Suppose A is a real n -by- n matrix. Define

$$e^A = \sum_{n=0}^{\infty} A^n / n!.$$

- a. If $AB = BA$, show that $e^{A+B} = e^A e^B$.
- b. Show that $e^A \otimes e^B = e^{A \otimes I + I \otimes B}$, where $I = I_n$.
44. A **Hadamard matrix** of order n is an n -by- n matrix H , each of whose entries is either $+1$ or -1 , that satisfies the condition $HH^t = nI_n$.
- a. Exhibit a Hadamard matrix of order 2.
- b. Prove that there is no Hadamard matrix of order 3.
- c. Prove that there is a Hadamard matrix of order $n = 2^k$ for every positive integer k . (Hint: If H_1 and H_2 are Hadamard matrices, prove that $H_1 \otimes H_2$ is a Hadamard matrix.)
- d. Can you find a Hadamard matrix of order 4 that is not the Kronecker product of two Hadamard matrices of order 2?

CHAPTER 6

Symmetry Classes of Tensors

For a fixed but arbitrary $\sigma \in S_m$, define

$$\Psi: \overbrace{V \times V \times \cdots \times V}^{m \text{ times}} \rightarrow V^{\otimes m}$$

by

$$\Psi(v_1, v_2, \dots, v_m) = v_{\sigma^{-1}(1)} \otimes v_{\sigma^{-1}(2)} \otimes \cdots \otimes v_{\sigma^{-1}(m)}.$$

Because Ψ is m -linear, there exists a unique¹ linear transformation $P(\sigma) \in L(V^{\otimes m}, V^{\otimes m})$ such that the diagram in Figure 6.1 is commutative. In other words,

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

for all decomposable tensors $v_1 \otimes v_2 \otimes \cdots \otimes v_m \in V^{\otimes m}$.

Observe that

$$\begin{aligned} & P(\tau)P(\sigma)(v_1 \otimes v_2 \otimes \cdots \otimes v_m) \\ &= P(\tau)(v_{\sigma^{-1}(1)} \otimes v_{\sigma^{-1}(2)} \otimes \cdots \otimes v_{\sigma^{-1}(m)}) \\ &= v_{\sigma^{-1}(\tau^{-1}(1))} \otimes v_{\sigma^{-1}(\tau^{-1}(2))} \otimes \cdots \otimes v_{\sigma^{-1}(\tau^{-1}(m))} \\ &= P(\tau\sigma)(v_1 \otimes v_2 \otimes \cdots \otimes v_m). \end{aligned}$$

¹Uniqueness is guaranteed by Exercise 9, Chapter 5.

$$\begin{array}{ccc}
 V \times V \times \dots \times V & \rightarrow & V^{\otimes m} \\
 \Psi \downarrow & & \downarrow P(\sigma) \\
 & & V^{\otimes m}
 \end{array}$$

FIGURE 6.1

Therefore (because the decomposable tensors span $V^{\otimes m}$),

$$P(\tau)P(\sigma) = P(\tau\sigma), \quad (6.1)$$

for all $\sigma, \tau \in S_m$. Moreover, $P(\sigma)$ is invertible, and

$$P(\sigma)^{-1} = P(\sigma^{-1}), \quad \sigma \in S_m, \quad (6.2)$$

so P is a representation of S_m .

With respect to the induced inner product,

$$\begin{aligned}
 & (P(\sigma)(v_1 \otimes v_2 \otimes \dots \otimes v_m), w_1 \otimes w_2 \otimes \dots \otimes w_m) \\
 &= (v_{\sigma^{-1}(1)} \otimes v_{\sigma^{-1}(2)} \otimes \dots \otimes v_{\sigma^{-1}(m)}, w_1 \otimes w_2 \otimes \dots \otimes w_m) \\
 &= \prod_{i=1}^m (v_{\sigma^{-1}(i)}, w_i) \\
 &= \prod_{i=1}^m (v_i, w_{\sigma(i)}) \\
 &= (v_1 \otimes v_2 \otimes \dots \otimes v_m, P(\sigma^{-1})(w_1 \otimes w_2 \otimes \dots \otimes w_m)),
 \end{aligned}$$

for all decomposable tensors $v_1 \otimes v_2 \otimes \dots \otimes v_m$ and $w_1 \otimes w_2 \otimes \dots \otimes w_m$ in $V^{\otimes m}$. Because the decomposable tensors span $V^{\otimes m}$, we conclude² that

$$P(\sigma)^* = P(\sigma^{-1}), \quad \sigma \in S_m. \quad (6.3)$$

²See Exercise 20, Chapter 2.

DEFINITION 6.1 Let G be a subgroup of S_m . If χ is an irreducible character of G . The symmetrizer, $T(G, \chi) \in L(V^{\otimes m}, V^{\otimes m})$, is defined by

$$T(G, \chi) = \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma) P(\sigma). \quad (6.4)$$

LEMMA 6.2 Let G be a subgroup of S_m . Let χ be an irreducible character of G . If $\tau \in G$ is fixed but arbitrary, then $P(\tau)T(G, \chi) = T(G, \chi)P(\tau)$.

Proof

$$\begin{aligned} P(\tau)T(G, \chi) &= \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma) P(\tau)P(\sigma) \\ &= \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma) P(\tau\sigma) \\ &= \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\tau^{-1}\sigma) P(\sigma) \\ &= \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma\tau^{-1}) P(\sigma) \\ &= \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma) P(\sigma\tau) \\ &= \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma) P(\sigma)P(\tau) \\ &= T(G, \chi)P(\tau), \end{aligned}$$

by Equation (6.1) and Exercise 14, Chapter 4. □

THEOREM 6.3 Suppose χ is an irreducible character of the subgroup G of S_m . If V is a vector space of dimension n , then $T(G, \chi)$ is an orthogonal projection on $V^{\otimes m}$.

Proof We need to show that $T(G, \chi)^* = T(G, \chi) = T(G, \chi)^2$. Observe that

$$\begin{aligned} T(G, \chi)^* &= \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \overline{\chi(\sigma)} P(\sigma)^* \\ &= \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma^{-1}) P(\sigma^{-1}) \\ &= T(G, \chi), \end{aligned}$$

by Equation (6.3); and

$$\begin{aligned}
 T(G, \chi)^2 &= \frac{\chi(e)^2}{o(G)^2} \left(\sum_{\sigma \in G} \chi(\sigma) P(\sigma) \right) \left(\sum_{\tau \in G} \chi(\tau) P(\tau) \right) \\
 &= \frac{\chi(e)^2}{o(G)^2} \sum_{\sigma, \tau \in G} \chi(\sigma) \chi(\tau) P(\sigma) P(\tau) \\
 &= \frac{\chi(e)^2}{o(G)^2} \sum_{\sigma, \tau \in G} \chi(\sigma) \chi(\tau) P(\sigma\tau) \\
 &= \frac{\chi(e)}{o(G)} \sum_{\mu \in G} \left(\frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma) \chi(\sigma^{-1}\mu) \right) P(\mu) \\
 &= \frac{\chi(e)}{o(G)} \sum_{\mu \in G} \chi(\mu) P(\mu) \\
 &= T(G, \chi),
 \end{aligned}$$

by Theorem 4.26. □

DEFINITION 6.4 The image of the projection operator $T(G, \chi)$, denoted $V_\chi(G)$, is called the **symmetry class of tensors** associated with G and χ .

THEOREM 6.5 Let G be a subgroup of S_m . Suppose $\chi, \xi \in I(G)$, the set of irreducible characters of G . If $\xi \neq \chi$, then $T(G, \xi)T(G, \chi) = 0$. Moreover,

$$\sum_{\chi \in I(G)} T(G, \chi) = I, \quad (6.5)$$

the identity operator in $L(V^{\otimes m}, V^{\otimes m})$.

Proof

$$\begin{aligned}
 T(G, \xi)T(G, \chi) &= \frac{\xi(e)\chi(e)}{o(G)^2} \left(\sum_{\sigma \in G} \xi(\sigma) P(\sigma) \right) \left(\sum_{\tau \in G} \chi(\tau) P(\tau) \right) \\
 &= \frac{\xi(e)\chi(e)}{o(G)^2} \sum_{\sigma, \tau \in G} \xi(\sigma) \chi(\tau) P(\sigma\tau) \\
 &= \frac{\xi(e)\chi(e)}{o(G)^2} \sum_{\sigma \in G} \left(\sum_{\mu \in G} \xi(\sigma) \chi(\sigma^{-1}\mu) \right) P(\mu) \\
 &= 0,
 \end{aligned}$$

by Theorem 4.26. Moreover,

$$\begin{aligned}
 \sum_{\chi \in I(G)} T(G, \chi) &= \frac{1}{o(G)} \sum_{\chi \in I(G)} \chi(e) \sum_{\sigma \in G} \chi(\sigma) P(\sigma) \\
 &= \sum_{\sigma \in G} \left(\frac{1}{o(G)} \sum_{\chi \in I(G)} \chi(e) \chi(\sigma) \right) P(\sigma) \\
 &= P(e) \\
 &= I,
 \end{aligned}$$

by the Orthogonality Relations of the Second Kind (Theorem 4.40). $\square$

COROLLARY 6.6 *If G is a subgroup of S_m , then $V^{\otimes m}$ is the orthogonal direct sum of the symmetry classes $V_\chi(G)$ as χ ranges over $I(G)$. In other words, if $I(G) = \{\chi_1, \chi_2, \dots, \chi_k\}$, then*

$$V^{\otimes m} = V_{\chi_1}(G) \perp V_{\chi_2}(G) \perp \dots \perp V_{\chi_k}(G).$$

Proof The result is immediate from Theorems 6.3, 6.5, and the definitions. $\square$

DEFINITION 6.7 Let χ be an irreducible character of the subgroup G of S_m . Let V be a vector space, and suppose $v_1, v_2, \dots, v_m \in V$. Then the **decomposable symmetrized tensor** $v_1 * v_2 * \dots * v_m$ is defined by

$$v_1 * v_2 * \dots * v_m = T(G, \chi)(v_1 \otimes v_2 \otimes \dots \otimes v_m).$$

This notation does not reflect the important fact that “ $*$ ” depends on G and χ . From Corollary 6.6, $v_1 * v_2 * \dots * v_m$ is the piece of $v_1 \otimes v_2 \otimes \dots \otimes v_m$ that belongs to $V_\chi(G)$. In particular, $V_\chi(G)$ is spanned by, but is not generally equal to, the set of its decomposable symmetrized tensors.

EXAMPLE 6.8 Suppose $m = 2$ and $G = S_2$. Then G has two irreducible characters, the principal character and the alternating character. If V is a fixed but arbitrary vector space, and $v_1, v_2 \in V$, let

$$\begin{aligned}
 v_1 \bullet v_2 &= T(S_2, 1)(v_1 \otimes v_2) \\
 &= \frac{1}{2}(v_1 \otimes v_2 + v_2 \otimes v_1)
 \end{aligned}$$

and

$$\begin{aligned} v_1 \wedge v_2 &= T(S_2, \varepsilon)(v_1 \otimes v_2) \\ &= \frac{1}{2}(v_1 \otimes v_2 - v_2 \otimes v_1). \end{aligned}$$

Observe that $v_1 \otimes v_2 = v_1 \bullet v_2 + v_1 \wedge v_2$. Moreover,

$$\begin{aligned} (v_1 \bullet v_2, v_1 \wedge v_2) &= \frac{1}{4}(v_1 \otimes v_2 + v_2 \otimes v_1, v_1 \otimes v_2 - v_2 \otimes v_1) \\ &= \frac{1}{4}((v_1 \otimes v_2, v_1 \otimes v_2) - (v_1 \otimes v_2, v_2 \otimes v_1) \\ &\quad + (v_2 \otimes v_1, v_1 \otimes v_2) - (v_2 \otimes v_1, v_2 \otimes v_1)) \\ &= \frac{1}{4}((v_1, v_1)(v_2, v_2) - (v_1, v_2)(v_2, v_1) \\ &\quad + (v_2, v_1)(v_1, v_2) - (v_2, v_2)(v_1, v_1)) \\ &= 0, \end{aligned}$$

confirming that $V^{\otimes 2} = V_1(S_2) \perp V_\varepsilon(S_2)$. $\square$

EXAMPLE 6.9 Let $V = \mathbb{C}_{1,n}$. Then (Example 5.18) $\mathbb{C}_{n,n}$ is a model for $V^{\otimes 2} = V \otimes V$ in which $X \otimes Y = X'Y$. In this model,

$$X \wedge Y = \frac{1}{2}(X'Y - Y'X),$$

the skew symmetric part of $X'Y$. Because it is spanned by tensors of this form, $V_\varepsilon(S_2) \subset \mathbb{C}_{n,n}$ is the space of skew symmetric matrices. Similarly, $V_1(S_2)$ is spanned by the decomposable symmetrized tensors

$$X \bullet Y = \frac{1}{2}(X'Y + Y'X),$$

from which it follows that $V_1(S_2)$ is the space of symmetric matrices. The fact that $(A, B) = \text{tr}(B^*A)$ is zero when A is skew symmetric and B is symmetric can, of course, be verified directly. $\square$

EXAMPLE 6.10 Let $V = \mathbf{C}_{1,n}$. As in Example 6.9 identify $V^{\otimes 2}$ with $\mathbf{C}_{n,n}$. For any fixed $A \in V^{\otimes 2}$, there exist $X_1, X_2, \dots, X_k$ and $Y_1, Y_2, \dots, Y_k$ in V such that

$$\begin{aligned} A &= \sum_{i=1}^k X_i \otimes Y_i \\ &= \sum_{i=1}^k X_i^t Y_i. \end{aligned}$$

If $\sigma = (12) \in S_2$, then

$$\begin{aligned} P(\sigma)(A) &= P(\sigma) \left(\sum_{i=1}^k X_i^t Y_i \right) \\ &= \sum_{i=1}^k P(\sigma)(X_i \otimes Y_i) \\ &= \sum_{i=1}^k Y_i \otimes X_i \\ &= \sum_{i=1}^k Y_i^t X_i \\ &= A^t. \end{aligned}$$

In other words, $P(\sigma) \in L(\mathbf{C}_{n,n}, \mathbf{C}_{n,n})$ is the transpose operator. $\square$

DEFINITION 6.11 Let χ be an irreducible character of the subgroup G of S_m . Suppose $\Phi: V \times V \times \dots \times V \rightarrow W$ is m -linear. If

$$\frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma^{-1}) \Phi(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(m)}) = \Phi(v_1, v_2, \dots, v_m),$$

for all $v_1, v_2, \dots, v_m \in V$, then Φ is symmetric with respect to G and χ .

LEMMA 6.12 Let χ be an irreducible character of the subgroup G of S_m . Let $\Phi: V \times V \times \dots \times V \rightarrow V_\chi(G)$ be defined by

$$\Phi(v_1, v_2, \dots, v_m) = v_1 * v_2 * \dots * v_m.$$

Then Φ is m -linear and symmetric with respect to G and χ .

EXAMPLE 6.13 Let χ be an irreducible character of G . If $\tau \in G$, then

$$\begin{aligned}
 P(\tau^{-1})(v_1 * v_2 * \cdots * v_m) &= P(\tau^{-1})T(G, \chi)(v_1 \otimes v_2 \otimes \cdots \otimes v_m) \\
 &= T(G, \chi)P(\tau^{-1})(v_1 \otimes v_2 \otimes \cdots \otimes v_m) \\
 &= T(G, \chi)(v_{\tau(1)} \otimes v_{\tau(2)} \otimes \cdots \otimes v_{\tau(m)}) \\
 &= v_{\tau(1)} * v_{\tau(2)} * \cdots * v_{\tau(m)}, \tag{6.6}
 \end{aligned}$$

by Lemma 6.2. If $\chi(e) = 1$, then

$$\begin{aligned}
 v_{\tau(1)} * v_{\tau(2)} * \cdots * v_{\tau(m)} &= \frac{1}{o(G)} \sum_{\sigma \in G} \chi(\sigma) P(\tau^{-1}\sigma)(v_1 \otimes v_2 \otimes \cdots \otimes v_m) \\
 &= \frac{1}{o(G)} \sum_{\sigma \in G} \chi(\tau\sigma) P(\sigma)(v_1 \otimes v_2 \otimes \cdots \otimes v_m) \\
 &= \frac{1}{o(G)} \sum_{\sigma \in G} \chi(\tau)\chi(\sigma) P(\sigma)(v_1 \otimes v_2 \otimes \cdots \otimes v_m) \\
 &= \chi(\tau)v_1 * v_2 * \cdots * v_m. \tag{6.7}
 \end{aligned}$$

□

THEOREM 6.14 (*Universal Factorization Property for Symmetric Multilinear Functions*). Let V and W be vector spaces. Suppose $\chi \in I(G)$. If $\Phi: V \times V \times \cdots \times V \rightarrow W$ is m -linear and symmetric with respect to G and χ , then there exists a unique linear transformation $h_\chi \in L(V_\chi(G), W)$ such that

$$h_\chi(v_1 * v_2 * \cdots * v_m) = \Phi(v_1, v_2, \dots, v_m),$$

for all $v_1, v_2, \dots, v_m \in V$.

Theorem 6.14 is perhaps best illustrated by the commutative diagram in Figure 6.2.

Proof By the (ordinary) Universal Factorization Property, there exists a unique linear transformation $h: V^{\otimes m} \rightarrow W$ such that

$$h(v_1 \otimes v_2 \otimes \cdots \otimes v_m) = \Phi(v_1, v_2, \dots, v_m),$$

for all $v_1, v_2, \dots, v_m \in V$. Therefore, for all $v_1, v_2, \dots, v_m \in V$ and all $\sigma \in G$,

$$h(v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \cdots \otimes v_{\sigma(m)}) = \Phi(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(m)}).$$

$$V \times V \times \dots \times V \rightarrow V_\chi(G)$$

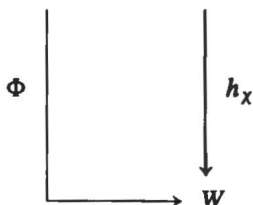


FIGURE 6.2

Multiply both sides of this equation by $\chi(e)\chi(\sigma^{-1})/o(G)$ and sum on $\sigma \in G$. Because h is linear and Φ is symmetric with respect to G and χ , we obtain

$$hT(G, \chi)(v_1 \otimes v_2 \otimes \dots \otimes v_m) = \Phi(v_1, v_2, \dots, v_m),$$

that is,

$$h(v_1 * v_2 * \dots * v_m) = \Phi(v_1, v_2, \dots, v_m).$$

Therefore $h_\chi = hT(G, \chi)$, the restriction of h to $V_\chi(G)$. $\square$

Theorem 6.14 illustrates a situation in which $V^{\otimes m}$ is larger than necessary. To “factor” a multilinear function that is symmetric with respect to G and χ , all we need is $V_\chi(G)$, a piece of $V^{\otimes m}$. On the other hand, taking χ to be the principal character on $G = \{e\}$, we see that $V^{\otimes m} = V_1(\{e\})$, itself, is a symmetry class of tensors.

LEMMA 6.15³ *Let χ be an irreducible character of the subgroup group G of S_m . Suppose $v_1, v_2, \dots, v_m$ and $w_1, w_2, \dots, w_m$ are vectors in V . If $v_1 * v_2 * \dots * v_m = w_1 * w_2 * \dots * w_m \neq 0$ then $\langle v_1, v_2, \dots, v_m \rangle = \langle w_1, w_2, \dots, w_m \rangle$.*

Proof Let $\{e_1, e_2, \dots, e_r\}$ be a basis of $W = \langle v_1, v_2, \dots, v_m \rangle$. If $w_k \notin W$ for some k , let $e_{r+1} = w_k$. Extend the linearly independent set $\{e_1, e_2, \dots, e_r, e_{r+1}\}$ to a basis $B = \{e_1, e_2, \dots, e_n\}$ of V . Let $\{f_1, f_2, \dots, f_n\}$ be the basis of V' dual to B . Because

$$\left\{ \prod_{i=1}^m f_{\alpha(i)} : \alpha \in \Gamma_{m,n} \right\}$$

³Extensions of this result can be found in [Merris (1975a)] and [Lim (1976)]. We shall have more to say about equality of decomposable tensors in Chapter 7.

is a basis of $M = M(V, V, \dots, V)$, and since $v_1 * v_2 * \dots * v_m \neq 0 \in M'$, there exists an element $\beta \in \Gamma_{m,n}$ such that

$$(v_1 * v_2 * \dots * v_m) \left(\prod_{i=1}^m f_{\beta(i)} \right) \neq 0. \quad (6.8)$$

Because,

$$v_1 * v_2 * \dots * v_m = \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma^{-1}) v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \dots \otimes v_{\sigma(m)},$$

the left-hand side of (6.8) takes the value

$$(v_1 * v_2 * \dots * v_m) \left(\prod_{i=1}^m f_{\beta(i)} \right) = \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma^{-1}) \prod_{i=1}^m f_{\beta(i)}(v_{\sigma(i)}). \quad (6.9)$$

It follows that $\beta \in \Gamma_{m,r}$ (otherwise the right-hand side of Equation (6.9) would be 0, contradicting (6.8)). Therefore,

$$\begin{aligned} 0 &\neq (v_1 * v_2 * \dots * v_m) \left(\prod_{i=1}^m f_{\beta(i)} \right) \\ &= (w_1 * w_2 * \dots * w_m) \left(\prod_{i=1}^m f_{\beta(i)} \right) \\ &= \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma^{-1}) \prod_{i=1}^m f_{\beta(i)}(w_{\sigma(i)}). \end{aligned} \quad (6.10)$$

Now, $f_{\beta(\sigma^{-1}(k))}(w_k) = 0$ because $w_k = e_{r+1}$ and $\beta(\sigma^{-1}(k)) \leq r$. Therefore, the right-hand side of Equation (6.10) is zero. This contradiction completes the proof. $\square$

Lemma 6.15 may be viewed as a partial analog of Theorem 5.16; the next result is a partial analog of Theorem 5.15.

LEMMA 6.16 *Let χ be an irreducible character of the permutation group $G \subset S_m$. Suppose $v_1, v_2, \dots, v_m \in V$. If $\{v_1, v_2, \dots, v_m\}$ is linearly independent, then $v_1 * v_2 * \dots * v_m \neq 0$.*

Proof If $\{v_1, v_2, \dots, v_m\}$ is linearly independent, then $m \leq n = \dim(V)$, and we may regard S_m as a subset of $\Gamma_{m,n}$. In particular, $\{v_\sigma^\otimes : \sigma \in G\}$ is part of an induced basis of $V^{\otimes m}$. Therefore,

$$v_1 * v_2 * \dots * v_m = \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma^{-1}) v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \dots \otimes v_{\sigma(m)}$$

is zero if and only if $\chi(\sigma) = 0$ for all $\sigma \in G$. However, $\chi(e) \neq 0$. $\square$

If $\{e_i : 1 \leq i \leq n\}$ is a basis of V , then $\{e_\alpha^\otimes : \alpha \in \Gamma_{m,n}\}$ is a basis of $V^{\otimes m}$. Therefore,

$$\{e_\alpha^* = T(G, \chi)(e_\alpha^\otimes) : \alpha \in \Gamma_{m,n}\} \quad (6.11)$$

spans $V_\chi(G)$. In general, however, some of the elements of (6.11) may be zero, and those that are not generally comprise more than a basis. This raises a number of interesting issues.

If $\alpha \in \Gamma_{m,n}$ and $\sigma \in S_m$, then the composition $\alpha\sigma \in \Gamma_{m,n}$; in sequence notation, it is obtained by rearranging the entries of α according to the permutation σ . That is, $\alpha\sigma = (\alpha\sigma(1), \alpha\sigma(2), \dots, \alpha\sigma(m))$. Evidently, each permutation $\sigma \in S_m$ induces a permutation of $\Gamma_{m,n}$.

DEFINITION 6.17 For each $\sigma \in S_m$, define $\bar{\sigma} : \Gamma_{m,n} \rightarrow \Gamma_{m,n}$ by $\bar{\sigma}(\alpha) = \alpha\sigma^{-1}$, $\alpha \in \Gamma_{m,n}$.

Because $\bar{\sigma} \circ \bar{\tau}(\alpha) = \bar{\sigma}(\alpha\tau^{-1}) = \alpha\tau^{-1}\sigma^{-1} = \alpha(\sigma\tau)^{-1} = \bar{\sigma\tau}(\alpha)$, $\alpha \in \Gamma_{m,n}$, $\sigma \rightarrow \bar{\sigma}$ is a homomorphism⁴ from S_m into the manifestation of S_{n^m} as a group of permutations of $\Gamma_{m,n}$. If G is a subgroup of S_m , let

$$\bar{G} = \{\bar{\sigma} : \sigma \in G\}.$$

Then $\bar{G}$ is a group of permutations of $\Gamma_{m,n}$.

Recall (Definition 3.20) that $\alpha, \beta \in \Gamma_{m,n}$ are equivalent modulo $\bar{G}$ if there is a $\bar{\sigma} \in \bar{G}$ such that $\bar{\sigma}(\alpha) = \beta$. Because $\beta = \alpha\sigma^{-1}$ if and only if $\alpha = \beta\sigma$, equivalence modulo $\bar{G}$ can be described entirely in terms of G . Abusing the language somewhat, we will adopt the following convention.

DEFINITION 6.18 Suppose G is a fixed but arbitrary subgroup of S_m . Let $\alpha, \beta \in \Gamma_{m,n}$. Then α is **equivalent** to β **modulo** G , written $\alpha \equiv \beta \pmod{G}$, if there exists a $\sigma \in G$ such that $\alpha = \beta\sigma$.

⁴If $n > 1$, it is an isomorphism.

EXAMPLE 6.19 Let $m = 3$ and $n = 2$. Then

$$\Gamma_{3,2} = \{(1, 1, 1), (1, 1, 2), (1, 2, 1), (1, 2, 2), (2, 1, 1), (2, 1, 2), (2, 2, 1), (2, 2, 2)\}.$$

There are four equivalence classes of $\Gamma_{3,2}$ modulo S_3 , namely,

$$\begin{aligned} &\{(1, 1, 1)\}, \\ &\{(1, 1, 2), (1, 2, 1), (2, 1, 1)\}, \\ &\{(1, 2, 2), (2, 1, 2), (2, 2, 1)\}, \quad \text{and} \\ &\{(2, 2, 2)\}. \end{aligned}$$

The six equivalence classes modulo $G = \{e, (13)\}$ are

$$\begin{aligned} &\{(1, 1, 1)\}, \\ &\{(1, 1, 2), (2, 1, 1)\}, \\ &\{(1, 2, 1)\}, \\ &\{(1, 2, 2), (2, 2, 1)\}, \\ &\{(2, 1, 2)\}, \quad \text{and} \\ &\{(2, 2, 2)\}. \end{aligned}$$

□

DEFINITION 6.20 Let G be a subgroup of S_m . If $\alpha \in \Gamma_{m,n}$, its **stabilizer subgroup** is $G_\alpha = \{\sigma \in G : \alpha\sigma = \alpha\}$.

Keeping strict faith with Definition 3.24, the stabilizer subgroup should be denoted $\overline{G}_\alpha$, because it is the $\overline{G}$ manifestation of G that acts on $\Gamma_{m,n}$.

LEMMA 6.21 Let $\{e_i : 1 \leq i \leq n\}$ be an orthonormal basis of the inner product space V . Then, with respect to the restriction of the induced inner product of $V^{\otimes m}$ to $V_X(G)$,

$$(e_\alpha^*, e_\beta^*) = \begin{cases} 0, & \text{if } \alpha \not\equiv \beta \pmod{G} \\ \frac{\chi(e)}{o(G)} \sum_{\sigma \in G_\alpha} \chi(\sigma), & \text{if } \alpha \equiv \beta. \end{cases}$$

Notice that Lemma 6.21 is silent about the case in which $\alpha \equiv \beta \pmod{G}$, but $\alpha \neq \beta$.

Proof Because $T(G, \chi)$ is hermitian and idempotent,

$$\begin{aligned}(e_\alpha^*, e_\beta^*) &= (T(G, \chi)(e_\alpha^\otimes), T(G, \chi)(e_\beta^\otimes)) \\ &= (T(G, \chi)(e_\alpha^\otimes), e_\beta^\otimes).\end{aligned}$$

Substituting for $T(G, \chi)$, we obtain

$$(e_\alpha^*, e_\beta^*) = \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma) \prod_{i=1}^m (e_{\alpha(i)}, e_{\beta\sigma(i)}).$$

If $\alpha \not\equiv \beta \pmod{G}$, then every term in this summand is zero. If $\alpha = \beta$, then only the terms corresponding to those σ for which $\alpha = \alpha\sigma$ survive. $\square$

The first conclusion to be drawn from Lemma 6.21 is that

$$\begin{aligned}\|e_\alpha^*\|^2 &= \frac{\chi(e)}{o(G)} \sum_{\sigma \in G_\alpha} \chi(\sigma) \\ &= \chi(e)(\chi, 1)_{G_\alpha} / [G : G_\alpha].\end{aligned}\tag{6.12}$$

Therefore, $e_\alpha^* \neq 0$, if and only if the restriction of χ to G_α contains the principal character as an irreducible constituent. Let

$$\Omega = \{\alpha \in \Gamma_{m,n} : (\chi, 1)_{G_\alpha} \neq 0\}.\tag{6.13}$$

Then $\{e_\alpha^* : \alpha \in \Omega\}$ is what remains after the zeros have been deleted from $\{e_\alpha^* : \alpha \in \Gamma_{m,n}\}$. In particular, $\{e_\alpha^* : \alpha \in \Omega\}$ spans $V_\chi(G)$. (Note that Ω depends on m, n, G , and χ .) If χ is the principal character, then $\Omega = \Gamma_{m,n}$. If $m \leq n$ and α is a one-to-one function (a sequence of distinct integers), then $G_\alpha = \{e\}$, so $\alpha \in \Omega$.

LEMMA 6.22 *Let G be a subgroup of S_m and suppose $\chi \in I(G)$. Then the set Ω defined in Equation (6.13) is a union of equivalence classes of $\Gamma_{m,n}$ modulo G . That is, $\alpha \in \Omega$ if and only if $\alpha\tau \in \Omega$, for all $\tau \in G$.*

DEFINITION 6.23 For a fixed but arbitrary m, n , and G , let Δ be the system of distinct representatives for the equivalence classes of $\Gamma_{m,n}$ modulo G , so chosen that $\alpha \in \Delta$ if and only if α is first, in lexicographic order, in its equivalence class.

LEMMA 6.24 *If $G = S_m$, then $\Delta = G_{m,n}$, the subset of $\Gamma_{m,n}$ consisting of all $C(n+m-1, m)$ nondecreasing functions (sequences).*

EXAMPLE 6.25 Returning to Example 6.19, suppose $m = 3$ and $n = 2$. If $G = S_3$ then $\Delta = \{(1, 1, 1), (1, 1, 2), (1, 2, 2), (2, 2, 2)\} = G_{3,2}$, confirming Lemma 6.24. If $G = \{e, (13)\}$, then, in lexicographic order, $\Delta = \{(1, 1, 1), (1, 1, 2), (1, 2, 1), (1, 2, 2), (2, 1, 2), (2, 2, 2)\}$. $\square$

EXAMPLE 6.26 Let $m = 3$, $n = 2$, and $G = S_3$. If $\pi = [2, 1]$, then $\chi_\pi(\sigma) = F(\sigma) - 1$, $\sigma \in S_3$, where $F(\sigma)$ is the number of fixed points of σ . Let us test the elements of $\Gamma_{3,2}$ for membership in Ω . By Lemma 6.22, it suffices to test the elements of Δ . From Lemma 6.24 (or Example 6.25), $\Delta = G_{3,2} = \{(1, 1, 1), (1, 1, 2), (1, 2, 2), (2, 2, 2)\}$. If $\alpha = (1, 1, 1)$ or $(2, 2, 2)$, then $G_\alpha = S_3$, and $(\chi_\pi, 1)_3 = 0$ by the Orthogonality Relations of the First Kind. Therefore, neither $(1, 1, 1)$, nor $(2, 2, 2)$ belongs to Ω . If $\alpha = (1, 1, 2)$ or $(1, 2, 2)$, then G_α is isomorphic to S_2 . By the Branching Theorem, the restriction of χ_π to S_2 is $1 + \varepsilon$, and

$$\begin{aligned}(1 + \varepsilon, 1)_2 &= (1, 1)_2 + (\varepsilon, 1)_2 \\ &= 1 + 0.\end{aligned}$$

Therefore, Ω is the union of the equivalence classes represented by $(1, 1, 2)$ and $(1, 2, 2)$, that is,

$$\Omega = \{(1, 1, 2), (1, 2, 1), (2, 1, 1)\} \cup \{(1, 2, 2), (2, 1, 2), (2, 2, 1)\}. \quad (6.14)$$

If $\pi = [1^3]$, then $\chi_\pi = \varepsilon$. Now, for every $\alpha \in \Gamma_{3,2}$, there exist $i \neq j$ such that $\alpha(i) = \alpha(j)$. Because $G = S_3$, this means a transposition $(ij) \in G_\alpha$. It follows that the restriction of ε to G_α is not identically 1. Therefore, $(1, \varepsilon)_{G_\alpha} = 0$, for every $\alpha \in \Gamma_{m,n}$. In other words, Ω is empty. Because $\{e_\alpha^* : \alpha \in \Omega\}$ spans $V_\chi(G)$, we conclude that $V_\varepsilon(S_3) = \{0\}$, for any vector space V of dimension 2. $\square$

DEFINITION 6.27 For a fixed but arbitrary subgroup G of S_m and $\chi \in I(G)$, let $\overline{\Delta} = \Delta \cap \Omega$. That is, $\overline{\Delta} = \{\alpha \in \Delta : (\chi, 1)_{G_\alpha} \neq 0\}$.

Like Ω , $\overline{\Delta}$ depends on m , n , G , and χ . Like Δ , each element of $\overline{\Delta}$ is first, in lexicographic order, in the equivalence class modulo G to which it belongs.

EXAMPLE 6.28 Let $G = S_3$ and $\chi = \chi_{[2,1]}$ then, from Equation (6.14), $\overline{\Delta} = \{(1, 1, 2), (1, 2, 2)\}$. $\square$

It follows from the definitions and Lemma 6.22 that

$$\Omega = \bigcup_{\alpha \in \overline{\Delta}} \{\alpha\sigma : \sigma \in G\}. \quad (6.15)$$

EXAMPLE 6.29 Recall that $Q_{m,n}$ is the subset of $\Gamma_{m,n}$ consisting of all $C(n, m)$ strictly increasing functions. If $\alpha \in Q_{m,n}$, and $\sigma \in S_m$, then α comes before $\alpha\sigma$ in lexicographic order. Therefore, for any subgroup G of S_m , α comes first in its equivalence class. That is, $\alpha \in \Delta$. In fact, more is true. Because α is a one-to-one function, $G_\alpha = \{e\}$. Therefore, $\alpha \in \overline{\Delta}$. In other words, $Q_{m,n} \subset \overline{\Delta}$, for every $\chi \in I(G)$. $\square$

THEOREM 6.30 If $G = S_m$ and $\chi = \varepsilon$, then $\overline{\Delta} = Q_{m,n}$.

Because $Q_{3,2}$ is empty, it follows from Theorem 6.30 and the definitions that $V_\epsilon(S_3) = \{0\}$ for any vector space V of dimension 2, confirming an observation already made in Example 6.26.

THEOREM 6.31 *Let χ be an irreducible character of the subgroup G of S_m . If $\mathcal{B} = \{e_1, e_2, \dots, e_n\}$ is a basis of the vector space V , then $V_\chi(G)$ is the direct sum of the subspaces $\langle e_{\alpha\sigma}^* : \sigma \in G \rangle$, as α ranges over $\overline{\Delta}$.*

Proof Because $\{e_\alpha^* : \alpha \in \overline{\Delta}\}$ spans $V_\chi(G)$, it follows from Equation (6.15) that $V_\chi(G)$ is the sum, over $\alpha \in \overline{\Delta}$, of the subspaces $\langle e_{\alpha\sigma}^* : \sigma \in G \rangle$. Consider the inner product on V with respect to which $\mathcal{B}$ is orthonormal. With respect to the induced inner product (Lemma 6.21), the sum of the subspaces $\langle e_{\alpha\sigma}^* : \sigma \in G \rangle$, $\alpha \in \overline{\Delta}$, is orthogonal. $\square$

COROLLARY 6.32 *If $\mathcal{B} = \{e_1, e_2, \dots, e_n\}$ is a basis of V , then*

$$\{e_\alpha^* : \alpha \in \overline{\Delta}\} \quad (6.16)$$

is a linearly independent set in $V_\chi(G)$. If $\mathcal{B}$ is orthogonal, then (6.16) is orthogonal. If $\chi(e) = 1$, then (6.16) is a basis of $V_\chi(G)$.

Proof The set (6.16) consists of one (nonzero) vector from each of the direct summands of Theorem 6.31. If $\mathcal{B}$ is orthogonal then, by Lemma 6.21, (6.16) is orthogonal. (If $\alpha, \beta \in \overline{\Delta}$, then $\alpha \equiv \beta \pmod{G}$ if and only if $\alpha = \beta$.) If $\chi(e) = 1$, it follows from Equation (6.7) that $e_{\alpha\sigma}^* = \chi(\sigma)e_\alpha^*$, for all $\sigma \in G$ and all $\alpha \in \Gamma_{m,n}$. In this case, the subspaces from Theorem 6.31 each have dimension 1. $\square$

DEFINITION 6.33 Suppose $\{e_1, e_2, \dots, e_n\}$ is a basis of the vector space V . Let χ be an irreducible character of G . If $\alpha \in \overline{\Delta}$, then $\langle e_{\alpha\sigma}^* : \sigma \in G \rangle$ is an **orbital subspace** of $V_\chi(G)$ corresponding to α . Let $s_\alpha = \dim(\langle e_{\alpha\sigma}^* : \sigma \in G \rangle)$.

FREESE'S THEOREM 6.34⁵ *Let χ be an irreducible character of G . If $\alpha \in \overline{\Delta}$, then $s_\alpha = \chi(e)(\chi, 1)_{G_\alpha}$.*

Proof Let $G = G_\alpha\sigma_1 \cup G_\alpha\sigma_2 \cup \dots \cup G_\alpha\sigma_r$, $r = [G : G_\alpha]$, be the right coset decomposition of G in G . Then, for any $\mu \in G$,

$$\begin{aligned} e_{\alpha\mu}^* &= \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma^{-1}) e_{\alpha\mu\sigma}^{\otimes} \\ &= \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma^{-1}\mu) e_{\alpha\sigma}^{\otimes} \\ &= \frac{\chi(e)}{o(G)} \sum_{i=1}^r \left(\sum_{\tau \in G_\alpha} \chi(\sigma_i^{-1}\tau^{-1}\mu) \right) e_{\alpha\sigma_i}^{\otimes}, \end{aligned}$$

⁵Theorem 6.34 appeared first in [Freese (1973)]. Note that it captures the result $s_\alpha=1$, $\alpha \in \overline{\Delta}$, when $\chi(e)=1$.

because $\alpha\tau = \alpha$, $\tau \in G_\alpha$. In particular,

$$e_{\alpha\sigma_j}^* = \frac{\chi(e)}{o(G)} \sum_{i=1}^r \left(\sum_{\tau \in G_\alpha} \chi(\sigma_i^{-1}\tau\sigma_j) \right) e_{\alpha\sigma_i}^\otimes. \quad (6.17)$$

Let $C = (c_{ij})$ be the r -by- r matrix defined by

$$c_{ij} = \frac{\chi(e)}{o(G)} \sum_{\tau \in G_\alpha} \chi(\sigma_i^{-1}\tau\sigma_j). \quad (6.18)$$

Because $\{e_{\alpha\sigma_i}^\otimes : 1 \leq i \leq r\}$ is linearly independent, it follows from Equations (6.17) and (6.18) that $s_\alpha = \text{rank}(C)$. Observe that

$$\begin{aligned} (C^2)_{ij} &= \frac{\chi(e)^2}{o(G)^2} \sum_{k=1}^r \sum_{\tau, \sigma \in G_\alpha} \chi(\sigma_i^{-1}\tau\sigma_k) \chi(\sigma_k^{-1}\sigma\sigma_j) \\ &= \frac{\chi(e)^2}{o(G)^2} \sum_{k=1}^r \sum_{\tau, \mu \in G_\alpha} \chi(\sigma_i^{-1}\tau\sigma_k) \chi(\sigma_k^{-1}\tau^{-1}\mu\sigma_j) \\ &= \frac{\chi(e)^2}{o(G)^2} \sum_{\mu \in G_\alpha} \sum_{\tau \in G} \chi(\sigma_i^{-1}\tau) \chi(\tau^{-1}\mu\sigma_j) \\ &= \frac{\chi(e)^2}{o(G)^2} \sum_{\mu \in G_\alpha} \sum_{\tau \in G} \chi(\tau) \chi(\tau^{-1}\sigma_i^{-1}\mu\sigma_j) \\ &= \frac{\chi(e)}{o(G)} \sum_{\mu \in G_\alpha} \chi(\sigma_i^{-1}\mu\sigma_j) \\ &= c_{ij}, \end{aligned}$$

by Theorem 4.26. Therefore, $C^2 = C$. Because C is a projection, $\text{rank}(C) = \text{tr}(C)$. Therefore,

$$\begin{aligned} s_\alpha &= \frac{\chi(e)}{o(G)} \sum_{\mu \in G_\alpha} \sum_{i=1}^r \chi(\sigma_i^{-1}\mu\sigma_i) \\ &= \frac{\chi(e)}{o(G_\alpha)} \sum_{\mu \in G_\alpha} \chi(\mu) \\ &= \chi(e)(\chi, 1)_{G_\alpha}. \end{aligned}$$

□

For each $\alpha \in \overline{\Delta}$, the matrix C defined by Equation (6.18) is not only idempotent but hermitian. It is a positive semidefinite matrix whose nonzero eigenvalues all equal 1. Using any of a number of matrix-theoretic approaches, one can find an s_α -by- s_α principal submatrix of C that is invertible. The row indices of such a submatrix comprise a subset F_α of $\{\alpha\sigma : \sigma \in G\}$ such that $\{e_\beta^* : \beta \in F_\alpha\}$ is a basis of the orbital subspace $\langle e_{\alpha\sigma}^* : \sigma \in G \rangle$. Putting these subsets together, we obtain a set

$$\hat{\Delta} = \bigcup_{\alpha \in \overline{\Delta}} F_\alpha \quad (6.19)$$

such that $\{e_\alpha^* : \alpha \in \hat{\Delta}\}$ is a basis of $V_\chi(G)$. Moreover, because C does not depend on the basis, $\mathcal{B} = \{e_i : 1 \leq i \leq n\}$, neither does $\hat{\Delta}$. (See, [Grone (1977a)], [Marcus & Chollet (1986)], [Merris (1978)], and [Merris & Pierce ((1973)] for variations on this theme.)

Suppose $\mathcal{B}$ is an orthogonal basis of V . If $\chi(e) = 1$, then (Corollary 6.32) $\{e_\alpha^* : \alpha \in \overline{\Delta}\}$ is an *orthogonal* basis of $V_\chi(G)$. If $\chi(e) > 1$, can the subsets F_α , $\alpha \in \overline{\Delta}$, be chosen in such a way that $\{e_\alpha^* : \alpha \in \hat{\Delta}\}$ is an orthogonal basis of $V_\chi(G)$? When $G = D_m$, the dihedral group of order $2m$, such a $\hat{\Delta}$ exists for every $\chi \in I(G)$ if and only if m is a power of 2. (See [Wang & Gong (1991a)] and [Holmes & Tam (1992)]). On the other hand, every doubly transitive subgroup G of S_m has an irreducible character for which no such $\hat{\Delta}$ exists [Holmes (1995)].

An explicit orthonormal basis of $\langle e_{\alpha\sigma}^* : \sigma \in G \rangle$ not comprised of decomposable symmetrized tensors was constructed in [Merris (1978)]. Other work on bases of orbital subspaces can be found in [Dias da Silva (1981)], [Kim, *et al.* (1987)], and [Wang & Gong (1991b)].

From a basis of $V_\chi(G)$ we can determine its dimension:

$$\begin{aligned} \dim(V_\chi(G)) &= o(\hat{\Delta}) \\ &= \sum_{\alpha \in \overline{\Delta}} s_\alpha \\ &= \chi(e) \sum_{\alpha \in \overline{\Delta}} (\chi, 1)_{G_\alpha}, \end{aligned} \quad (6.20)$$

a somewhat useless formula, at least in its present form. Another approach to computing the dimension takes advantage of the fact that $V_\chi(G)$ is the image of the projection $T(G, \chi)$. That is,

$$\begin{aligned} \dim(V_\chi(G)) &= \text{tr}(T(G, \chi)) \\ &= \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma) \text{tr}(P(\sigma)). \end{aligned} \quad (6.21)$$

Denote by η the character of the representation P of S_m . Since it is integer valued, $\eta = \bar{\eta}$, and Equation (6.21) can be written as

$$\dim(V_\chi(G)) = \chi(e)(\chi, \eta)_G. \quad (6.22)$$

Because $P(\sigma)$ permutes the elements of the induced basis,

$$\begin{aligned} \eta(\sigma) &= o(\{\alpha \in \Gamma_{m,n} : P(\sigma)(e_\alpha^\otimes) = e_\alpha^\otimes\}) \\ &= o(\{\alpha \in \Gamma_{m,n} : \alpha = \alpha\sigma\}) \\ &= o(\{\alpha \in \Gamma_{m,n} : \alpha(i) = \alpha(\sigma(i)), \quad 1 \leq i \leq m\}). \end{aligned}$$

Evidently, α is a fixed point of $P(\sigma)$ if and only if $\alpha(i) = \alpha(j)$ whenever $\sigma(i) = j$, if and only if α is constant on the cycles of σ . To count the number of fixed points, we may use the Fundamental Counting Principle. There are $c(\sigma)$ decisions to be made, namely the value of α on each cycle of σ . Because there are n choices for each decision, $\eta(\sigma) = n^{c(\sigma)}$. Therefore, from Equation (6.21),

$$\dim(V_\chi(G)) = \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma) n^{c(\sigma)}. \quad (6.23)$$

Another formula for the dimension of $V_\chi(G)$ will emerge from the representation theory of $GL(n, \mathbb{C})$. (See Corollary 8.19.)

When χ is the principal character of G ,

$$\dim(V_1(G)) = \frac{1}{o(G)} \sum_{\sigma \in G} n^{c(\sigma)}. \quad (6.24)$$

When $\chi = 1$, $\bar{\Delta} = \Delta$ and $\dim(V_1(G)) = o(\Delta)$. Therefore, Equation (6.24) can be viewed as a manifestation of Burnside's Lemma!

From Definition 6.18, functions $\alpha, \beta \in \Gamma_{m,n}$ are equivalent modulo G if and only if there is a $\sigma \in G$ such that $\alpha = \beta\sigma$, if and only if there is a $\bar{\sigma} \in \bar{G} = \{\bar{\sigma} : \sigma \in G\}$ such that $\bar{\sigma}(\alpha) = \beta$. Thus,

$$\begin{aligned} o(\Delta) &= \frac{1}{o(\bar{G})} \sum_{\bar{\sigma} \in \bar{G}} F(\bar{\sigma}) \\ &= \frac{1}{o(G)} \sum_{\sigma \in G} F(\bar{\sigma}), \end{aligned}$$

where $F(\bar{\sigma})$ is the number of fixed points of $\bar{\sigma}$. As we have just seen, $\eta(\sigma) = F(\bar{\sigma}) = n^{c(\sigma)}$. For future reference, we summarize this observation in the following.

THEOREM 6.35 *Let G be a subgroup of S_m . Let Δ be a system of distinct representatives for the equivalence classes of $\Gamma_{m,n}$ modulo G . Then*

$$o(\Delta) = \frac{1}{o(G)} \sum_{\sigma \in G} n^{c(\sigma)}.$$

The next result, an explicit description of Ω when $G = S_m$, requires the notion of a multiplicity partition.

DEFINITION 6.36 Let $\alpha \in \Gamma_{m,n}$ be fixed but arbitrary. For each $y \in \{1, 2, \dots, n\}$, $o(\alpha^{-1}(y)) = o(\{x : \alpha(x) = y\})$ is the **multiplicity** of y in α . Of these multiplicities, suppose m_1 is the largest, m_2 the next largest, and so on. If the image of α contains r integers altogether, then $m_1 \geq m_2 \geq \dots \geq m_r \geq 1$, and $m_1 + m_2 + \dots + m_r = m$. (We are not necessarily assuming that $m_i = o(\alpha^{-1}(i))$.) In particular, $\mu = \mu(\alpha) = [m_1, m_2, \dots, m_r]$ is a partition of m called the **multiplicity partition** of α .

Suppose $G = S_m$. Let μ be the multiplicity partition of a fixed but arbitrary $\alpha \in \Gamma_{m,n}$. Then G_α is conjugate to the Young Subgroup

$$S_\mu = S_{m_1} \times S_{m_2} \times \dots \times S_{m_r}.$$

THEOREM 6.37 *Suppose $G = S_m$. Let π be a fixed but arbitrary partition of m and take $\chi = \chi_\pi$. Let μ be the multiplicity partition of a fixed but arbitrary $\alpha \in \Gamma_{m,n}$. Then $e_\alpha^* \neq 0$ if and only if π majorizes μ .*

Proof Without loss of generality (Lemma 6.22), we may assume $G_\alpha = S_\mu$. Denoting the principal character of S_μ by 1_μ we have, from Equation (6.12), that

$$\begin{aligned} [S_m : S_\mu] \|e_\alpha^*\|^2 / \chi_\pi(e) &= (\chi_\pi, 1_\mu)_{S_\mu} \\ &= (1_\mu, \chi_\pi)_{S_\mu} \end{aligned}$$

because the left-hand side is real. By the Frobenius Reciprocity Theorem,

$$\begin{aligned} (1_\mu, \chi_\pi)_{S_\mu} &= (1_\mu^{S_m}, \chi_\pi)_m \\ &= K_{\pi, \mu}. \end{aligned}$$

By Theorem 4.55, the Kostka coefficient $K_{\pi, \mu} \neq 0$ if and only if $\pi \succ \mu$. □

COROLLARY 6.38 *Let V be a vector space of dimension n . Let $\pi = [\pi_1, \pi_2, \dots, \pi_k]$ be a partition of m . Suppose $\chi = \chi_\pi$. Then the symmetry class of tensors $V_\chi(S_m) \neq \{0\}$ if and only if $k \leq n$.*

Proof $V_\chi(S_m) \neq \{0\}$ if and only if $\Omega \neq \emptyset$ if and only if (Theorem 6.37) there is an $\alpha \in \Gamma_{m,n}$ whose multiplicity partition, $\mu = [m_1, m_2, \dots, m_r]$, is majorized by π . However, $\pi \succ \mu$ only if $k \leq r$. Because $r \leq n$, necessity is proved. Conversely, if $k \leq n$, there is an $\alpha \in \Gamma_{m,n}$ such that $\mu(\alpha) = \pi$. Because $\pi \succ \pi$, $\alpha \in \Omega$. $\square$

COROLLARY 6.39 Suppose $\{e_i: 1 \leq i \leq n\}$ is a basis of V . Let $G = S_m$ and $\chi = \chi_\pi$, where $\pi = [\pi_1, \pi_2, \dots, \pi_k] \vdash m$. Suppose $\alpha \in \Gamma_{m,n}$ has multiplicity partition $\mu = [m_1, m_2, \dots, m_r]$. If $m_1 > \pi_1$, then $e_\alpha^* = 0$.

Proof π cannot majorize a partition whose largest part is greater than π_1 . $\square$

J.A. Dias da Silva and Amélia Fonseca were able to obtain a generalization of Theorem 6.37 using the notion of a “multilinearity partition” (not to be confused with a multiplicity partition).

DEFINITION 6.40 Let G be a subgroup of S_m and suppose $\xi \in I(G)$. The **multilinearity partition**, $MP(\xi)$, is the least upper bound (with respect to majorization) of the partitions $\pi \vdash m$ for which $(\xi, \chi_\pi)_G \neq 0$.

EXAMPLE 6.41 Suppose H is the subgroup of $G = S_5$ generated by $\{(12345), (1325)\}$. Then $o(H) = 20$, and $K = \{e_5\}$ is the only subgroup of H that is normal in G . By Theorem 3.36, G is isomorphic to a transitive subgroup $\hat{G}$ of S_6 (arising from the action of G on the left cosets of H).⁶

Let ξ be the irreducible character of S_5 corresponding to the partition $[2, 1^3]$. Never mind that $\xi(\sigma) = \varepsilon(\sigma)(F(\sigma) - 1)$, $\sigma \in G = S_5$. From the perspective of $\hat{G} \subset S_6$, ξ is the irreducible character of degree 4 that takes the value $+1$ on 6-cycles. The character of S_6 induced by $\xi \in I(\hat{G})$ is

$$\xi^{S_6} = \chi_{[4, 1^2]} + \chi_{[3^2]} + \chi_{[2^2, 1^2]}.$$

By the Frobenius Reciprocity Theorem, $(\xi, \chi_\pi)_{\hat{G}} \neq 0$ if and only if $\pi \in \{[4, 1^2], [3^2], [2^2, 1^2]\}$. With respect to majorization, the least upper bound of these three partitions is $[4, 2]$. (See Figure 1.3.) Therefore, the multilinearity partition $MP(\xi) = [4, 2]$. In this particular example, it turns out that

$$(\xi, \chi_\nu)_{\hat{G}} = 0,$$

when $\nu = MP(\xi)$. $\square$

THEOREM 6.42 Let G be a subgroup of S_m and suppose $\xi \in I(G)$. Let μ be the multiplicity partition of $\alpha \in \Gamma_{m,n}$. If $e_\alpha^* = T(G, \xi)e_\alpha^\otimes \neq 0$, then $MP(\xi)$ majorizes μ .

A proof of Theorem 6.42 can be found in [Dias da Silva & Fonseca (1987)] and [Fonseca (1989)]. The following partial converse was obtained in [Dias da Silva & Fonseca (1990) & (1995)].

⁶It is $\hat{G}$ that accounts for the outer automorphisms of S_6 . (See, e.g., [Passman (1968), pp. 35–36].)

THEOREM 6.43 *Suppose ξ is an irreducible character of the subgroup G of S_m . Let χ_ν be the irreducible character of S_m corresponding to the partition $\nu = MP(\xi)$. Suppose $MP(\xi)$ majorizes the multiplicity partition of $\alpha \in \Gamma_{m,n}$. If $(\xi, \chi_\nu)_G \neq 0$, then there exists a permutation $\sigma \in S_m$ such that $e_{\alpha\sigma}^* = T(G, \xi)e_{\alpha\sigma}^\otimes \neq 0$.*

We turn now to another generalization of Theorem 6.37.

DEFINITION 6.44 Suppose V is a vector space. The vectors $v_1, v_2, \dots, v_m \in V$ are said to **conform** to the partition $\pi \vdash m$ if it is possible to distribute the m vectors among the m boxes of the Ferrers diagram $F(\pi)$ so that the vectors in each column are linearly independent.

If one of $v_1, v_2, \dots, v_m$ is zero, then the vectors conform to no partition of m ; if $v_1 = v_2$, then any conforming distribution of vectors to boxes must place v_1 and v_2 in different columns.

EXAMPLE 6.45 Let $\{e_1, e_2\}$ be a linearly independent set in V . Suppose $v_1 = v_2 = e_1$, $v_3 = v_4 = e_2$, $v_5 = e_1 + e_2$, and $v_6 = e_1 - e_2$. Because $\{v_1, v_3\}$, $\{v_2, v_4\}$, and $\{v_5, v_6\}$ are all linearly independent, $v_1, v_2, \dots, v_6$ conform to $\pi = [3, 3]$. If $\rho \vdash 6$ and ρ majorizes $[3, 3]$, then $v_1, v_2, \dots, v_6$ also conform to ρ . $\square$

EXAMPLE 6.46 Suppose $B = \{e_1, e_2, \dots, e_n\}$ is a basis of the vector space V . Let $\mu = [m_1, m_2, \dots, m_r]$ be the multiplicity partition afforded by some fixed but arbitrary $\alpha \in \Gamma_{m,n}$. Then the vectors $e_{\alpha(1)}, e_{\alpha(2)}, \dots, e_{\alpha(m)}$ conform to π if and only if π majorizes μ . $\square$

In view of Example 6.46, Theorem 6.37 can be restated as follows: Suppose $\pi \vdash m$. Let $G = S_m$ and $\chi = \chi_\pi$. Then $e_\alpha^* \neq 0$ if and only if the vectors $e_{\alpha(1)}, e_{\alpha(2)}, \dots, e_{\alpha(m)}$ conform to π .

GAMAS'S THEOREM 6.47 *Suppose $\pi \vdash m$. Let $G = S_m$ and $\chi = \chi_\pi$. Then the decomposable symmetrized tensor $v_1 * v_2 * \dots * v_m \neq 0$ if and only if the vectors $v_1, v_2, \dots, v_m$ conform to π .*

The original version of Theorem 6.47 appeared in [Gamas (1988)]. Our approach has followed [Pate (1990)]. Observe that Gamas's Theorem also generalizes Lemma 6.16 because a linearly independent set of m vectors conforms to every partition of m . A common generalization of Theorems 6.42, 6.43, and 6.47 appeared in [Dias da Silva & Fonseca (1990)]:

THEOREM 6.48 *Let ξ be an irreducible character of the subgroup G of S_m . Let χ_ν be the irreducible character of S_m corresponding to the multilinearity partition $\nu = MP(\xi)$. Suppose $(\xi, \chi_\nu)_G \neq 0$. Then there exists a permutation $\sigma \in S_m$ such that*

$$T(G, \xi)(v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \cdots \otimes v_{\sigma(m)}) = v_{\sigma(1)} * v_{\sigma(2)} * \cdots * v_{\sigma(m)} \neq 0,$$

if and only if the vectors $v_1, v_2, \dots, v_m$ conform to ν .

Suppose $v_1, v_2, \dots, v_m$ are nonzero vectors in some vector space V . It was proved in [Dias da Silva (1990)] (also see [Dias da Silva (1996)]) that, with respect to majorization, there is a unique minimum partition to which $v_1, v_2, \dots, v_m$ conform.

DEFINITION 6.49 Let $R = \{v_1, v_2, \dots, v_m\}$ be a multiset of nonzero vectors in V . The minimum partition to which $v_1, v_2, \dots, v_m$ conform is the **nullity partition** of R , denoted $\eta(R)$. The **rank partition** of R is $\rho(R) = \eta(R)^*$, the conjugate of its nullity partition.

In the language of Definition 6.49, Theorem 6.48 can be restated as follows:

THEOREM 6.50 Let ξ be an irreducible character of the subgroup G of S_m . Let χ_ν the irreducible character of S_m corresponding to the partition $\nu = MP(\xi)$. Suppose $(\xi, \chi_\nu)_G \neq 0$. Let $R = \{v_1, v_2, \dots, v_m\}$ be a multiset of nonzero vectors in V . Then there exists a permutation $\sigma \in S_m$ such that

$$T(G, \xi)(v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \cdots \otimes v_{\sigma(m)}) = v_{\sigma(1)} * v_{\sigma(2)} * \cdots * v_{\sigma(m)} \neq 0,$$

if and only if $\nu \succ \eta(R)$.

Among the most important symmetry classes is $V^{\otimes m}$ itself, occurring when $G = \{e\}$ and $\chi = 1$. We proceed to discuss another.

DEFINITION 6.51 The symmetry class $V_\varepsilon(S_m)$, corresponding to the symmetric group and its alternating character, is called the **space of skew-symmetric tensors**, the m -th **Grassmann space**, or the m -th **exterior power** of V . The special notation $\wedge^m V$ is used for $V_\varepsilon(S_m)$, and the decomposable symmetrized tensors are written $v_1 \wedge v_2 \wedge \cdots \wedge v_m$.

LEMMA 6.52 Let V be a vector space. Then for all $\sigma \in S_m$ and all $v_1, v_2, \dots, v_m \in V$,

$$v_{\sigma(1)} \wedge v_{\sigma(2)} \wedge \cdots \wedge v_{\sigma(m)} = \varepsilon(\sigma) v_1 \wedge v_2 \wedge \cdots \wedge v_m. \quad (6.25)$$

In particular, if $v_i = v_j$ for some pair $i \neq j$, then $v_1 \wedge v_2 \wedge \cdots \wedge v_m = 0$.

If $\{e_i : 1 \leq i \leq n\}$ is a basis of V , then (Theorem 6.30 and Corollary 6.32) $\{e_\alpha^\wedge : \alpha \in Q_{m,n}\}$ is a basis of $\wedge^m V$, where $e_\alpha^\wedge = e_{\alpha(1)} \wedge e_{\alpha(2)} \wedge \cdots \wedge e_{\alpha(m)}$. In particular,

$$\dim(\wedge^m V) = C(n, m). \quad (6.26)$$

THEOREM 6.53 Suppose $v_1, v_2, \dots, v_m \in V$. Then $v_1 \wedge v_2 \wedge \cdots \wedge v_m \neq 0$ if and only if $\{v_1, v_2, \dots, v_m\}$ is linearly independent.

Proof Recall that $\varepsilon = \chi_{[1^m]}$. Because $v_1, v_2, \dots, v_m$ conform to $[1^m]$ if and only if $\{v_1, v_2, \dots, v_m\}$ is linearly independent, the result is an immediate consequence of Gamas's Theorem. $\square$

Theorem 6.53 shows that the converse of Lemma 6.16 is valid in $\wedge^m V$. The next result may be viewed as a partial converse of Lemma 6.15.

THEOREM 6.54 Let $v_1, v_2, \dots, v_m \in V$. If

$$w_j = \sum_{i=1}^m a_{ij} v_i, \quad 1 \leq j \leq m,$$

then $w_1 \wedge w_2 \wedge \cdots \wedge w_m = \det(a_{ij}) v_1 \wedge v_2 \wedge \cdots \wedge v_m$.

Proof Observe that

$$\begin{aligned} w_1 \wedge w_2 \wedge \cdots \wedge w_m &= \left(\sum_{i=1}^m a_{i1} v_i \right) \wedge \left(\sum_{i=1}^m a_{i2} v_i \right) \wedge \cdots \wedge \left(\sum_{i=1}^m a_{im} v_i \right) \\ &= \sum_{\alpha \in \Gamma_{m,m}} \left(\prod_{t=1}^m a_{\alpha(t)t} \right) v_{\alpha(1)} \wedge v_{\alpha(2)} \wedge \cdots \wedge v_{\alpha(m)} \\ &= \sum_{\sigma \in S_m} \left(\prod_{t=1}^m a_{\sigma(t)t} \right) v_{\sigma(1)} \wedge v_{\sigma(2)} \wedge \cdots \wedge v_{\sigma(m)} \\ &= \left(\sum_{\sigma \in S_m} \varepsilon(\sigma) \prod_{t=1}^m a_{\sigma(t)t} \right) v_1 \wedge v_2 \wedge \cdots \wedge v_m, \end{aligned}$$

by multilinearity, Theorem 6.53, and Lemma 6.52. $\square$

The proof of Theorem 6.54 depended on the following fact: If $A = (a_{ij})$ is an m -by- m matrix, then

$$\det(A) = \sum_{\sigma \in S_m} \varepsilon(\sigma) \prod_{t=1}^m a_{t\sigma(t)}. \quad (6.27)$$

Taking Equation (6.27) as the definition of determinant, the usual expansion theorem can be derived using multilinear techniques: Suppose $\dim(V) = m$. Let $u \in V$. Define

$$\Phi: \overbrace{V \times V \times \cdots \times V}^{(m-1)\text{-times}} \rightarrow \wedge^m V$$

by $\Phi(v_1, v_2, \dots, v_{m-1}) = v_1 \wedge v_2 \wedge \cdots \wedge v_{m-1} \wedge u$. Then Φ is $(m-1)$ -linear and symmetric with respect to S_{m-1} and ε . By Theorem 6.14, there is (unique) $h_\varepsilon \in L(\wedge^{m-1} V, \wedge^m V)$ such that

$$h_\varepsilon(v_1 \wedge v_2 \wedge \cdots \wedge v_{m-1}) = v_1 \wedge v_2 \wedge \cdots \wedge v_{m-1} \wedge u,$$

for all $v_1, v_2, \dots, v_{m-1} \in V$. Suppose $\{e_1, e_2, \dots, e_m\}$ is a basis of V . Let

$$v_j = \sum_{i=1}^m a_{ij} e_i, \quad 1 \leq j < m, \quad \text{and} \quad u = \sum_{i=1}^m a_{im} e_i.$$

Then, by Theorem 6.54,

$$v_1 \wedge v_2 \wedge \cdots \wedge v_{m-1} \wedge u = \det(A) e_1 \wedge e_2 \wedge \cdots \wedge e_m. \quad (6.28)$$

On the other hand,

$$\begin{aligned} v_1 \wedge v_2 \wedge \cdots \wedge v_{m-1} &= \sum_{\alpha \in \Gamma_{m-1, m}} \left(\prod_{t=1}^{m-1} a_{\alpha(t)t} \right) e_\alpha^\wedge \\ &= \sum_{\alpha \in Q_{m-1, m}} \sum_{\sigma \in S_{m-1}} \left(\prod_{t=1}^{m-1} a_{\alpha\sigma(t)t} \right) e_{\alpha\sigma}^\wedge \\ &= \sum_{\alpha \in Q_{m-1, m}} \left(\sum_{\sigma \in S_{m-1}} \varepsilon(\sigma) \prod_{t=1}^{m-1} a_{\alpha(t)\sigma(t)} \right) e_\alpha^\wedge. \end{aligned} \quad (6.29)$$

DEFINITION 6.55 Suppose $A \in \mathbb{C}_{m, m}$. If $i, j \in \{1, 2, \dots, m\}$, let $A(i|j)$ be the submatrix of A obtained by deleting its i -th row and j -th column. If $\alpha \in Q_{r, m}$, denote by $\alpha^c \in Q_{m-r, m}$ the (increasing) sequence **complementary** to α .

If, for example, $\alpha = (2, 4) \in Q_{2, 5}$, then $\alpha^c = (1, 3, 5) \in Q_{3, 5}$. If $\alpha = (1, 2, 3, 5) \in Q_{4, 5}$, then $\alpha^c = (4) \in Q_{1, 5}$. (We will feel free to abuse the notation by writing $\alpha^c = 4$ in the latter case.)

In the notation of Definition 6.55, Equation (6.29) becomes

$$v_1 \wedge v_2 \wedge \cdots \wedge v_{m-1} = \sum_{\alpha \in Q_{m-1, m}} \det A(\alpha^c | m) e_\alpha^\wedge.$$

Applying the linear transformation h_ε to this equation produces

$$\begin{aligned} h_\varepsilon(v_1 \wedge v_2 \wedge \cdots \wedge v_{m-1}) &= \sum_{\alpha \in Q_{m-1,m}} \det A(\alpha^c | m) h_\varepsilon(e_\alpha^\wedge) \\ &= \sum_{\alpha \in Q_{m-1,m}} \det A(\alpha^c | m) e_{\alpha(1)} \wedge e_{\alpha(2)} \wedge \cdots \wedge e_{\alpha(m-1)} \wedge u. \end{aligned} \quad (6.30)$$

By multilinearity,

$$e_{\alpha(1)} \wedge e_{\alpha(2)} \wedge \cdots \wedge e_{\alpha(m-1)} \wedge u = \sum_{i=1}^m a_{im} e_{\alpha(1)} \wedge e_{\alpha(2)} \wedge \cdots \wedge e_{\alpha(m-1)} \wedge e_i,$$

only one term of which survives, namely $i = \alpha^c$. Because $i = \alpha^c$ if and only if $\alpha = (1, 2, \dots, i-1, i+1, \dots, m)$ we have, from Lemma 6.52, that

$$e_{\alpha(1)} \wedge e_{\alpha(2)} \wedge \cdots \wedge e_{\alpha(m-1)} \wedge e_i = (-1)^{m-i} e_1 \wedge e_2 \wedge \cdots \wedge e_m.$$

Therefore, from Equation (6.30),

$$v_1 \wedge v_2 \wedge \cdots \wedge v_{m-1} \wedge u = \sum_{i=1}^m (-1)^{i+m} a_{im} \det A(i|m) e_1 \wedge e_2 \wedge \cdots \wedge e_m.$$

Comparing with Equation (6.28) (and using the fact that $e_1 \wedge e_2 \wedge \cdots \wedge e_m \neq 0$), we obtain

$$\det(A) = \sum_{i=1}^m (-1)^{i+m} a_{im} \det A(i|m), \quad (6.31)$$

the familiar expansion for $\det(A)$ along the last column.

In order to state the next result, we need the following extension of Definition 6.55.

DEFINITION 6.56 Suppose $A = (a_{ij}) \in \mathbb{C}_{m,n}$. If $\alpha \in \Gamma_{p,m}$ and $\beta \in \Gamma_{q,n}$, then $A[\alpha|\beta] \in \mathbb{C}_{p,q}$ is the matrix whose (i, j) -entry is the $(\alpha(i), \beta(j))$ -entry of A . If $\alpha \in Q_{p,m}$ and $\beta \in Q_{q,n}$, then $A(\alpha|\beta) = A[\alpha^c|\beta^c]$ is the matrix obtained by deleting from A the rows whose indices appear in α and the columns whose indices appear in β .

LAPLACE EXPANSION THEOREM 6.57 Suppose $1 \leq p < n$. Let $\alpha \in Q_{p,n}$. If $A \in \mathbb{C}_{n,n}$, then

$$\det(A) = \sum_{\beta \in Q_{p,n}} (-1)^{r(\alpha)+r(\beta)} \det(A[\alpha|\beta]) \det(A(\alpha|\beta)), \quad (6.32)$$

where $r(\alpha) = \alpha(1) + \alpha(2) + \cdots + \alpha(p)$.

While it is understandably more complicated, the proof of Theorem 6.57 is analogous to the derivation of Equation (6.31).⁷

⁷There is a comparable Laplace Expansion Theorem for permanents (see, e.g., [Marcus & Minc (1964)]). Generalizations to other matrix functions appear in [Marcus & Soules (1967)].

EXAMPLE 6.58 Let's expand the determinant of

$$A = \begin{pmatrix} 3 & 1 & 0 & 5 \\ 4 & 3 & 1 & 9 \\ 0 & 2 & 2 & 3 \\ 2 & 4 & 2 & 7 \end{pmatrix}$$

along rows 1 and 3, that is, let $\alpha = (1, 3)$. Summing over $\beta \in Q_{2,4} = \{(1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (3, 4)\}$, Equation (6.32) becomes

$$\begin{aligned} \det(A) &= \sum_{\beta \in Q_{2,4}} (-1)^{r(\beta)} \det(A[1, 3|\beta]) \det(A(1, 3|\beta)) \\ &= (-1)^{1+2} \det(A[1, 3|1, 2]) \det(A(1, 3|1, 2)) \\ &\quad + (-1)^{1+3} \det(A[1, 3|1, 3]) \det(A(1, 3|1, 3)) \\ &\quad + (-1)^{1+4} \det(A[1, 3|1, 4]) \det(A(1, 3|1, 4)) + \dots \\ &= -\det \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix} \det \begin{pmatrix} 1 & 9 \\ 2 & 7 \end{pmatrix} + \det \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \det \begin{pmatrix} 3 & 9 \\ 4 & 7 \end{pmatrix} \\ &= -\det \begin{pmatrix} 3 & 5 \\ 0 & 3 \end{pmatrix} \det \begin{pmatrix} 3 & 1 \\ 4 & 2 \end{pmatrix} - \det \begin{pmatrix} 1 & 0 \\ 2 & 2 \end{pmatrix} \det \begin{pmatrix} 4 & 9 \\ 2 & 7 \end{pmatrix} \\ &= +\det \begin{pmatrix} 1 & 5 \\ 2 & 3 \end{pmatrix} \det \begin{pmatrix} 4 & 1 \\ 2 & 2 \end{pmatrix} - \det \begin{pmatrix} 0 & 5 \\ 2 & 3 \end{pmatrix} \det \begin{pmatrix} 4 & 3 \\ 2 & 4 \end{pmatrix} \\ &= -6 \times (-11) + 6 \times (-15) - 9 \times 2 - 2 \times 10 \\ &\quad + (-7) \times 6 - (-10) \times (10) \\ &= -4. \end{aligned}$$

□

Another important family of examples occurs when G is arbitrary but $\chi = 1$ is the principal character. In this case

$$\begin{aligned} P(\tau^{-1})(v_1 * v_2 * \dots * v_m) &= v_{\tau(1)} * v_{\tau(2)} * \dots * v_{\tau(m)} \\ &= v_1 * v_2 * \dots * v_m, \end{aligned} \quad (6.33)$$

for all $\tau \in G$. (See Equations (6.6)–(6.7).) Moreover, $\bar{\Delta} = \Delta$, and, from Equation (6.24),

$$\dim(V_1(G)) = \frac{1}{o(G)} \sum_{\sigma \in G} n^{c(\sigma)}.$$

LEMMA 6.59 Let G be a subgroup of S_m . If $v \in V$, then

$$\overbrace{v \otimes v \otimes \cdots \otimes v}^{m\text{-times}} \in V_1(G).$$

THEOREM 6.60⁸ Let $\xi = 1$ be the principal character of the subgroup G of S_m . Suppose V is an inner product space. In the symmetry class of tensors $V_1(G)$,

- (i) $v_1 * v_2 * \cdots * v_m = 0$ if and only if $v_i = 0$ for some i ;
- (ii) if $v_1 * v_2 * \cdots * v_m = w_1 * w_2 * \cdots * w_m \neq 0$, then there exists a permutation $\sigma \in S_m$ and constants $d_i \in \mathbb{C}$ such that $v_i = d_i w_{\sigma(i)}$, $1 \leq i \leq m$, and the product $d_1 \times d_2 \times \cdots \times d_m = 1$.

Proof If $G = \{e\}$ part (i) reduces to Theorem 5.15; if $G = S_m$, it is an easy consequence of Gamas's Theorem because $\xi = \chi_{[m]}$, and any collection of m nonzero vectors conforms to $[m]$. So, we may assume G is a proper subgroup of S_m . (If G were normal in S_m , part (i) would follow from Theorem 6.50 and the invertibility of $P(\sigma)$.)

If $v_i = 0$ for some i , then $v_1 * v_2 * \cdots * v_m = 0$ by multilinearity. Conversely, suppose $v_1 * v_2 * \cdots * v_m = 0$. Let $u \in V$ be fixed but arbitrary. Then

$$\begin{aligned} 0 &= (u \otimes u \otimes \cdots \otimes u, v_1 * v_2 * \cdots * v_m) \\ &= (u \otimes u \otimes \cdots \otimes u, T(G, 1)v_1 \otimes v_2 \otimes \cdots \otimes v_m) \\ &= (T(G, 1)u \otimes u \otimes \cdots \otimes u, v_1 \otimes v_2 \otimes \cdots \otimes v_m) \\ &= (u \otimes u \otimes \cdots \otimes u, v_1 \otimes v_2 \otimes \cdots \otimes v_m) \\ &= \prod_{i=1}^m (u, v_i). \end{aligned} \tag{6.34}$$

If $\{e_1, e_2, \dots, e_n\}$ is an orthonormal basis of V , there exist constants a_{ij} such that

$$v_j = \sum_{i=1}^n a_{ij} e_i, \quad 1 \leq j \leq m.$$

Let $u = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$ where, for the moment, we view the coefficients,

⁸Theorem 6.60 first appeared in [Marcus & Newman (1962)].

$x_1, x_2, \dots, x_n$, as independent indeterminates. Then

$$\begin{aligned} \prod_{t=1}^m (u, v_t) &= \prod_{t=1}^m \left(u, \sum_{i=1}^n a_{it} e_i \right) \\ &= \prod_{t=1}^m \left(\sum_{i=1}^n \bar{a}_{it} (u, e_i) \right) \\ &= \prod_{t=1}^m \left(\sum_{i=1}^n \bar{a}_{it} x_i \right). \end{aligned} \quad (6.35)$$

If $v_t \neq 0$ for all t , the right-hand side of Equation (6.35) is a polynomial in n variables (the x_i) which is not identically zero. On the other hand, according to Equation (6.34), any substitution for the x_i produces a value of 0. This contradiction establishes that $v_t = 0$ for some t , and part (i) is proved.

If $\dim(V) = 1$, part (ii) is immediate. Thus, we may assume $\dim(V) > 1$. Because $v_1 * v_2 * \dots * v_m = w_1 * w_2 * \dots * w_m$ we have, as in the derivation of Equation (6.34), that

$$\prod_{t=1}^m (u, w_t) = \prod_{t=1}^m (u, v_t). \quad (6.36)$$

For each i , let $v_i = y_i + z_i$, where $y_i \in \langle w_1 \rangle$ and $z_i \in \langle w_1 \rangle^\perp$. It follows from Equation (6.36) that, for all $u \in \langle w_1 \rangle^\perp$,

$$\begin{aligned} 0 &= \prod_{t=1}^m (u, v_t) \\ &= \prod_{t=1}^m (u, z_t). \end{aligned}$$

As in the proof of part (i), $z_k = 0$ for some k , which means that $v_k = y_k \in \langle w_1 \rangle$. In other words, $v_k = d_1 w_1$ for some $d_1 \in \mathbb{C}$. Substituting in Equation (6.36), we obtain

$$0 = (u, w_1) \left(\prod_{t=2}^m (u, w_t) - d_1 \prod_{t \neq k} (u, v_t) \right), \quad (6.37)$$

for all $u \in V$. Because $w_1 \neq 0$, a modification of the previous argument shows that the second factor on the right-hand side of Equation (6.37) is zero, and the rest of the proof follows by induction. $\square$

The next example, taken from [Marcus (1973)], shows that the permutation σ in Theorem 6.60, part (ii) need not belong to G .

EXAMPLE 6.61 Let $G = A_3$ and $\chi = 1$. Suppose $\{e_1, e_2\}$ is a basis of V . Let $w_1 = e_1$, $w_2 = e_2$, and $w_3 = e_1 + e_2$. Then

$$\begin{aligned} w_1 * w_2 * w_3 &= e_1 * e_2 * (e_1 + e_2) \\ &= e_1 * e_2 * e_1 + e_1 * e_2 * e_2. \end{aligned}$$

Because $\tau = (123) \in G$, it follows from Equation (6.33) that

$$\begin{aligned} P(\tau^{-1})(v_1 * v_2 * v_3) &= v_2 * v_3 * v_1 \\ &= v_1 * v_2 * v_3 \end{aligned}$$

for any three vectors v_1, v_2 , and $v_3 \in V$. In particular, letting $v_1 = e_1$ and $v_2 = v_3 = e_2$, we obtain $e_2 * e_2 * e_1 = e_1 * e_2 * e_2$. Hence,

$$\begin{aligned} w_1 * w_2 * w_3 &= e_1 * e_2 * e_1 + e_2 * e_2 * e_1 \\ &= (e_1 + e_2) * e_2 * e_1 \\ &= w_3 * w_2 * w_1 \\ &= w_{\sigma(1)} * w_{\sigma(2)} * w_{\sigma(3)}, \end{aligned}$$

where $\sigma = (13) \notin G$. □

Next we consider, in some detail, the symmetry class of tensors $V_1(S_m)$ corresponding to $G = S_m$ and $\chi = 1$.

DEFINITION 6.62 Let V be a vector space. The symmetry class $V_1(S_m)$ corresponding to the principal character of the symmetric group is the space of **completely symmetric tensors**. The decomposable symmetrized tensors in $V_1(S_m)$ are written $v_1 \bullet v_2 \bullet \cdots \bullet v_m$.

It follows from Equation (6.33) that

$$\begin{aligned} P(\tau^{-1})(v_1 \bullet v_2 \bullet \cdots \bullet v_m) &= v_{\tau(1)} \bullet v_{\tau(2)} \bullet \cdots \bullet v_{\tau(m)} \\ &= v_1 \bullet v_2 \bullet \cdots \bullet v_m, \end{aligned}$$

for all $\tau \in S_m$. This “complete” symmetry characterizes $V_1(S_m)$.

THEOREM 6.63 Let $w \in V^{m\otimes}$. Then $w \in V_1(S_m)$ if and only if $P(\tau)(w) = w$ for all $\tau \in S_m$.

Proof If $P(\tau)(w) = w$ for all $\tau \in S_m$, then

$$\begin{aligned} w &= \frac{1}{m!} \sum_{\tau \in S_m} P(\tau)(w) \\ &= T(S_m, 1)(w) \in V_1(S_m). \end{aligned}$$

Conversely, if $w \in V_1(S_m)$, then

$$\begin{aligned} P(\tau)(w) &= P(\tau)(T(S_m, 1)(w)) \\ &= (P(\tau)T(S_m, 1))(w) \\ &= T(S_m, 1)(w) \\ &= w. \end{aligned}$$

□

The next theorem is merely a specialization of previous results to the space of completely symmetric tensors.

THEOREM 6.64 *Let $\{e_1, e_2, \dots, e_n\}$ be a basis of the vector space V . Then $\{e_\alpha^\bullet = e_{\alpha(1)} \bullet e_{\alpha(2)} \bullet \dots \bullet e_{\alpha(m)} : \alpha \in G_{m,n}\}$ is a basis of $V_1(S_m)$. In particular, $\dim(V_1(S_m)) = C(n + m - 1, m)$.*

Suppose $\{x_1, x_2, \dots, x_n\}$ is a set of independent indeterminates over the complex numbers. Denote by $C_m[x_1, x_2, \dots, x_n]$ the subset of $C[x_1, x_2, \dots, x_n]$ consisting of the zero polynomial together with all homogeneous polynomials of (total) degree m in the n variables. Then $C_m[x_1, x_2, \dots, x_n]$ is a vector space over C .

LEMMA 6.65 *If $x_1, x_2, \dots, x_n$ are independent indeterminates over the complex numbers, then*

$$B = \left\{ \prod_{i=1}^m x_{\alpha(i)} : \alpha \in G_{m,n} \right\} \quad (6.38)$$

is a basis of $C_m[x_1, x_2, \dots, x_n]$.

The following definition will be useful in the proof of Lemma 6.65.

DEFINITION 6.66 If $\alpha \in \Gamma_{m,n}$, denote by $m_t(\alpha) = o(\alpha^{-1}(t))$ the multiplicity of t in the sequence α .

The nonzero integers in the multiset $\{m_t(\alpha) : 1 \leq t \leq n\}$ are the parts of the multiplicity partition $\mu(\alpha)$. If $\alpha = (2, 1, 2, 1, 2) \in \Gamma_{5,3}$, for example, then $m_1(\alpha) = 2$, $m_2(\alpha) = 3$, and $m_3(\alpha) = 0$. If $\beta = (2, 1, 1, 2, 2)$, then

$m_t(\beta) = m_t(\alpha)$, $1 \leq t \leq 3$. This sharing of multiplicities cannot happen for two different sequences of $G_{m,n}$. If $\alpha, \beta \in G_{m,n}$, then $\alpha = \beta$ if and only if $m_t(\alpha) = m_t(\beta)$, $1 \leq t \leq m$.⁹

Proof (of Lemma 6.65). Suppose

$$\sum_{\alpha \in G_{m,n}} c_\alpha \prod_{t=1}^m x_{\alpha(t)} = 0. \quad (6.39)$$

Let $\beta \in G_{m,n}$ be fixed but arbitrary. Take the partial derivative of both sides of Equation (6.39) with respect to x_1 , a total of $m_1(\beta)$ times. Then take partial derivatives with respect to x_2 , a total of $m_2(\beta)$ times, and so on. Finally take partial derivatives with respect to x_n , a total of $m_n(\beta)$ times. The result is

$$\left(\prod_{t=1}^n m_t(\beta)! \right) c_\beta = 0,$$

that is, $c_\beta = 0$. Therefore, $\mathcal{B}$ is linearly independent. The proof that $\mathcal{B}$ spans $\mathbb{C}_m[x_1, x_2, \dots, x_n]$ is left to the exercises. $\square$

LEMMA 6.67 *Let V be a vector space of dimension n . Suppose $x_1, x_2, \dots, x_n$ are independent indeterminates over $\mathbb{C}$. Then the vector space $V_1(S_m)$ is isomorphic to $\mathbb{C}_m[x_1, x_2, \dots, x_n]$.*

Proof Let $\{e_1, e_2, \dots, e_n\}$ be a basis of V . Define $T: V_1(S_m) \rightarrow \mathbb{C}_m[x_1, x_2, \dots, x_n]$ by $T(e_\alpha^*) = \prod x_{\alpha(t)}$, $\alpha \in G_{m,n}$, and linear extension. By Theorem 6.64 and Lemma 6.65, T is a vector space isomorphism. $\square$

Completely ignored in Lemma 6.67, of course, is the "additional structure" that makes $V_1(S_m)$ more than a vector space.¹⁰

THEOREM 6.68¹¹ *Let $\{e_1, e_2, \dots, e_n\}$ be a basis of V . For each $\alpha \in G_{m,n}$, define a vector*

$$y(\alpha) = \sum_{t=1}^n m_t(\alpha) e_t \in V.$$

Let $z_\alpha = y(\alpha) \otimes y(\alpha) \otimes \dots \otimes y(\alpha) \in V_1(S_m)$. Then $\{z_\alpha: \alpha \in G_{m,n}\}$ is a basis of $V_1(S_m)$. In particular, $V_1(S_m)$ is spanned by tensors of the form $v \otimes v \otimes \dots \otimes v (v = v \bullet v \bullet \dots \bullet v)$.

⁹If $\alpha = (1, 1, 1, 2, 2)$ and $\beta = (1, 1, 3, 3, 3)$, then $\mu(\alpha) = \mu(\beta)$, but $m_t(\alpha) \neq m_t(\beta)$, $1 \leq t \leq 3$.

¹⁰See, e.g., [Grone (1977b)] for a discussion of this issue.

¹¹Initially a conjecture of S. Pierce, Theorem 6.68 was proved in [Marcus & Gordon (1970)]. Our proof is taken from [Marcus (1973)].

Proof Let $T: V_1(S_m) \rightarrow \mathbb{C}_m[x_1, x_2, \dots, x_n]$ be the isomorphism defined by $T(e_\alpha^\circ) = \prod x_{\alpha(i)}$, $\alpha \in G_{m,n}$, and linear extension. Then

$$\begin{aligned} T(z_\alpha) &= T(y(\alpha) \bullet y(\alpha) \bullet \cdots \bullet y(\alpha)) \\ &= \left(\sum_{i=1}^n m_i(\alpha) x_i \right)^m. \end{aligned}$$

Because there are the right number of them, it remains to show that the polynomials

$$\left(\sum_{i=1}^n m_i(\alpha) x_i \right)^m, \quad \alpha \in G_{m,n},$$

are linearly independent. Let

$$R = \sum c(a_1, a_2, \dots, a_n) \left(\sum_{i=1}^n a_i x_i \right)^m, \quad (6.40)$$

where the first summation is over all $C(m+n-1, m)$ nonnegative integer solutions to $a_1 + a_2 + \cdots + a_n = m$. Setting $R = 0$, we will prove that the coefficients, $c(a_1, a_2, \dots, a_n)$, are all zero by induction on $m+n$.

If $n = 1$, then $0 = R = c(m)(mx_1)^m$ implies $c(m) = 0$. Thus, we may assume $n > 1$ and $m+n = k > 2$. Let $R = R_0 + R_1$, where R_0 is the sum of those terms in Equation (6.40) for which $a_n = 0$, and R_1 is the sum of the terms with $a_n \geq 1$. Of course, $\partial R_0 / \partial x_n = 0$. Because $R = 0$, it follows that

$$\begin{aligned} 0 &= \frac{\partial R_1}{\partial x_n} \\ &= m \sum_{\substack{a_1 + \cdots + a_n = m \\ a_i \geq 0, a_n \geq 1}} a_n c(a_1, a_2, \dots, a_n) \left(\sum_{i=1}^n a_i x_i \right)^{m-1} \\ &= m \sum_{\substack{a_1 + \cdots + a_n = m-1 \\ a_i \geq 0}} (a_n + 1) c(a_1, \dots, a_{n-1}, a_n + 1) (a_1 x_1 + \cdots + a_n x_n + x_n)^{m-1}, \end{aligned}$$

where the last step is obtained by replacing a_n with $a_n + 1$ and summing over $a_n \geq 0$. Next, we are going to substitute $x_i - x_n/m$ for x_i , $1 \leq i \leq n$. Because the summation is now over $a_1 + a_2 + \cdots + a_n = m-1$,

$$\begin{aligned} &a_1(x_1 - x_n/m) + \cdots + a_n(x_n - x_n/m) + (x_n - x_n/m) \\ &= a_1 x_1 + \cdots + a_n x_n - x_n(a_1 + \cdots + a_n)/m + x_n(m-1)/m \\ &= a_1 x_1 + \cdots + a_n x_n. \end{aligned}$$

This second change of variables results in the identity

$$0 = m \sum_{\substack{a_1 + \dots + a_n = m-1 \\ a_i \geq 0}} (a_n + 1) c(a_1, \dots, a_{n-1}, a_n + 1) (a_1 x_1 + \dots + a_n x_n)^{m-1}.$$

It follows from the induction hypothesis that

$$c(a_1, \dots, a_{n-1}, a_n + 1) = 0,$$

whenever $a_i \geq 0$, and $a_1 + \dots + a_n = m - 1$, that is,

$$c(a_1, \dots, a_{n-1}, a_n) = 0$$

whenever $a_1 + \dots + a_n = m$ and $a_n \geq 1$.

Setting $0 = R = R_0 + R_1$, we established that each of the coefficients appearing in the summation comprising R_1 is zero. At this point we can apply the induction hypothesis to

$$\begin{aligned} 0 &= R_0 \\ &= m \sum_{\substack{a_1 + \dots + a_{n-1} = m \\ a_i \geq 0}} c(a_1, a_2, \dots, a_{n-1}) \left(\sum_{t=1}^{n-1} a_t x_t \right)^m \end{aligned}$$

and conclude that the remaining coefficients of R are zero. $\square$

EXAMPLE 6.69 It follows from Theorem 6.68 that $V_1(S_2)$ is spanned by the decomposable tensors $v \otimes v$, $v \in V$. In particular, if $\{e_1, e_2, \dots, e_n\}$ is a basis of V , it must be the case that $e_\alpha^\bullet \in \langle v \otimes v : v \in V \rangle$, $\alpha \in G_{2,n}$. Clearly, $e_i \bullet e_i = e_i \otimes e_i \in \langle v \otimes v : v \in V \rangle$, but what about $e_i \bullet e_j$ when $i \neq j$? Observe that

$$\begin{aligned} 2e_i \bullet e_j &= e_i \otimes e_j + e_j \otimes e_i \\ &= (e_i + e_j) \otimes (e_i + e_j) - e_i \otimes e_i - e_j \otimes e_j. \end{aligned}$$

$\square$

Recall (Theorem 5.22) that $L(V^{\otimes m}, V^{\otimes m})$ is a model for $L(V, V)^{\otimes m}$ in which $T_1 \otimes T_2 \otimes \dots \otimes T_m$ is a typical decomposable tensor.

DEFINITION 6.70 Let $\mathcal{P}(\sigma^{-1}): L(V^{\otimes m}, V^{\otimes m}) \rightarrow L(V^{\otimes m}, V^{\otimes m})$, $\sigma \in S_m$, denote the linear operator on the vector space $L(V^{\otimes m}, V^{\otimes m})$ that satisfies

$$\mathcal{P}(\sigma^{-1})(T_1 \otimes T_2 \otimes \cdots \otimes T_m) = T_{\sigma(1)} \otimes T_{\sigma(2)} \otimes \cdots \otimes T_{\sigma(m)}, \quad (6.41)$$

for all $T_1, T_2, \dots, T_m \in L(V, V)$.

LEMMA 6.71 Suppose V is a vector space and $L \in L(V^{\otimes m}, V^{\otimes m})$. If $\sigma \in S_m$ is fixed but arbitrary, then $\mathcal{P}(\sigma)(L) = P(\sigma)LP(\sigma^{-1})$, where $P(\sigma) \in L(V^{\otimes m}, V^{\otimes m})$.

Proof Consider

$$\begin{aligned} & (\mathcal{P}(\sigma^{-1})(T_1 \otimes T_2 \otimes \cdots \otimes T_m))(v_1 \otimes v_2 \otimes \cdots \otimes v_m) \\ &= (T_{\sigma(1)} \otimes T_{\sigma(2)} \otimes \cdots \otimes T_{\sigma(m)})(v_1 \otimes v_2 \otimes \cdots \otimes v_m) \\ &= T_{\sigma(1)}(v_1) \otimes T_{\sigma(2)}(v_2) \otimes \cdots \otimes T_{\sigma(m)}(v_m) \\ &= P(\sigma^{-1})(T_1(v_{\sigma^{-1}(1)}) \otimes T_2(v_{\sigma^{-1}(2)}) \otimes \cdots \otimes T_m(v_{\sigma^{-1}(m)})) \\ &= (P(\sigma^{-1})T_1 \otimes T_2 \otimes \cdots \otimes T_m P(\sigma))(v_1 \otimes v_2 \otimes \cdots \otimes v_m). \end{aligned}$$

Because the linear operators, $\mathcal{P}(\sigma^{-1})(T_1 \otimes T_2 \otimes \cdots \otimes T_m)$ and $P(\sigma^{-1})T_1 \otimes T_2 \otimes \cdots \otimes T_m P(\sigma)$ agree on all decomposable tensors, $v_1 \otimes v_2 \otimes \cdots \otimes v_m$, they are equal. Because

$$\mathcal{P}(\sigma^{-1})(T_1 \otimes T_2 \otimes \cdots \otimes T_m) = P(\sigma^{-1})T_1 \otimes T_2 \otimes \cdots \otimes T_m P(\sigma)$$

for all decomposable tensors, $T_1 \otimes T_2 \otimes \cdots \otimes T_m$, it must be that $\mathcal{P}(\sigma^{-1})(L) = P(\sigma^{-1})LP(\sigma)$, for all $L \in L(V^{\otimes m}, V^{\otimes m})$. $\square$

It follows from Lemma 6.71 that $L \in L(V^{\otimes m}, V^{\otimes m})$ is a fixed point of $\mathcal{P}(\sigma) \in L(L(V^{\otimes m}, V^{\otimes m}), L(V^{\otimes m}, V^{\otimes m}))$ if and only if L commutes with $P(\sigma) \in L(V^{\otimes m}, V^{\otimes m})$.

DEFINITION 6.72 Suppose $L \in L(V^{\otimes m}, V^{\otimes m})$. Then L is **bisymmetric** if it commutes with $P(\sigma)$, for all $\sigma \in S_m$.

We now deduce a technical result of considerable importance to the determination of the rational irreducible representations of $GL(n, \mathbb{C})$.

COROLLARY 6.73 Suppose V is a vector space and L is a linear operator on $V^{\otimes m}$. Then L is bisymmetric if and only if L belongs to $\langle T^{\otimes m} : T \in L(V, V) \rangle$, the linear closure of $\{T^{\otimes m} : T \in L(V, V)\}$.

Proof By Lemma 6.71, $P(\sigma)L = LP(\sigma)$ if and only if $\mathcal{P}(\sigma)(L) = L$. By Theorem 6.63, $\mathcal{P}(\sigma)(L) = L$, for all $\sigma \in S_m$ if and only if L belongs to the symmetry class of $L(V^{\otimes m}, V^{\otimes m}) = [L(V, V)]^{\otimes m}$ consisting of the completely symmetric tensors, that is, to $[L(V, V)]_1(S_m)$. By Theorem 6.68, this symmetry class is spanned by $\{T \otimes T \otimes \cdots \otimes T : T \in L(V, V)\}$. $\square$

Because $T^{\otimes m}$ commutes with $P(\sigma)$ for all $\sigma \in S_m$, it commutes with any projection of the form $T(G, \chi)$. Hence, for any subgroup G of S_m , for all $\chi \in I(G)$, and for all $T \in L(V, V)$, the symmetry class of tensors, $V_\chi(G)$, is an invariant subspace of $T^{\otimes m}$.

DEFINITION 6.74 Let G be a subgroup of S_m and suppose $\chi \in I(G)$. Let V be a vector space and suppose $T \in L(V, V)$. Denote by $K(T)$ the restriction of $T^{\otimes m}$ to $V_\chi(G)$. Then $K(T) \in L(V_\chi(G), V_\chi(G))$ is the induced (or associated) transformation determined by G and χ .

Understand that $K(T)$ depends on G and χ as well as T .

THEOREM 6.75 Let χ be an irreducible character of the subgroup G of S_m . Let V be an inner product space and suppose $S, T \in L(V, V)$. If $V_\chi(G) \neq 0$, then

- (i) $K(S)K(T) = K(ST)$;
- (ii) $K(S)(v_1 * v_2 * \cdots * v_m) = S(v_1) * S(v_2) * \cdots * S(v_m)$;
- (iii) $K(S^*) = K(S)^*$, the adjoint of $K(S)$ with respect to the restriction of the induced inner product to $V_\chi(G)$;
- (iv) $K(S)$ is invertible for all invertible S ;
- (v) $K(S + T) \geq K(S) + K(T)$ whenever $S, T \geq 0$.

Proof For the most part, these results are obtained by restricting $S^{\otimes m}$ and $T^{\otimes m}$ to $V_\chi(G)$. Part (i) is a consequence of Theorem 5.20. To prove part (ii), observe that

$$\begin{aligned}
 K(S)(v_1 * v_2 * \cdots * v_m) &= (S \otimes S \otimes \cdots \otimes S)T(G, \chi)(v_1 \otimes v_2 \otimes \cdots \otimes v_m) \\
 &= T(G, \chi)(S \otimes S \otimes \cdots \otimes S)(v_1 \otimes v_2 \otimes \cdots \otimes v_m) \\
 &= T(G, \chi)(S(v_1) \otimes S(v_2) \otimes \cdots \otimes S(v_m)) \\
 &= S(v_1) * S(v_2) * \cdots * S(v_m).
 \end{aligned}$$

Part (iii) follows from Corollary 5.33, part (iv) from Corollary 5.21, and part (v) from Corollary 5.39. $\square$

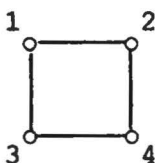


FIGURE 6.3

Applications to the Theory of Enumeration

Suppose we color each vertex of the square in Figure 6.3 either red, white, or blue. Any such coloring can be described by means of a function $f: \{1, 2, 3, 4\} \rightarrow \{r, w, b\}$. To count the different red-white-blue vertex colorings $f = (f(1), f(2), f(3), f(4))$, observe that there are four decisions to be made, each having three choices. By the Fundamental Counting Principle, there are $3^4 = 81$ colorings. The set C , of all functions $f: \{1, 2, 3, 4\} \rightarrow \{r, w, b\}$, contains 81 elements.

Four of these 81 colorings are illustrated in Figure 6.4. Look at them carefully. How different will they be after the paint dries and the colored squares are free to rotate? It seems 81 is the right answer to the wrong question. Let's formulate the right question. Say two colorings (functions in C) are **equivalent** if one can be obtained from the other by a rotation of the square. This equivalence relation partitions C into equivalence classes called **color patterns**. The four colorings in Figure 6.4 comprise one color pattern. The right question is, how many color patterns are there?

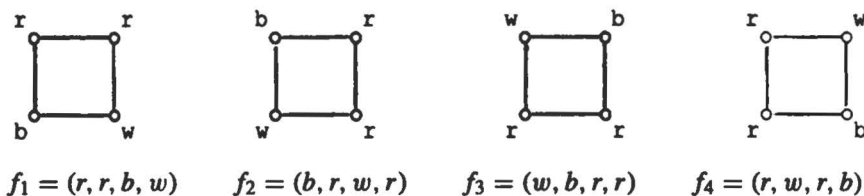


FIGURE 6.4

Let's begin by figuring out exactly how the functions f_1 and f_2 in Figure 6.4 are related. Geometrically, the *coloring* f_2 can be obtained from f_1 by a 90° clockwise rotation, the symmetry associated with $\sigma = (1243)$ (with respect to the numbering exhibited in Figure 6.3). However, the *function* $f_2 \neq \sigma f_1$. In fact, σf_1 is

meaningless. The composition of σ and f_1 makes sense, but only in the order $f_1\sigma$. On the other hand, $f_1\sigma = (r, w, r, b) = f_4$, not f_2 . The correct combination of f_1 , f_2 , and σ is $f_2 = f_1\sigma^{-1}$. In general, when symmetry σ is applied to an f -colored square, it produces another colored square, the one corresponding to $f\sigma^{-1}$. This is interesting. Associated with each symmetry of the square is a permutation of *colored* squares. The permutation $\sigma \in S_4$ induces an action $\bar{\sigma}: C \rightarrow C$ defined by $\bar{\sigma}(f) = f\sigma^{-1}$. This induced action is familiar. If we identify C with $\Gamma_{4,3}$, it is identical to the action described in Definition 6.17. Evidently, two colorings of the square are equivalent if and only if the corresponding functions are equivalent modulo the group $G = \{e_4, (1243), (14)(23), (1342)\}$ of its plane rotations. It follows that Δ is a system of distinct representatives for the red-white-blue color patterns, and the number of inequivalent color patterns may be obtained directly from Theorem 6.35, that is,

$$\begin{aligned} o(\Delta) &= \frac{1}{o(G)} \sum_{\sigma \in G} 3^{c(\sigma)} \\ &= (3^4 + 3 + 3^2 + 3)/4 \\ &= 24. \end{aligned}$$

Of the 81 different red-white-blue colorings of the square, only 24 are inequivalent modulo its group of plane rotations. A system of distinct representatives for these 24 color patterns (a “colorful” rendering of Δ) is exhibited in Figure 6.5.

Suppose it were *your* task to come up with a system of distinct representatives like the one illustrated in Figure 6.5. It would, of course, be enormously helpful to know that there are (only) 24 color patterns. (Once 24 inequivalent colorings are found, the job is finished.) It would be even more helpful to know, for example, that there are exactly three inequivalent colorings in which two vertices are red, one is white, and one blue. Let’s define the *weight* of such a color pattern to be r^2wb . Then the *pattern inventory* corresponding to the inequivalent red-white-blue vertex colorings of the square (Figure 6.5) is

$$\begin{aligned} W_G(r, w, b) &= (r^4 + w^4 + b^4) + (r^3w + r^3b + rw^3 + rb^3 + w^3b + wb^3) \\ &\quad + 2(r^2w^2 + r^2b^2 + w^2b^2) + 3(r^2wb + rw^2b + rwb^2). \end{aligned} \tag{6.42}$$

Note that $W_G(1, 1, 1) = 24$, reflecting the fact that each pattern contributes one monomial (its weight) to $W_G(r, w, b)$. In terms of monomial symmetric functions,

$$W_G(r, w, b) = M_{[4]}(r, w, b) + M_{[3,1]}(r, w, b) + 2M_{[2^2]}(r, w, b) + 3M_{[2,1^2]}(r, w, b).$$

DEFINITION 6.77 Let χ be an irreducible character of the subgroup G of S_m . If $x_1, x_2, \dots, x_n$ are independent indeterminates, the **character weighted pattern inventory** is defined by

$$\begin{aligned} W_G^\chi(x_1, x_2, \dots, x_n) &= \sum_{\alpha \in \bar{\Delta}} (\chi, 1)_{G_\alpha} w(\alpha) \\ &= \sum_{\alpha \in \Delta} (\chi, 1)_{G_\alpha} w(\alpha). \end{aligned}$$

When χ is the principle character, $W_G^1(x_1, x_2, \dots, x_n)$ is the traditional **pattern inventory**, abbreviated

$$W_G(x_1, x_2, \dots, x_n) = \sum_{\alpha \in \Delta} w(\alpha).$$

Suppose $G = S_m$ and $\chi = \chi_\pi$ for some $\pi \vdash m$. If $\alpha \in \Delta$, then G_α is isomorphic to the Young subgroup S_μ , where $\mu = \mu(\alpha)$ is its multiplicity partition, and $(\chi, 1)_{S_\mu}$ is the Kostka coefficient $K_{\pi, \mu}$. In this case,

$$W_{S_m}^\chi(x_1, x_2, \dots, x_n) = \sum_{\alpha \in G_{m,n}} K_{\pi, \mu(\alpha)} \prod_{i=1}^m x_{\alpha(i)} \quad (6.44)$$

is a so-called **Schur polynomial**. (We shall have more to say about Schur polynomials in Chapter 8.)

DEFINITION 6.78 Suppose χ is an irreducible character of the subgroup G of S_m . Let $y_1, y_2, \dots, y_m$ be independent indeterminates over the complex numbers. The **character weighted cycle index polynomial** afforded by G and χ is

$$Z_G^\chi(y_1, y_2, \dots, y_m) = \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma) \prod_{i=1}^m y_i^{c_i(\sigma)},$$

where, recall, $c_i(\sigma)$ is the number of cycles of length i in the disjoint cycle factorization of σ . When $\chi = 1$, $Z_G^\chi(y_1, y_2, \dots, y_m)$ is the traditional **cycle index polynomial**, abbreviated $Z_G(y_1, y_2, \dots, y_m)$.

EXAMPLE 6.79 Let $m = 4$. The cycle index polynomial of $G = \{e_4, (12), (34), (12)(34)\}$ is

$$Z_G(y_1, y_2, y_3, y_4) = (y_1^4 + 2y_1^2y_2 + y_2^2)/4;$$

the cycle index polynomial of $H = \{e_4, (12)(34), (13)(24), (14)(23)\}$ is

$$Z_H(y_1, y_2, y_3, y_4) = (y_1^4 + 3y_2^2)/4.$$

Observe that neither y_3 nor y_4 actually appears in either Z_G or Z_H . Moreover, while G and H are isomorphic (as abstract groups), they have different cycle index polynomials. $\square$

These definitions are brought together in the following result of [Williamson (1971)], [Merris (1980)], and [White (1980b)].

THEOREM 6.80 *Suppose χ is an irreducible character of the subgroup G of S_m . Then*

$$\chi(e)W_G^\chi(x_1, x_2, \dots, x_n) = Z_G^\chi(P_1, P_2, \dots, P_m), \quad (6.45)$$

where $P_t = P_t(x_1, x_2, \dots, x_n) = x_1^t + x_2^t + \dots + x_n^t$ is the t -th power sum, $1 \leq t \leq m$.

COROLLARY 6.81 (*Pólya's Theorem*). *Let G be a permutation group of degree m . Then the pattern inventory, $W_G(x_1, x_2, \dots, x_n)$, is obtained from the cycle index polynomial, $Z_G(y_1, y_2, \dots, y_m)$, by the substitution $y_t = P_t(x_1, x_2, \dots, x_n)$, $1 \leq t \leq m$.*

Proof (of Theorem 6.80): Let U be a vector space of dimension n , over the scalar field of rational functions (ratios of polynomials) in the indeterminates $x_1, x_2, \dots, x_n$. Let $\{u_1, u_2, \dots, u_n\}$ be a basis of U . If $\alpha \in \Gamma_{m,n}$, then $U_\alpha^\otimes = \langle u_{\alpha\tau}^\otimes : \tau \in G \rangle$ is invariant under $\{P(\sigma) : \sigma \in G\}$. Denote by $P_\alpha(\sigma)$ the restriction of $P(\sigma)$ to $U_\alpha^\otimes$, $\sigma \in G$. Following [Williamson (1971)], we define $\tilde{P}(\sigma) : U^{\otimes m} \rightarrow U^{\otimes m}$ by

$$\tilde{P}(\sigma) = \bigoplus_{\alpha \in \Delta} w(\alpha) P_\alpha(\sigma),$$

and

$$\tilde{T}(G, \chi) = \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma) \tilde{P}(\sigma). \quad (6.46)$$

Because $U_\alpha^\otimes$ is invariant under $\{P(\sigma) : \sigma \in G\}$, it is invariant under $T(G, \chi)$. Denote the restriction of $T(G, \chi)$ to $U_\alpha^\otimes$ by $T_\alpha(G, \chi)$. Then $T_\alpha(G, \chi)$ is a projection operator on $U_\alpha^\otimes$, and

$$\tilde{T}(G, \chi) = \bigoplus_{\alpha \in \Delta} w(\alpha) T_\alpha(G, \chi). \quad (6.47)$$

From Equation (6.47), the trace of $\tilde{T}(G, \chi)$ is

$$\sum_{\alpha \in \Delta} w(\alpha) \operatorname{tr}(T_\alpha(G, \chi)) = \sum_{\alpha \in \Delta} s_\alpha w(\alpha).$$

From Equation (6.46), the trace of $\tilde{T}(G, \chi)$ is

$$\frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma) \operatorname{tr}(\tilde{P}(\sigma)) = \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma) \sum_{\alpha \in \Delta} w(\alpha) \operatorname{tr}(P_\alpha(\sigma)).$$

Therefore,

$$\sum_{\alpha \in \Delta} s_{\alpha} w(\alpha) = \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma) \sum_{\alpha \in \Delta} w(\alpha) \operatorname{tr}(P_{\alpha}(\sigma)) \quad (6.48)$$

For a particular weight

$$w = \prod_{i=1}^n x_i^{k_i},$$

(where each $k_i \geq 0$, and $k_1 + k_2 + \cdots + k_n = m$), denote by Δ_w the sequences in Δ of weight w , that is, $\Delta_w = \{\beta \in \Delta : w(\beta) = w\}$. Then

$$\sum_{\alpha \in \Delta} \operatorname{tr}(P_{\alpha}(\sigma)) \prod_{i=1}^m x_{\alpha(i)} = \sum_w w \sum_{\alpha \in \Delta_w} \operatorname{tr}(P_{\alpha}(\sigma)). \quad (6.49)$$

Summing $\operatorname{tr}(P_{\alpha}(\sigma)) = o(\{\gamma \in \Gamma_{m,n} : \gamma \equiv \alpha \pmod{G} \text{ and } \gamma\sigma = \gamma\})$ over $\alpha \in \Delta_w$ produces $\sum \operatorname{tr}(P_{\alpha}(\sigma)) = o(\{\gamma \in \Gamma_{m,n} : w(\gamma) = w \text{ and } \gamma\sigma = \gamma\})$. Substituting this in Equation (6.49) produces

$$\sum_{\alpha \in \Delta} w(\alpha) \operatorname{tr}(P_{\alpha}(\sigma)) = \sum_w o(\{\gamma \in \Gamma_{m,n} : w(\gamma) = w \text{ and } \gamma\sigma = \gamma\}) w. \quad (6.50)$$

Now, $\gamma\sigma = \gamma$ if and only if γ is constant on the cycles of σ . The right-hand side of Equation (6.50) is an inventory of the weights of those γ 's that are constant on the cycles of σ . Another way to inventory the same weights is to compute the product,

$$(x_1 + x_2 + \cdots + x_n)^{c_1(\sigma)} (x_1^2 + x_2^2 + \cdots + x_n^2)^{c_2(\sigma)} \cdots (x_1^m + x_2^m + \cdots + x_n^m)^{c_m(\sigma)}.$$

Substituting this product in the right-hand side of Equation (6.50) and then substituting the result in the right-hand side of Equation (6.48), we obtain

$$\sum_{\alpha \in \Delta} s_{\alpha} w(\alpha) = \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma) \prod_{i=1}^m \left(\sum_{i=1}^n x_i^i \right)^{c_i(\sigma)}.$$

□

EXAMPLE 6.82 Suppose $m = 4$, $n = 3$, $G = \langle (1243) \rangle = \{e_4, (1243), (14)(23), (1342)\}$, and $\chi = 1$. Then

$$Z_G(y_1, y_2, y_3, y_4) = \frac{1}{4}(y_1^4 + 2y_2^4 + y_2^2).$$

Letting $x_1 = r$, $x_2 = w$, and $x_3 = b$, Corollary 6.81 yields

$$\begin{aligned} W_G(r, w, b) &= \frac{1}{4}((r + w + b)^4 + 2(r^4 + w^4 + b^4) + (r^2 + w^2 + b^2)^2) \\ &= (r^4 + w^4 + b^4) + (r^3w + r^3b + rw^3 + rb^3 + w^3b + wb^3) \\ &\quad + 2(r^2w^2 + r^2b^2 + w^2b^2) + 3(r^2wb + rw^2b + rwb^2), \end{aligned}$$

precisely Equation (6.42). $\square$

EXAMPLE 6.83 Modulo

$$D_4 = \{e_4, (1243), (14)(23), (1342), (14), (23), (12)(34), (13)(24)\} \quad (6.51)$$

the group of all eight symmetries of the square, how many red-white-blue vertex color patterns are there? Does doubling the symmetry group halve the number of patterns? Let's see. By Theorem 6.35,

$$\begin{aligned} o(\Delta) &= \frac{1}{8} \sum_{\sigma \in D_4} 3^{c(\sigma)} \\ &= (3^4 + 3 + 3^2 + 3 + 3^3 + 3^3 + 3^2 + 3^2)/8 \\ &= 21. \end{aligned}$$

So, going from the group of plane rotations, G , to the full symmetry group, D_4 , reduces the number of patterns from 24 to 21. Take another look at Figure 6.5 and see if you can determine which of the colorings, inequivalent modulo G , are equivalent modulo D_4 . (At the conclusion of this little exercise, you should be able to use the figure to write out the pattern inventory for D_4 .)

From Pólya's Theorem (Corollary 6.81), the pattern inventory for D_4 is

$$\begin{aligned} W_{D_4}(r, w, b) &= Z_{D_4}(r + w + b, r^2 + w^2 + b^2, r^3 + w^3 + b^3, r^4 + w^4 + b^4) \\ &= \frac{1}{8}((r + w + b)^4 + 2(r^4 + w^4 + b^4) + 3(r^2 + w^2 + b^2)^2 \\ &\quad + 2(r + w + b)^2(r^2 + w^2 + b^2)) \\ &= (r^4 + w^4 + b^4) + (r^3w + r^3b + rw^3 + rb^3 + w^3b + wb^3) \\ &\quad + 2(r^2w^2 + r^2b^2 + w^2b^2) + 2(r^2wb + rw^2b + rwb^2). \end{aligned}$$

$\square$

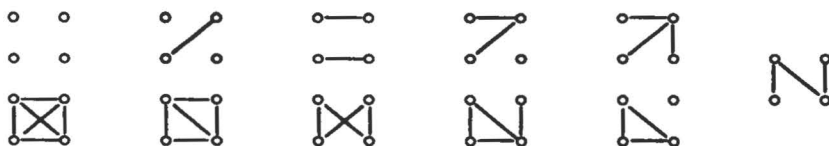


FIGURE 6.6

The Enumeration of Graphs

Pólya's Theorem can be found in the 1937 paper, "Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen" (Combinatorial Enumeration of Groups, Graphs, and Chemical Compounds). To discuss Pólya's approach to graph enumeration, let $g(n, m)$ be the number of nonisomorphic graphs with n vertices and m edges. Define

$$f_n(x) = \sum_{m=0}^{C(n,2)} g(n, m)x^m. \quad (6.52)$$

Then $f_n(x)$ is a generating function for the numbers of nonisomorphic graphs having n vertices.

EXAMPLE 6.84 The 11 nonisomorphic graphs on four vertices are illustrated in Figure 6.6. Using these pictures, it is easy to produce

$$f_4(x) = 1 + x + 2x^2 + 3x^3 + 2x^4 + x^5 + x^6. \quad (6.53)$$

□

Because there is a unique graph, namely $K_n = (V, V^{(2)})$, having n vertices and $C(n, 2)$ edges, $g(n, C(n, 2)) = 1$, that is, $f_n(x)$ is a monic polynomial of degree $C(n, 2)$. Because $G = (V, E)$ and $H = (W, F)$ are isomorphic if and only if their complements, $G^c = (V, V^{(2)} \setminus E)$ and $H^c = (W, W^{(2)} \setminus F)$, are isomorphic, the coefficients of $f_n(x)$ are symmetric.

If we had a picture, comparable to Figure 6.6, for the 34 non-isomorphic graphs on five vertices, it would be a simple matter to write down

$$f_5(x) = x^{10} + x^9 + 2x^8 + 4x^7 + 6x^6 + 6x^5 + 6x^4 + 4x^3 + 2x^2 + x + 1. \quad (6.54)$$

On the other hand, if it were your job to produce such a picture, it would surely be useful to know, for example, that the coefficient of x^4 in $f_5(x)$ is 6, in other

words, that there exactly six nonisomorphic graphs having five vertices and four edges. Okay, how does one generate $f_n(x)$ without any pictures?

To begin, set $V = \{1, 2, \dots, n\}$. Then (Definition 1.31) $G = (V, E)$ and $H = (V, F)$ are isomorphic if and only if there is a permutation $\sigma \in S_n$ such that

$$\{i, j\} \in E \text{ if and only if } \{\sigma(i), \sigma(j)\} \in F. \quad (6.55)$$

DEFINITION 6.85 For each $\sigma \in S_n$, let $\tilde{\sigma}$ be the natural induced action of σ on $V^{(2)}$, the 2-element subsets of $V = \{1, 2, \dots, n\}$, that is, $\tilde{\sigma}(\{i, j\}) = \{\sigma(i), \sigma(j)\}$, $\{i, j\} \in V^{(2)}$. The pair group $S_n^{(2)} = \{\tilde{\sigma} : \sigma \in S_n\}$.

With respect to this induced action, (6.55) can be expressed as $e \in E$ if and only if $\tilde{\sigma}(e) \in F$. In other words, $G = (V, E)$ is isomorphic to $H = (V, F)$ if and only if there is a permutation $\tilde{\sigma}$ in the pair group $S_n^{(2)}$ such that

$$\tilde{\sigma}(E) = F. \quad (6.56)$$

There is another way to look at $S_n^{(2)}$. Consider the complete graph $K_n = (V, V^{(2)})$. As a geometric object, it has a symmetry group, namely its group of automorphisms.¹² Let's call it $\mathcal{G}$. When viewed as permutations of the vertices of K_n , $\mathcal{G} = S_n$. However, when viewed as permutations of the edges of K_n , $\mathcal{G} = S_n^{(2)}$.

Let's color the edges of K_n using two colors, say black and white. Any such coloring corresponds to a partition of $V^{(2)}$ into $E = \{e \in V^{(2)} : e \text{ is black}\}$ and $V^{(2)} \setminus E = \{e \in V^{(2)} : e \text{ is white}\}$. Moreover, if E is the set of black edges from one coloring and F the set of black edges from another, the two colorings are equivalent modulo $\mathcal{G}$ if and only if there is a permutation $\tilde{\sigma}$ in the pair group $S_n^{(2)}$ such that $\tilde{\sigma}(E) = F$, precisely the criterion of Equation (6.56). In the natural one-to-one correspondence between graphs on n vertices and black-white edge colorings of K_n , two graphs are isomorphic if and only if the corresponding edge colorings are equivalent modulo $S_n^{(2)}$. Letting x_1 correspond to white and x_2 to black, the pattern inventory for the inequivalent black-white colorings of the edges of K_n is

$$W_G(x_1, x_2) = Z_{S_n^{(2)}}(x_1 + x_2, x_1^2 + x_2^2, \dots, x_1^m + x_2^m), \quad (6.57)$$

where $m = C(n, 2)$ is the degree of $S_n^{(2)}$. It seems that substituting $x_1 = 1$ and $x_2 = x$ in Equation (6.57) proves the following:

THEOREM 6.86 *The generating function for the numbers of nonisomorphic graphs on n vertices is*

$$f_n(x) = Z_{S_n^{(2)}}(1 + x, 1 + x^2, \dots, 1 + x^{C(n,2)}). \quad (6.58)$$

To find $f_n(x)$, it remains to compute the cycle index polynomial for $S_n^{(2)}$.

¹² An automorphism of G is an isomorphism from G onto G .

EXAMPLE 6.87 If $n = 4$ then $V = \{1, 2, 3, 4\}$, and $V^{(2)} = \{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \{3, 4\}\}$. Numbering the elements of $V^{(2)}$ lexicographically, we have

$$1 = \{1, 2\}, \quad 2 = \{1, 3\}, \quad 3 = \{1, 4\}, \quad 4 = \{2, 3\}, \quad 5 = \{2, 4\}, \quad 6 = \{3, 4\}.$$

Suppose $\sigma = (1234) \in S_4$. Let's compute the disjoint cycle factorization of $\tilde{\sigma} \in S_6$:

$$\tilde{\sigma}(1) = \tilde{\sigma}(\{1, 2\}) = \{\sigma(1), \sigma(2)\} = \{2, 3\} = 4;$$

$$\tilde{\sigma}(4) = \tilde{\sigma}(\{2, 3\}) = \{\sigma(2), \sigma(3)\} = \{3, 4\} = 6;$$

$$\tilde{\sigma}(6) = \tilde{\sigma}(\{3, 4\}) = \{\sigma(3), \sigma(4)\} = \{4, 1\} = 3;$$

$$\tilde{\sigma}(3) = \tilde{\sigma}(\{1, 4\}) = \{\sigma(1), \sigma(4)\} = \{2, 1\} = 1.$$

Therefore, (1463) is a cycle of $\tilde{\sigma}$. Continuing,

$$\tilde{\sigma}(2) = \tilde{\sigma}(\{1, 3\}) = \{\sigma(1), \sigma(3)\} = \{2, 4\} = 5;$$

$$\tilde{\sigma}(5) = \tilde{\sigma}(\{2, 4\}) = \{\sigma(2), \sigma(4)\} = \{3, 1\} = 2.$$

Hence, $\tilde{\sigma} = (1463)(25)$. Figure 6.7 displays $\tilde{\sigma}$ for all 24 permutations $\sigma \in S_4$. $\square$

EXAMPLE 6.88 From Figure 6.7, the cycle index polynomial

$$Z_{S_4^{(2)}}(y_1, y_2, \dots, y_6) = \frac{1}{24} (y_1^6 + 9y_1^2y_2^2 + 8y_3^2 + 6y_2y_4). \quad (6.59)$$

σ	$\tilde{\sigma}$	σ	$\tilde{\sigma}$	σ	$\tilde{\sigma}$
e_4	e_6	(123)	(142)(356)	(1234)	(1463)(25)
(12)	(24)(35)	(124)	(153)(246)	(1243)	(1562)(34)
(13)	(14)(36)	(132)	(124)(365)	(1324)	(16)(2453)
(14)	(15)(26)	(134)	(145)(263)	(1342)	(1265)(34)
(23)	(12)(56)	(142)	(135)(264)	(1423)	(16)(2354)
(24)	(13)(46)	(143)	(154)(236)	(1432)	(1364)(25)
(34)	(23)(45)	(234)	(123)(465)	(12)(34)	(25)(34)
(13)(24)	(16)(34)	(243)	(132)(456)	(14)(23)	(16)(25)

FIGURE 6.7

In view of Theorem 6.86, substituting $y_i = 1 + x^i$ in Equation (6.59) produces

$$\begin{aligned}
 f_4(x) &= \left((1+x)^6 + 9(1+x)^2(1+x^2)^2 + 8(1+x^3)^2 + 6(1+x^2)(1+x^4) \right) / 24 \\
 &= \left((1+6x+15x^2+20x^3+15x^4+x^5+x^6) \right. \\
 &\quad \left. + 9(1+2x+3x^2+4x^3+3x^4+2x^5+x^6) \right. \\
 &\quad \left. + 8(1+2x^3+x^6) + 6(1+x^2+x^4+x^6) \right) / 24 \\
 &= 1 + x + 2x^2 + 3x^3 + 2x^4 + x^5 + x^6,
 \end{aligned}$$

precisely Equation (6.53). $\square$

The generation of cycle index polynomials for the pair group is not as difficult as Example 6.87 makes it seem:

LEMMA 6.89 *Let $\tilde{\sigma}$ and $\tilde{\tau}$ be the elements of $S_n^{(2)}$ induced by the permutations σ and τ of S_n , respectively. If σ and τ have the same cycle structure, then $\tilde{\sigma}$ and $\tilde{\tau}$ have the same cycle structure.*

EXAMPLE 6.90 The cycle index polynomial for S_5 is easy to obtain from Figure 4.4, namely,

$$\frac{1}{120} \left(y_1^5 + 10y_1^3y_2 + 15y_1y_2^2 + 20y_1^2y_3 + 20y_2y_3 + 30y_1y_4 + 24y_5 \right).$$

Given this expression and Lemma 6.89, only seven (as opposed to 120) computations of the type carried out in Example 6.87 suffice to yield the cycle index polynomial for $S_5^{(2)}$:

$$\frac{1}{120} \left(y_1^{10} + 10y_1^4y_2^3 + 15y_1^2y_2^4 + 20y_1y_3^3 + 20y_1y_3y_6 + 30y_2y_4^2 + 24y_5^2 \right).$$

$\square$

COROLLARY 6.91 *The total number of nonisomorphic graphs on n vertices (without regard to the number of edges) is*

$$f_n(1) = \frac{1}{n!} \sum_{\sigma \in S_n} 2^{c(\tilde{\sigma})}, \quad (6.60)$$

where $c(\tilde{\sigma})$ is the total number of cycles, including cycles of length 1, in the disjoint cycle factorization of $\tilde{\sigma}$.

Proof From Theorem 6.86,

$$f_n(1) = Z_{S_n^{(2)}}(2, 2, \dots, 2).$$

□

Combining Example 6.90 with Corollary 6.91, we see that there are

$$(2^{10} + 10 \times 2^7 + 15 \times 2^6 + 20 \times 2^4 + 20 \times 2^3 + 30 \times 2^3 + 24 \times 2^2) / 120 = 34$$

nonisomorphic graphs on 5 vertices (as predicted by Equation (6.54)).

Application to NMR Spectroscopy

We turn now to the character weighted cycle index polynomials.

EXAMPLE 6.92 Consider the red-white-blue vertex color patterns of the square, modulo D_4 , in which *opposite vertices are colored differently*. If $\chi = \varepsilon$ then, from Equation (6.51),

$$Z_{D_4}^\varepsilon(y_1, y_2, y_3, y_4) = (y_1^4 - 2y_4 + 3y_2^2 - 2y_1^2y_2) / 8.$$

Thus,

$$\begin{aligned} W_{D_4}^\varepsilon(r, w, b) &= ((r + w + b)^4 - 2(r^4 + w^4 + b^4) + 3(r^2 + w^2 + b^2)^2 \\ &\quad - 2(r + w + b)^2(r^2 + w^2 + b^2)) / 8 \\ &= (r^2w^2 + r^2b^2 + w^2b^2) + (r^2wb + rw^2b + rwb^2). \end{aligned}$$

Sure enough, these are the weights of the six inequivalent colorings illustrated in Figure 6.8. What may not be so obvious is why the alternating character should “select” precisely those patterns in which opposite vertices are colored differently.¹³

r r	b b	w w	r r	b b	w w
w w	r r	b b	w b	r w	r b

FIGURE 6.8

¹³Hint: The restriction of ε to $((12)(34))$ is equal to the principal character while its restrictions to $((14))$ is not.

EXAMPLE 6.93 From the algebraic perspective of symmetry classes of tensors, $\bar{\Delta}$ is the subset of Δ “selected” by an irreducible character. From the combinatorial perspective of pattern inventory, the role of the character in “selecting” patterns is less well understood. Exceptions sometimes occur in instances where characters are associated with some physical quantity. One example involves high resolution nuclear magnetic resonance (NMR) spectroscopy, an application that requires a brief excursion into quantum mechanics. (For a nice overview of this subject, see [Hollas (1987)] or [Schensted (1976)].)

The quantum mechanical description of the *state* of a system is given by a complex valued “wave function”, $\Psi(z)$, of the coordinates z of the system. Wave functions are governed by the Schrödinger equation

$$H\Psi = E\Psi, \quad (6.61)$$

linking the Hamiltonian (differential) operator H with the energy E . Solutions of Equation (6.61) describe the *stationary* states of the system. Transitions between these states correspond to spectral lines. For our purposes, the exact nature of H is unimportant. What matters is that it is linear,¹⁴ so the set of all solutions corresponding to a fixed (eigenvalue) E comprises a vector space, and that the Schrödinger equation for an n -particle system is invariant with respect to the symmetries of the system. It follows from these facts that the eigenspace of each stationary state corresponds to a representation of $G = \{\sigma \in S_n : H\sigma = \sigma H\}$. In other words, a character of G is associated with each stationary energy E . It is this natural occurrence of characters that makes Theorem 6.80 relevant to the discussion.

Skipping over many details, we come to the work of K. Balasubramanian on NMR spectroscopy.¹⁵ Nuclear magnetic resonance is produced by a magnetic field associated with unpaired nuclear spins. The phenomenon is observed by placing a sample in a steady magnetic field and exposing it to radio waves. The frequency of the radiation and the strength of the magnetic field are adjusted to produce absorption of the radio waves.

Among the triumphs of quantum mechanics is a theoretical insight leading to predictions associated with these (and other) “spectral lines”. Among its frustrations is the difficulty of obtaining exact solutions to the Schrödinger equation. The 1927 Born-Oppenheimer approximation leads to a factorization of the wave function $\Psi = \Psi_e \Psi_n$ as a product of an electron part and a nuclear part. A further splitting of the nuclear part leads to $\Psi_n = \Psi_v \Psi_r \Psi_s$, where Ψ_v is a vibrational part, Ψ_r a rotational piece, and Ψ_s a nuclear spin factor. Associated with the spin factor is a reducible character χ_s . The key to predicting relative intensities of the NMR spectral lines is a reduction of χ_s .

¹⁴This is a consequence of the principal of superposition of states, an axiom of quantum mechanics.

¹⁵See [Balasubramanian (1981)–(1985)].



FIGURE 6.9

The nuclear spin quantum number, I , is of the form $k/2$, for some nonnegative integer k . Hydrogen (atomic weight 1) has nuclear spin quantum number $1/2$. For carbon-12, $I = 0$. Nitrogen-14 has quantum number 1 and $I = 3/2$ for chlorine-35. Corresponding to quantum number I , there are $n = 2I + 1$ solutions Ψ_s to the NMR wave equation.

Consider, for example, the (nonrigid) hydrazine molecule N_2H_4 illustrated in Figure 6.9(a). We will focus on the “proton nuclear spin species” arising from the four hydrogen-1 atoms numbered as in Figure 6.9(b). Because $I = 1/2$, Ψ_s has $n = 2$ solutions, $\Psi_s = \mu$ and $\Psi_s = \delta$, corresponding to “spin up” and “spin down”. Let V be the free vector space spanned by these two wave functions and consider $V^{\otimes 4}$, where $m = 4$ because we are dealing with four protons. Setting $e_1 = \mu$ and $e_2 = \delta$, the decomposable tensor

$$\mu \otimes \mu \otimes \delta \otimes \delta = e_1 \otimes e_1 \otimes e_2 \otimes e_2$$

corresponds to spin up for protons 1 and 2, and spin down for protons 3 and 4.

The nuclear spin character, χ_s , is the (permutation) character afforded by the representation $\sigma \rightarrow P(\sigma) \in L(V^{\otimes 4}, V^{\otimes 4})$, $\sigma \in G$. Let $\Delta_i = \{\gamma \in \Delta : w(\gamma) = x_1^i x_2^{4-i}\}$, $0 \leq i \leq 4$, and denote by ρ_i the character of G afforded by the restriction of P to $\langle e_{\gamma\sigma}^{\otimes} : \gamma \in \Delta_i, \sigma \in G \rangle$. Then

$$\chi_s = \sum_{i=0}^4 \rho_i. \quad (6.62)$$

Therefore, it suffices to reduce ρ_i , $0 \leq i \leq 4$.

	C_1	C_2	C_3	C_4	C_5
χ_1	1	1	1	1	1
χ_2	1	1	-1	-1	1
χ_3	1	1	-1	1	-1
χ_4	1	1	1	-1	-1
χ_5	2	-2	0	0	0

FIGURE 6.10 The character table for D_4 .



FIGURE 6.11

The symmetry group of nonrigid hydrazine is $G = D_4$. From Example 4.42, the conjugacy classes of D_4 are $C_1 = \{e\}$, $C_2 = \{(14)(23)\}$, $C_3 = \{(1243), (1342)\}$, $C_4 = \{(14), (23)\}$, and $C_5 = \{(12)(34), (13)(24)\}$; its character table (Figure 4.3) is reproduced in Figure 6.10. Theorem 6.80, applied to $G = D_4$ and $\chi = \chi_i$, $1 \leq i \leq 5$, produces

$$W_G^{\chi_1}(x_1, x_2) = x_1^4 + x_1^3 x_2 + 2x_1^2 x_2^2 + x_1 x_2^3 + x_2^4;$$

$$W_G^{\chi_2}(x_1, x_2) = x_1^2 x_2^2;$$

$$W_G^{\chi_3}(x_1, x_2) = x_1^3 x_2 + x_1^2 x_2^2 + x_1 x_2^3;$$

$$W_G^{\chi_4}(x_1, x_2) = 0;$$

and (after dividing both sides by $\chi_5(e) = 2$)

$$W_G^{\chi_5}(x_1, x_2) = x_1^3 x_2 + x_1^2 x_2^2 + x_1 x_2^3.$$

Consider the term $2x_1^2 x_2^2$ in the pattern inventory $W_G(x_1, x_2) = W_G^{\chi_1}(x_1, x_2)$. Evidently, there are two inequivalent spin species of weight $x_1^2 x_2^2$. In the current example, the situation is clear. Representatives of the two species are illustrated in Figure 6.11, and $(e_{\gamma\sigma}^{\otimes} : \gamma \in \Delta_2, \sigma \in G)$ is the direct sum of the invariant subspaces

$$\begin{aligned}
 U_1 &= (\mu \otimes \mu \otimes \delta \otimes \delta, \delta \otimes \delta \otimes \mu \otimes \mu, \mu \otimes \delta \otimes \mu \otimes \delta, \delta \otimes \mu \otimes \delta \otimes \mu) \\
 &= (e_{\alpha\sigma}^{\otimes} : \sigma \in G),
 \end{aligned}$$

where $\alpha = (1, 1, 2, 2)$, and

$$\begin{aligned}
 U_2 &= (\mu \otimes \delta \otimes \delta \otimes \mu, \delta \otimes \mu \otimes \mu \otimes \delta) \\
 &= (e_{\beta\sigma}^{\otimes} : \sigma \in G),
 \end{aligned}$$

where $\beta = (1, 2, 2, 1)$. Therefore, $\rho_2 = \xi_1 + \xi_2$, where ξ_i is the character of the restriction of P to U_i , $i = 1, 2$. Because the ξ_i are transitive permutation characters, they are induced from principal characters of stabilizer subgroups. The stabilizer subgroup of the spin species $e_\alpha^\otimes$, for example, is $G_\alpha = \{e, (12)(34)\}$. By the Frobenius Reciprocity Theorem, $(\xi_1, \chi_j)_G = (1, \chi_j)_{G_\alpha}$, and $(\xi_2, \chi_j)_G = (1, \chi_j)_{G_\beta}$, $1 \leq j \leq 5$. Thus,

$$\begin{aligned}(\rho_2, \chi_j)_G &= (\xi_1 + \xi_2, \chi_j)_G \\&= (\xi_1, \chi_j)_G + (\xi_2, \chi_j)_G \\&= (1, \chi_j)_{G_\alpha} + (1, \chi_j)_{G_\beta} \\&= \sum_{\gamma \in \Delta_2} (1, \chi_j)_{G_\gamma}.\end{aligned}$$

In general,

$$(\rho_i, \chi_j)_G = \sum_{\gamma \in \Delta_i} (1, \chi_j)_{G_\gamma},$$

precisely the coefficient of $x_1^i x_2^{4-i}$ in $W_G^{X_j}(x_1, x_2)$. (See Definition 6.77.) Because

$$\rho_i = \sum_{j=0}^5 (\rho_i, \chi_j)_G \chi_j, \quad 0 \leq i \leq 4,$$

we can obtain a reduction of ρ_i , by inspection, from the coefficients of $x_1^i x_2^{4-i}$ in $W_G^{X_j}(x_1, x_2)$, $1 \leq j \leq 5$. In particular,

$$\rho_0 = \chi_1 = \rho_4;$$

$$\rho_1 = \chi_1 + \chi_3 + \chi_5 = \rho_3;$$

and

$$\rho_2 = 2\chi_1 + \chi_2 + \chi_3 + \chi_5.$$

Substituting these values into Equation (6.62), yields the nuclear spin character

$$\chi_s = 6\chi_1 + \chi_2 + 3\chi_3 + 3\chi_5.$$

□

Exercises

1. Suppose χ is a linear character of G . Prove that $P(\tau)T(G, \chi) = \chi(\tau^{-1})T(G, \chi)$.
2. Let $m = 2$, $G = S_2$, and $\chi = \varepsilon$. Confirm Equation (6.7) by showing directly that $v_1 * v_2 = -v_2 * v_1$.
3. Mimic Example 6.8 when $G = S_3$.
4. Let $\lambda_1, \lambda_2, \dots, \lambda_p$ be the eigenvalues of $A \in \mathbf{C}_{p,p}$ (multiplicities included), and $\omega_1, \omega_2, \dots, \omega_q$ the eigenvalues of $B \in \mathbf{C}_{q,q}$. Find the eigenvalues of
 - a. $A \otimes B - B \otimes A$.
 - b. $A \otimes B + B \otimes A$.
5. Let $A, B \in \mathbf{C}_{n,n}$. If A is skew symmetric and B is symmetric, show that $\text{tr}(B^*A) = 0$.
6. Suppose $\Phi: V \times V \times \dots \times V \rightarrow W$ is m -linear. Let χ be a linear character of the subgroup G of S_m . Prove that Φ is symmetric with respect to G and χ if and only if $\Phi(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(m)}) = \chi(\sigma)\Phi(v_1, v_2, \dots, v_m)$, $\sigma \in G$.
7. Prove Lemma 6.12.
8. Let $f_1, f_2, \dots, f_m \in V' = L(V, \mathbf{C})$. Prove that

$$\Phi = \sum_{\tau \in G} \chi(\tau) \prod_{i=1}^m f_{\tau(i)}$$

is multilinear and symmetric with respect to G and χ .

9. Prove that equivalence modulo G is an equivalence relation on $\Gamma_{m,n}$.
10. Prove that the mapping $\sigma \rightarrow \bar{\sigma}$, from S_m into S_n^m , introduced in Definition 6.17, is one-to-one for all $n > 1$.
11. Suppose G is a subgroup of S_m . Let $\alpha, \beta \in \Gamma_{m,n}$.
 - a. Prove that G_α is a subgroup of G .
 - b. If $\alpha \equiv \beta \pmod{G}$, prove that their stabilizer subgroups, G_α and G_β , are conjugate in G .
 - c. Suppose $G_\alpha \subset G_\beta$. If $\beta \in \Omega$, prove that $\alpha \in \Omega$.
12. Prove Lemma 6.22. (Hint: Exercise 11b.)
13. Let $\chi \in I(G)$ and $\alpha, \beta \in \Gamma_{m,n}$, where G is a subgroup of S_m .
 - a. Prove that

$$\frac{1}{o(G_\alpha)} \sum_{\sigma \in G_\alpha} \chi(\sigma)$$

is a nonnegative integer.

- b. Complete Lemma 6.21: If $\alpha = \beta\tau$, prove that

$$(e_\alpha^*, e_\beta^*) = \frac{\chi(e)}{o(G)} \sum_{\sigma \in G_\alpha} \chi(\sigma\tau).$$

14. Prove Lemma 6.24.
15. Let χ be an irreducible character of the subgroup G of S_m .
- Prove that $\Delta = \overline{\Delta}$ if and only if χ is the principal character.
 - Prove Theorem 6.30.
16. Let χ be an irreducible character of G . Let $\overline{\chi}$ be the character defined by $\overline{\chi}(\sigma) = \overline{\chi(\sigma)}$, $\sigma \in G$.
- Prove that $\overline{\chi} \in I(G)$.
 - Show that the $\overline{\Delta}$ set corresponding to G and χ is identical to the $\overline{\Delta}$ set arising from G and $\overline{\chi}$.
17. Suppose $\alpha \in \overline{\Delta}$. If $\{e_i : 1 \leq i \leq n\}$ is an orthonormal basis of V , prove that
- $\chi(e) \leq s_\alpha \leq \chi(e)^2$.
 - $s_\alpha \leq [G : G_\alpha]$.
 - equality holds in part (b) if and only if χ is the principal character and $G_\alpha = G$.
 - $\|e_\alpha^*\|^2 = s_\alpha/[G : G_\alpha]$.
18. Let $\{e_1, e_2, \dots, e_n\}$, $n > 1$, be a basis of the vector space V . Suppose G is a permutation group of degree m . For a fixed but arbitrary $\alpha \in \Gamma_{m,n}$ define $V_\alpha^\otimes = \langle e_{\alpha\tau}^\otimes : \tau \in G \rangle$.
- Prove that $V_\alpha^\otimes$ is an invariant subspace of $P(\sigma)$, $\sigma \in G$. That is, show that $P(\sigma)(w) \in V_\alpha^\otimes$ for all $w \in V_\alpha^\otimes$ and all $\sigma \in G$.
 - Denote by $P_\alpha(\sigma)$ the restriction of $P(\sigma)$ to $V_\alpha^\otimes$, $\sigma \in G$. Prove that P_α is a representation of G .
 - Show that the action of $P_\alpha(\sigma)$ on $\{e_{\alpha\tau}^\otimes : \tau \in G\}$ is identical to the action of $\overline{\sigma}$ on $\{\alpha\tau : \tau \in G\}$, where $\overline{\sigma}$ is defined in Definition 6.17.
 - Suppose G is a subgroup of S_m and $\chi \in I(G)$. Let $\alpha \in \overline{\Delta}$. Show that s_α , the dimension of the orbital subspace $\langle \{e_{\alpha\tau}^* : \tau \in G\} \rangle$, is $\chi(e)(\chi, \xi)_G$, where ξ is the character of G afforded by P_α .
 - Show that $\xi = 1_\alpha^G$, where 1_α is the principal character of G_α .
 - Use these facts, along with the Frobenius Reciprocity Theorem, to give another proof of Freese's Theorem.
 - Use part (e) to show that Equations (6.20) and (6.22) are equivalent.
19. Prove Lemma 6.52.

20. Let χ be an irreducible character of the subgroup G of S_m . Suppose $\alpha \in \Gamma_{m,n}$ is fixed but arbitrary. Let $B = \{e_i : 1 \leq i \leq n\}$ be a basis of V . Define $G^\alpha = \{\sigma \in G : \text{there exists a number, } \xi(\sigma), \text{ such that } e_{\alpha\sigma}^* = \xi(\sigma)e_\alpha^*\}$. Prove that
- G^α is a subgroup of G .
 - ξ is a linear character of G^α .
 - G_α is a subgroup of G^α .
 - when restricted to G_α , $\xi = 1$.
 - G^α does not depend on B .
 - $(\chi, 1)_{G_\alpha} = (\chi, \xi)_{G^\alpha}$.
21. Let $\{v_i : 1 \leq i \leq k\}$ be a linearly independent set of vectors in V . Suppose the vectors $u_1, u_2, \dots, u_k \in V$ satisfy the identity

$$\sum_{i=1}^k v_i \wedge u_i = 0.$$

Prove that $u_i \in \langle v_1, v_2, \dots, v_k \rangle$, $1 \leq i \leq k$. (Compare with Chapter 5, Exercise 7a.)

22. Prove Theorem 6.53 directly, without using Gamas's Theorem. (Hint: Lemma 6.16 and Lemma 6.52.)
23. Show how Gamas's Theorem can be used to prove Lemma 6.16.

24. Let $A = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 0 & 1 & 2 & 3 \\ 2 & 1 & 0 & 1 & 2 \\ 3 & 2 & 1 & 0 & 1 \\ 4 & 3 & 2 & 1 & 0 \end{pmatrix}$.

Write down $A[\alpha|\beta]$ and $A(\alpha|\beta)$ when

- $\alpha = (1, 2)$ and $\beta = (4, 5)$
 - $\alpha = (2, 3)$ and $\beta = (4, 5)$
 - $\alpha = (2, 4)$ and $\beta = (4, 5)$
 - $\alpha = (2, 4)$ and $\beta = (2, 4)$
25. Let $\pi = [r, 1^{m-r}]$, $G = S_m$, and $\chi = \chi_\pi$. If $\{v_1, v_2, \dots, v_m\}$ is a set of nonzero vectors in the vector space V , prove that $v_1 * v_2 * \dots * v_m = 0$ if and only if $\dim(\langle v_1, v_2, \dots, v_m \rangle) < m - r + 1$.
26. Let $G = D_4$. Use Figure 6.10 to show that the multilinearity partition
- $MP(\chi_1) = [4]$.
 - $MP(\chi_2) = [2, 2]$.
 - $MP(\chi_3) = [3, 1]$.

- d. $MP(\chi_4) = [2, 1, 1]$.
 e. $MP(\chi_5) = [3, 1]$.
27. ([Fonseca (1992)]) Let G be a subgroup of S_m . Let f be the function from $I(G)$ to the partitions of m defined by $f(\chi) = MP(\chi)$. Prove that f need be neither one-to-one nor onto. (Hint: Exercise 26.)
28. Suppose V is an inner product space of dimension n . Then (Equation (6.26)), $\dim(\wedge^m V) = C(n, m) = C(n, n - m) = \dim(\wedge^{n-m} V)$. Therefore, as vector spaces, $\wedge^m V$ and $\wedge^{n-m} V$ are isomorphic. The obvious question is whether there is some natural isomorphism that preserves decomposable symmetrized tensors. Let $E = \{e_1, e_2, \dots, e_n\}$ be an ordered orthonormal basis of V . Define a linear transformation $H_E^+ : \wedge^m V \rightarrow \wedge^{n-m} V$ by $H_E^+(e_\alpha^\wedge) = (-1)^{t(\alpha)} e_{\alpha^c}^\wedge$, $\alpha \in Q_{m,n}$, and linear extension, where $t(\alpha) = (1 + \alpha(1)) + (2 + \alpha(2)) + \dots + (m + \alpha(m))$ and $\alpha^c \in Q_{n-m,n}$ is the sequence complementary to α .¹⁶
- a. Prove that H_E^+ is invertible.
 b. Let $V = \mathbb{R}^3$, real three-dimensional space. Let E be the standard ordered basis of V , that is, $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, and $e_3 = (0, 0, 1)$. Suppose $m = 2$. If $u, v \in V$, prove that

$$H_E^+(u \wedge v) = u \times v,$$

the cross product of u and v . (Hint: Let $u = (u_1, u_2, u_3) = u_1 e_1 + u_2 e_2 + u_3 e_3$ and $v = v_1 e_1 + v_2 e_2 + v_3 e_3$. Express $u \wedge v$ in terms of the basis $\{e_\alpha^\wedge : \alpha \in Q_{2,3}\}$ and then apply the linear map H_E^+ .)

- c. Let E be the standard ordered basis of $V = \mathbb{R}^3$. If $u, v, w \in V$, prove that $u \wedge v \wedge w = (u \circ (v \times w)) e_1 \wedge e_2 \wedge e_3$.
- d. If u, v , and w are vectors in $\mathbb{R}^3$, prove that $|u \circ (v \times w)|$ is the volume of the parallelepiped they determine. (The exterior product, $e_1 \wedge e_2 \wedge \dots \wedge e_n$, is sometimes referred to as a *volume element*.)
- e. Suppose u, v , and w are vectors in $\mathbb{R}^3$. Without using H_E^+ , prove that $u \circ (v \times w) = 0$ if and only if $\{u, v, w\}$ is linearly dependent.
29. Prove Lemma 6.59.
30. Let V be a vector space. Prove that $V_1(S_m) \subset V_1(G)$, for every subgroup G of S_m .
31. Prove Theorem 6.64.
32. Finish the proof of Lemma 6.65 by explaining why the set $\mathcal{B}$, defined in Equation (6.38), spans the vector space $\mathbb{C}_m[x_1, x_2, \dots, x_n]$.

¹⁶It can be shown that the Hodge star mapping, H_E^+ , preserves decomposability. (See, e.g., [Marcus (1975), pp. 21–31].)

33. Let V be a vector space and suppose $T \in L(V, V)$ has rank r . Let χ be a linear character of the subgroup G of S_m .
- Prove that $\text{rank}(K(T)) = o(\Gamma_{m,r} \cap \overline{\Delta})$.
 - Suppose $G = S_m$ and $\chi = \varepsilon$. If $r < m$, prove that $K(T) = 0$.
34. Let V be an inner product space. If T is a unitary operator on V , prove that $K(T)$ is unitary.
35. Let V be an inner product space. Suppose $S, T \in L(V, V)$. If $S \geq T$, prove that $K(S) \geq K(T)$.
36. Let χ be a fixed but arbitrary irreducible character of the subgroup G of S_m . Let V be a vector space of dimension $n \geq m$.
- Show that $V_\chi(G)$ is an invariant subspace for $P(\sigma)$, $\sigma \in G$.
 - Denote by $P_\chi(\sigma)$ the restriction of $P(\sigma)$ to $V_\chi(G)$, $\sigma \in G$. Prove that $\sigma \rightarrow P_\chi(\sigma)$ is a representation of G .
 - If $\xi \in I(G)$, let

$$Z_\xi = \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \xi(\sigma) P_\chi(\sigma).$$

Prove that Z_χ is the identity operator on $V_\chi(G)$.

- Prove that Z_ξ is the zero operator on $V_\chi(G)$ if $\chi \neq \xi$.
- Prove that $(\xi, \eta)_G = 0$ if $\xi \neq \overline{\chi}$, where η is the character of G afforded by P_χ .
- Let A be a representation of G affording $\overline{\chi}$. Prove that P_χ is equivalent to the direct sum of A with itself $(\overline{\chi}, \eta)_G$ times.

Applications Exercises

37. Consider a regular pentagon whose vertices have been numbered consecutively 1–5.
- Show that, modulo the group $G = \langle (12345) \rangle$ of its plane rotations, the pentagon has 51 inequivalent red-white-blue vertex colorings.
 - Show that, modulo the group D_5 of all ten symmetries (Exercise 32, Chapter 3), it has 39 inequivalent red-white-blue vertex colorings.
 - If a fourth color becomes available show that, modulo D_5 , there are 136 inequivalent vertex colorings.
38. Suppose each face of a cube is (uniformly) painted red, white, or blue. There are $3^6 = 729$ different ways to do it. Let's consider two red-white-blue colored cubes to be equivalent if they can be rotated into positions that exhibit the same

coloring. Use Theorem 6.35 to show that of the 729 different colorings, only 57 are inequivalent. (Hint: Figure 3.4.)

39. A **binary code of length n** is a set of n -letter “words” constructed from the 2-letter “alphabet” $\{0, 1\}$.
- Show that there is a natural one-to-one correspondence between the $2^3 = 8$ three-letter binary words and the eight vertices of a cube in three-dimensional Euclidean space.
 - Show that a binary code of length 3 can be illustrated by means of a black-white coloring of the vertices of a cube.
 - Show that there are 2^8 different binary codes of length 3.
 - Two binary codes of length 3 are *equivalent* if their black-white illustrations can be rotated into positions that exhibit the same coloring. How many *inequivalent* binary codes of length 3 are there? (Hint: Figure 3.5.)
40. Suppose $\sigma = (12 \dots m) \in S_m$ is a full m -cycle. Let $G = \langle \sigma \rangle = \{\sigma^i : 1 \leq i \leq m\}$ be the cyclic group generated by σ . Prove that the cycle index polynomial

$$Z_G(y_1, y_2, \dots, y_m) = \frac{1}{m} \sum_{k|m} \varphi(k) y_k^{m/k},$$

where $\varphi(k)$ is the number of positive integers $i \leq k$ that are relatively prime to k .

41. Recall (Definition 1.23) that the m -th homogeneous symmetric function

$$H_m(x_1, x_2, \dots, x_n) = \sum_{\pi \vdash m} M_\pi(x_1, x_2, \dots, x_n)$$

- a. Prove that $H_m(x_1, x_2, \dots, x_n)$ is the result of substituting

$$y_i = \sum_{i=1}^n x_i^i$$

in the cycle index polynomial $Z_{S_m}(y_1, y_2, \dots, y_m)$.

- b. Prove that

$$C(m+n-1, m) = \frac{1}{m!} \sum_{\sigma \in S_m} n^{c(\sigma)},$$

where $c(\sigma)$ is the total number of cycles, including cycles of length 1, in the disjoint cycle factorization of σ .

42. Let G be the group of plane rotations of the regular pentagon (Exercise 37).

- a. Show that $W_G(r, w, b) = [(r + w + b)^5 + 4(r^5 + w^5 + b^5)]/5$.
 - b. Show that $W_G(r, w, b) = M_{[5]} + M_{[4,1]} + 2M_{[3,2]} + 4M_{[3,1^2]} + 6M_{[2^2,1]}$, where M_π is an abbreviation for the monomial symmetric function $M_\pi(r, w, b)$.
 - c. Draw a system of distinct representatives for the four color patterns of weight rw^3b .
 - d. Draw a system of distinct representatives for the six color patterns of weight rw^2b^2 .
43. Let D_5 be the group of all ten symmetries of the regular pentagon (Exercise 37).
- a. Show that

$$W_{D_5}(r, w, b) = ((r + w + b)^5 + 5(r + w + b)(r^2 + w^2 + b^2)^2 + 4(r^5 + w^5 + b^5))/10.$$

- b. Express $W_{D_5}(r, w, b)$ as a linear combination of monomial symmetric functions.
 - c. Draw a system of distinct representatives for the two color patterns of weight rw^3b .
 - d. Draw a system of distinct representatives for the four color patterns of weight rw^2b^2 .
44. Let G be the group of 12 rotational symmetries of the regular tetrahedron.
- a. Express $W_G(r, w, b, g)$ as a linear combination of monomial symmetric functions when G is viewed as a group of permutations of the faces of the tetrahedron.
 - b. Express $W_G(w, b)$ as a linear combination of monomial symmetric functions when G is viewed as a group of permutations of the *edges* of the tetrahedron.
45. Consider the group G of 24 rotational symmetries of the cube.
- a. If G is expressed as permutations of the faces of the cube, show that

$$Z_G(y_1, y_2, \dots, y_6) = \frac{1}{24} (y_1^6 + 3y_1^2y_2^2 + 6y_1^2y_4 + 6y_2^3 + 8y_3^2).$$

- b. Show that the pattern inventory for the 57 red-white-blue colorings of the faces of the cube (Exercise 38) is

$$W_G(r, w, b) = M_{[6]} + M_{[5,1]} + 2M_{[4,2]} + 2M_{[4,1^2]} + 2M_{[3^2]} + 3M_{[3,2,1]} + 6M_{[2^3]},$$

where M_π is an abbreviation for the monomial symmetric function $M_\pi(r, w, b)$.

- c. If G is expressed as permutations of the vertices of the cube, show that

$$Z_G(y_1, y_2, \dots, y_6) = \frac{1}{24} (y_1^8 + 9y_2^4 + 6y_4^2 + 8y_1^2 y_3^2).$$

- d. Recall that G is abstractly isomorphic to S_4 (Exercise 36, Chapter 3) and, therefore, to $S_4^{(2)}$. Compare and contrast your answers to parts (a) and (c) with Equation (6.59).
46. Prove Lemma 6.89.
47. Consider the nonisomorphic graphs on 5 vertices.
- Use Example 6.90 with Theorem 6.86 to obtain Equation (6.54).
 - Illustrate the six nonisomorphic graphs having five vertices and six edges.
 - Illustrate the six nonisomorphic graphs having five vertices and five edges.
 - Illustrate the six nonisomorphic graphs having five vertices and four edges.
48. Confirm Example 6.90.
49. Let $G = S_m$. Suppose $\pi \vdash m$. Let χ_π be the irreducible character of S_m corresponding to π . Show that, in the sense of Example 6.92, χ_π selects those patterns whose multiplicity partitions are majorized by π . (Hint: Theorem 6.37.)

	C_1	C_2	C_3	C_4	C_5
χ_1	1	-1	1	1	-1
χ_2	3	-1	-1	0	1
χ_3	2	0	2	-1	0
χ_4	3	1	-1	0	-1
χ_5	1	1	1	1	1

FIGURE 6.12 Character table of the group of rotational symmetries of the cube.

50. Let G be the group of 24 rotational symmetries of the cube. Because G is abstractly isomorphic to S_4 (Exercise 36, Chapter 3), it must have five conjugacy classes. From Figure 3.4, we see that the elements of G exhibit five different cycle structures. It must be, therefore, that two elements of G are conjugate in G if and only if they are conjugate in S_6 . So, the conjugacy classes of G are $C_1 = \{e\} = G \cap C_{[1^6]}$, $C_2 = G \cap C_{[2^3]}$, $C_3 = G \cap C_{[2^2, 1^2]}$,

$C_4 = G \cap C_{[3^2]}$, and $C_5 = G \cap C_{[4,1^2]}$. It is now an easy matter to transcribe the character table of S_4 (Exercise 36, Chapter 4) and obtain the character table of G shown in Figure 6.12.

- a. Show that $\chi_1 = \varepsilon$, the restriction of the alternating character of S_6 to G .
- b. Show that

$$Z_G^\varepsilon(y_1, y_2, \dots, y_6) = [y_1^6 - 6y_2^3 + 3y_1^2y_2^2 + 8y_3^2 - 6y_1^2y_4]/24.$$

(Hint: Exercise 45.)

- c. Show that, after substituting $y_i = r^i + w^i + b^i$, $1 \leq i \leq 6$, into $Z_G^\varepsilon(y_1, y_2, \dots, y_6)$, the result is

$$\sum_{\alpha \in \Delta} w(\alpha) = M_{[4,1^2]} + 2M_{[3^2]} + 3M_{[3,2,1]} + 3M_{[2^3]},$$

where M_π is the monomial symmetric function $M_\pi(r, w, b)$.

- d. If $\chi = \chi_2$, and $M_\pi = M_\pi(r, w, b)$, show that

$$W_G^\chi(r, w, b) = M_{[5,1]} + M_{[4,2]} + 4M_{[4,1^2]} + 2M_{[3^2]} + 7M_{[3,2,1]} + 9M_{[2^3]}.$$

- e. If $\chi = \chi_3$, and $M_\pi = M_\pi(r, w, b)$, show that

$$W_G^\chi(r, w, b) = M_{[5,1]} + 2M_{[4,2]} + 3M_{[4,1^2]} + 2M_{[3^2]} + 6M_{[3,2,1]} + 9M_{[2^3]}.$$

- f. If $\chi = \chi_4$, and $M_\pi = M_\pi(r, w, b)$, show that

$$W_G^\chi(r, w, b) = 2M_{[4,2]} + 3M_{[4,1^2]} + 2M_{[3^2]} + 7M_{[3,2,1]} + 12M_{[2^3]}.$$

51. Prove that $\sigma \rightarrow \bar{\sigma}$ (Definition 6.85) is an isomorphism from S_n onto $S_n^{(2)}$.
52. Derive Corollary 6.91 directly from Theorem 6.35.
53. Denote the cycle index polynomial of S_m by Z_m .
 - a. Prove that

$$Z_m = \sum \prod_{i=1}^m (y_i/i)^{k_i} / k_i!$$

where the sum is over all nonnegative integer sequences $k_1, k_2, \dots, k_m$ satisfying $k_1 + 2k_2 + 3k_3 + \dots + mk_m = m$.

b. Prove Cauchy's Identity:

$$\sum (1^{k_1} k_1! 2^{k_2} k_2! \dots m^{k_m} k_m!)^{-1} = 1,$$

where the sum is over all nonnegative integer sequences $k_1, k_2, \dots, k_m$ satisfying $k_1 + 2k_2 + 3k_3 + \dots + mk_m = 1$.

CHAPTER 7

Generalized Matrix Functions

In 1893, J. Hadamard published his celebrated theorem on determinants (Theorem 2.41): If $A \geq 0$, then

$$h(A) \geq \det(A),$$

where, recall, $h(A)$ is the product of the main diagonal entries of A . Suppose $A \geq 0$ is partitioned into blocks,

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad (7.1)$$

where A_{11} and A_{22} are square. It follows from Hadamard's Theorem that

$$h(A) \geq \det(A_{11}) \det(A_{22}), \quad A \geq 0. \quad (7.2)$$

In 1907, the first of what has become a glittering array of extensions, generalizations, and improvements of Hadamard's Theorem was obtained by E. Fischer:

$$\det(A_{11}) \det(A_{22}) \geq \det(A), \quad A \geq 0. \quad (7.3)$$

Incidental to his work on group representation theory, I. Schur introduced the following notion.

DEFINITION 7.1 Suppose χ is a character of the subgroup G of S_m . The **generalized matrix function** $d_\chi: C_{m,m} \rightarrow C$ is defined by

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

where $A = (a_{ij})$. When $G = S_m$ and $\chi = \chi_\pi$ for some $\pi \vdash m$, d_χ is called an **immanant**; the cumbersome expression d_{χ_π} is sometimes abbreviated d_π .

EXAMPLE 7.2 The determinant is the immanant afforded by $\chi = \varepsilon$. Thus, $\det(A) = d_\varepsilon(A) = d_{[1^m]}(A)$, $A \in \mathbb{C}_{m,m}$. If $G = S_m$ and $\chi = \chi_{[m]}$, the principal character, then $d_{[m]}$ is the permanent,

$$\text{per}(A) = \sum_{\sigma \in S_m} \prod_{i=1}^m a_{i\sigma(i)}.$$

When χ is the principal character of $G = \{e\}$, $d_\chi = h$, is Hadamard's function.

Suppose $p + q = m$, where p and q are positive integers. Let $H_1 = \{\sigma \in S_m : \sigma(i) = i, p < i \leq m\}$ and $H_2 = \{\sigma \in S_m : \sigma(i) = i, 1 \leq i \leq p\}$. Then H_1 is easily identified with S_p , and H_2 is isomorphic to S_q . Partition $A \in \mathbb{C}_{m,m}$ as in Equation (7.1), where $A_{11} \in \mathbb{C}_{p,p}$ and $A_{22} \in \mathbb{C}_{q,q}$. If $G = H_1 \times H_2$ and $\chi = \varepsilon$, then

$$d_\chi(A) = \det(A_{11}) \det(A_{22}).$$

If $G = H_1 \times H_2$ and $\chi = 1$, then

$$d_\chi(A) = \text{per}(A_{11}) \text{per}(A_{22}).$$

□

In 1918, Schur obtained the following dramatic generalization of Fischer's Inequality.

SCHUR'S INEQUALITY 7.3 Let χ be a character of the subgroup G of S_m . If $A \in \mathbb{C}_{m,m}$ is positive semidefinite hermitian, then

$$d_\chi(A) \geq \chi(e) \det(A).$$

DEFINITION 7.4 Suppose G is a subgroup of S_m . Let χ be an irreducible character of G . The corresponding normalized generalized matrix function is defined by $\bar{d}_\chi(A) = d_\chi(A)/\chi(e)$, $A \in \mathbb{C}_{m,m}$.

In view of Definition 7.4, and because any character of G is a sum of irreducible characters, Schur's Inequality is equivalent to

$$\bar{d}_\chi(A) \geq \det(A), \quad A \geq 0. \quad (7.4)$$

We begin the proof of Schur's Inequality by placing it in a more general setting. If $c: S_m \rightarrow \mathbb{C}$ is a fixed but arbitrary function, define

$$d_c(A) = \sum_{\sigma \in S_m} c(\sigma) \prod_{i=1}^m a_{i\sigma(i)},$$

Then the generalized matrix function afforded by G and χ is the d_c function corresponding to

$$c(\sigma) = \begin{cases} \chi(\sigma), & \text{if } \sigma \in G, \\ 0, & \text{if } \sigma \in S_m \setminus G. \end{cases} \quad (7.5)$$

(Called the “trivial extension” of χ to S_m , the c -function defined in Equation (7.5) typically is not a character of S_m .)

DEFINITION 7.5 Let $\mathcal{H}_m = \{A \in \mathbf{C}_{m,m} : A \geq 0\}$. Denote by $\mathcal{C}_m$ the set of functions $c: S_m \rightarrow \mathbb{C}$ such that $d_c(A) \geq 0$ for all $A \in \mathcal{H}_m$.

WATKINS'S THEOREM 7.6¹ If $c \in \mathcal{C}_m$ and $A \in \mathcal{H}_m$, then

$$d_c(A) \geq c(e) \det(A). \quad (7.6)$$

In other words, if $d_c(A) \geq 0$ for all $A \geq 0$, then $d_c(A) \geq c(e) \det(A)$ for all $A \geq 0$.

Proof If $A = (a_{ij})$ is singular then $\det(A) = 0$, and the conclusion follows from the definition of $\mathcal{C}_m$. If A is positive definite then, by the interlacing inequalities, each of its principal submatrices is positive definite. In particular, $a_{mm} > 0$. Denote by $E_{mm} \in \mathbf{C}_{m,m}$ the matrix whose only non zero entry is a 1 in position (m, m) . Let $r = \det(A) / \det(A(m|m))$, where $A(m|m)$ is the matrix obtained from A by deleting its last row and column. Then (Corollary 2.44), r is the largest positive real number such that $A \geq r E_{mm}$. Let $A_0 = A - r E_{mm}$ and define $A_x = A_0 + x E_{mm}$. If f is the (linear) function defined by

$$f(x) = d_c(A_x) - c(e) \det(A_x),$$

then

$$f'(x) = d_c(A(m|m) \oplus (1)) - c(e) \det(A(m|m)).$$

It follows by induction that $f'(x) \geq 0$ for all $x \geq 0$. Because $f(0) = d_c(A_0) \geq 0$, it must be that $f(x) \geq 0$ for all $x \geq 0$. In particular, $f(r) \geq 0$. This completes the proof because $A_r = A$. $\square$

In view of Watkins's Theorem, to prove Schur's Inequality it suffices to show that $d_\chi(A) \geq 0$, $A \in \mathcal{H}_m$. We will return to this point later. For the present, let $\{e_1, e_2, \dots, e_m\}$ be an orthonormal basis of the inner product space V . Let $W = \langle e_\tau^\otimes : \tau \in S_m \rangle$, the subspace of $V^{\otimes m}$ spanned by the tensors

$$e_{\tau(1)} \otimes e_{\tau(2)} \otimes \cdots \otimes e_{\tau(m)}, \quad \tau \in S_m.$$

Because $P(\sigma)(e_\tau^\otimes) = e_{\tau\sigma^{-1}}^\otimes$, for all $\sigma, \tau \in S_m$, W is an invariant subspace of $P(\sigma)$. Denote by $P_W(\sigma)$ the restriction of $P(\sigma)$ to W .

¹Theorem 7.6 first appeared in [Grone, Merris & Watkins (1988)].

LEMMA 7.7 *Let $R(\sigma)$ be the matrix representation of $P_W(\sigma)$ with respect to the lexicographically ordered basis $\{e_\tau^\otimes : \tau \in S_m\}$. Then $\sigma \rightarrow R(\sigma)$ is the (right) regular representation of S_m .*

For each function $c: S_m \rightarrow \mathbb{C}$, define

$$M_c = \sum_{\sigma \in S_m} c(\sigma^{-1}) R(\sigma).$$

Observe that M_c is an $m!$ -by- $m!$ matrix whose rows and columns are indexed by the permutations of S_m . The (μ, τ) -entry of M_c is

$$\sum_{\sigma \in S_m} c(\sigma^{-1}) \delta_{\mu, \tau \sigma^{-1}} = c(\tau^{-1} \mu). \quad (7.7)$$

In particular, $M_c \geq 0$ if and only if

$$\sum_{\mu, \tau \in S_m} \overline{x(\mu)} c(\tau^{-1} \mu) x(\tau) \geq 0,$$

for all $x: S_m \rightarrow \mathbb{C}$.

DEFINITION 7.8 The complex valued function c of S_m is **positive semidefinite** if

$$\sum_{\mu, \tau \in S_m} \overline{x(\mu)} c(\tau^{-1} \mu) x(\tau) \geq 0,$$

for all $x: S_m \rightarrow \mathbb{C}$. Denote by $\mathcal{C}_m^+$ the set of all positive semidefinite functions $c: S_m \rightarrow \mathbb{C}$.

LEMMA 7.9 *Let c be a complex valued function of S_m . Then $c \in \mathcal{C}_m^+$ if and only if there exists a function $b: S_m \rightarrow \mathbb{C}$ such that*

$$c(\sigma) = \sum_{\tau \in S_m} \overline{b(\tau)} b(\sigma \tau), \quad \sigma \in S_m. \quad (7.8)$$

Proof Suppose b is a complex valued function of S_m . Define $c: S_m \rightarrow \mathbb{C}$ by Equation (7.8). Because $R(\sigma)$ is a permutation matrix, $R(\sigma)^* = R(\sigma^{-1})$, and

$$M_b^* = \sum_{\tau \in S_m} \overline{b(\tau)} R(\tau).$$

Therefore,

$$\begin{aligned}
 M_b^* M_b &= \sum_{\tau, \sigma \in S_m} \overline{b(\tau)} b(\sigma^{-1}) R(\tau \sigma) \\
 &= \sum_{\sigma \in S_m} \left(\sum_{\tau \in S_m} \overline{b(\tau)} b(\sigma^{-1} \tau) \right) R(\sigma) \\
 &= \sum_{\sigma \in S_m} c(\sigma^{-1}) R(\sigma) \\
 &= M_c.
 \end{aligned}$$

Hence, $M_c \geq 0$, which is equivalent to $c \in \mathcal{C}_m^+$.

Conversely, suppose $M_c \geq 0$. Let U be an $m!$ -by- $m!$ unitary matrix such that

$$U^* R(\sigma) U = \text{diag} (R_1(\sigma), R_2(\sigma), \dots, R_k(\sigma)), \quad \sigma \in S_m,$$

where each R_i is irreducible and R_i is equivalent to R_j only if $R_i = R_j$. Recall that, up to equivalence, every irreducible representation of S_m occurs among the R_i with multiplicity equal to its degree. Without loss of generality, we may assume that R_i , $1 \leq i \leq p$, is a complete set of inequivalent irreducible representations of S_m . Let

$$S(\sigma) = \text{diag} (R_1(\sigma), R_2(\sigma), \dots, R_p(\sigma)), \quad \sigma \in S_m.$$

Because $\{R(\sigma): \sigma \in S_m\}$ is linearly independent, $\{U^* R(\sigma) U: \sigma \in S_m\}$ is linearly independent. As it merely eliminates redundancies, $\{S(\sigma): \sigma \in S_m\}$ is linearly independent. If the degree of R_i is n_i , $1 \leq i \leq p$, then, because $n_1^2 + n_2^2 + \dots + n_p^2 = m!$, $\{S(\sigma): \sigma \in S_m\}$ must be a basis for the direct sum $\mathbb{C}_{n_1, n_1} \oplus \mathbb{C}_{n_2, n_2} \oplus \dots \oplus \mathbb{C}_{n_p, n_p}$.

Because $U^* M_c U \geq 0$, its direct summands

$$M_i = \sum_{\sigma \in S_m} c(\sigma^{-1}) R_i(\sigma) \in \mathcal{H}_{n_i},$$

$1 \leq i \leq k$. Suppose $M_i = B_i^* B_i$, $1 \leq i \leq k$. If $j > p$ then $R_j = R_i$ for some $i \leq p$; set $B_j = B_i$. If $B = \text{diag} (B_1, B_2, \dots, B_k)$, then $B^* B = U^* M_c U$. Moreover, by our previous remarks, $D = \text{diag} (B_1, B_2, \dots, B_p)$ is in the space spanned by $\{S(\sigma): \sigma \in S_m\}$, that is, there exists a function $b: S_m \rightarrow \mathbb{C}$ such that

$$D = \sum_{\sigma \in S_m} b(\sigma^{-1}) S(\sigma).$$

It follows that

$$B = \sum_{\sigma \in S_m} b(\sigma^{-1}) U^* R(\sigma) U.$$

Therefore,

$$\begin{aligned} M_c &= U B^* U^* U B U^* \\ &= \left(\sum_{\sigma \in S_m} b(\sigma^{-1}) R(\sigma) \right)^* \left(\sum_{\sigma \in S_m} b(\sigma^{-1}) R(\sigma) \right) \\ &= \sum_{\sigma \in S_m} \left(\sum_{\tau \in S_m} \overline{b(\tau)} b(\sigma^{-1} \tau) \right) R(\sigma). \end{aligned}$$

Equation (7.8) now follows by another application of the linear independence of $\{R(\sigma) : \sigma \in S_m\}$. $\square$

DEFINITION 7.10 Suppose $A = (a_{i,j}) \in \mathbf{C}_{m,m}$. The **Schur Power** matrix, $\Pi(A)$, is the $m!$ -by- $m!$ matrix whose rows and columns are indexed by the lexicographically ordered sequences $(\tau(1), \tau(2), \dots, \tau(m))$, $\tau \in S_m$, and whose (μ, τ) -entry is

$$\prod_{i=1}^m a_{\mu(i), \tau(i)}.$$

Observe that $\Pi(A)$ is a principal submatrix of $A^{\otimes m}$. Hence,

$$A \geq 0 \implies \Pi(A) \geq 0.$$

THEOREM 7.11 $\mathcal{C}_m^+ \subset \mathcal{C}_m$.

Proof Let $A \in \mathcal{H}_m$ be fixed but arbitrary. Then $A^t \in \mathcal{H}_m$ and $\Pi(A^t) \in \mathcal{H}_{m!}$. If $c \in \mathcal{C}_m^+$, then $M_c \in \mathcal{H}_{m!}$. Consider the matrix M^c obtained from M_c by interchanging the rows and columns corresponding to τ and τ^{-1} , $\tau \in S_m$. Then the (μ, τ) -entry of M^c is $c(\tau \mu^{-1})$. Moreover, because M^c is permutation similar to M_c , $M^c \geq 0$. It follows from Example 5.29 that the Hadamard product

$M^c \cdot \Pi(A) \in \mathcal{H}_m!$. If x is the vector of all 1's, then

$$\begin{aligned}
 0 &\leq x^*(M^c \cdot \Pi(A))x \\
 &= \sum_{\mu, \tau \in S_m} c(\tau\mu^{-1}) \prod_{i=1}^m a_{\mu(i), \tau(i)} \\
 &= \sum_{\tau \in S_m} \left(\sum_{\mu \in S_m} c(\tau\mu^{-1}) \prod_{i=1}^m a_{i, \tau\mu^{-1}(i)} \right) \\
 &= \sum_{\tau \in S_m} \left(\sum_{\sigma \in S_m} c(\sigma) \prod_{i=1}^m a_{i, \sigma(i)} \right) \\
 &= m! d_c(A).
 \end{aligned}$$

□

It follows from Theorems 7.6 and 7.11 that $d_c(A) \geq c(e) \det(A)$, $c \in \mathcal{C}_m^+$, $A \in \mathcal{H}_m$.² An obvious question raised by Theorem 7.11 is whether $\mathcal{C}_m^+ = \mathcal{C}_m$.

EXAMPLE 7.12³ Define $c: S_4 \rightarrow \mathbb{C}$ by $c((12)(34)) = 1$, and $c(\sigma) = 0$, $\sigma \neq (12)(34)$. If $A = (a_{ij})$ is a 4-by-4 hermitian matrix, then $d_c(A) = |a_{12}|^2 |a_{34}|^2$. Thus, $c \in \mathcal{C}_4$. Define $x: S_4 \rightarrow \mathbb{C}$ by $x((12)) = 1$, $x((34)) = -1$, and $x(\sigma) = 0$, when $(12) \neq \sigma \neq (34)$. Then

$$\sum_{\mu, \tau \in S_4} \overline{x(\mu)} c(\tau^{-1}\mu) x(\tau) = -2,$$

so $c \notin \mathcal{C}_4^+$.

□

EXAMPLE 7.13 Let G be a subgroup of S_m . Suppose $\chi \in I(G)$. By Theorem 4.70, there is a unitary representation $\sigma \rightarrow B(\sigma) = (b_{ij}(\sigma))$, $\sigma \in G$, that affords χ . Fix i and let c be the trivial extension of b_{ii} to S_m , that is, $c: S_m \rightarrow \mathbb{C}$ is the function

$$c(\sigma) = \begin{cases} b_{ii}(\sigma), & \text{if } \sigma \in G, \\ 0, & \text{otherwise.} \end{cases}$$

²The case of equality has been analyzed in [Dias da Silva (1979)]. Also see [Chan (1990)].

³This example is taken from [Watkins (1993)].

Then, for any $x: S_m \rightarrow \mathbb{C}$,

$$\begin{aligned}
 \sum_{\mu, \tau \in S_m} \overline{x(\mu)} c(\tau^{-1} \mu) x(\tau) &= \sum_{\mu, \tau \in G} \overline{x(\mu)} \left(\sum_{j=1}^n b_{ij}(\tau^{-1}) b_{ji}(\mu) \right) x(\tau) \\
 &= \sum_{j=1}^m \left(\sum_{\mu \in G} \overline{x(\mu)} b_{ji}(\mu) \right) \left(\sum_{\tau \in G} \overline{b_{ji}(\tau)} x(\tau) \right) \\
 &= \sum_{j=1}^m \left| \sum_{\mu \in G} \overline{x(\mu)} b_{ji}(\mu) \right|^2 \\
 &\geq 0.
 \end{aligned}$$

Therefore, $c \in C_m^+$. □

COROLLARY 7.14 Suppose $\chi \in I(G)$. Let $\sigma \rightarrow B(\sigma) = (b_{ij}(\sigma))$ be a unitary representation of G that affords χ . Then, (defining $b_{ii}(\sigma) = 0$, $\sigma \notin G$)

$$d_{b_{ii}}(A) \geq \det(A), \quad A \in \mathcal{H}_m, \quad (7.9)$$

$$1 \leq i \leq \chi(e).$$

Proof This is an immediate consequence of Example 7.13, Theorem 7.11, Theorem 7.6, and the fact that $b_{ii}(e) = 1$. □

Observe that Schur's Inequality is obtained by summing both sides of Inequality (7.9) from $i = 1$ to $\chi(e)$. In other words, Inequality (7.4) is the statement that the $\chi(e)$ inequalities in (7.9) hold on the average.⁴

Let $A \in \mathbf{C}_{m,m}$. If $\varepsilon: S_m \rightarrow \mathbb{C}$ is the alternating character, then the entry in row μ of $\Pi(A)\varepsilon$ is

$$\begin{aligned}
 \sum_{\tau \in S_m} (\Pi(A)_{\mu\tau}) \varepsilon(\tau) &= \sum_{\tau \in S_m} \varepsilon(\tau) \prod_{i=1}^m a_{\mu(i), \tau(i)} \\
 &= \sum_{\tau \in S_m} \varepsilon(\tau) \prod_{i=1}^m a_{i, \tau \mu^{-1}(i)} \\
 &= \sum_{\sigma \in S_m} \varepsilon(\sigma \mu) \prod_{i=1}^m a_{i, \sigma(i)} \\
 &= \varepsilon(\mu) \det(A).
 \end{aligned}$$

It seems that $\det(A)$ is an eigenvalue of $\Pi(A)$ corresponding to the eigenvector ε .

⁴Partitions of the form $[2^p, 1^q]$, $2p+q=m$, arise when the Pauli Exclusion Principal is applied to the quantum mechanical spin functions for a system of electrons [Schensted (1976)]. The corresponding $d_{b_{ii}}$ functions are called *algebraic* [Poshusta (1991)] & [Poshusta & Kinghorn (1992)].

The next result, implicit in [Schur (1918)], was proved explicitly in [Bapat & Sunder (1986)]. Our proof follows [Pate (1994a)].

THEOREM 7.15 *If $A \geq 0$, then $\det(A)$ is the smallest eigenvalue of the Schur power matrix $\Pi(A)$.*

By Theorem 2.34, it suffices to show that $\det(A)$ is the infimum of the numerical range, $F(\Pi(A))$. In view of Theorems 7.6 and 7.11, this is a consequence of the following.

LEMMA 7.16 *Suppose $A = (a_{ij}) \in \mathcal{H}_m$. Then $r \in F(\Pi(A))$ if and only if there exists a function $c \in C_m^+$ such that $c(e) = 1$ and $r = d_c(A)$.*

Proof By definition, $r \in F(\Pi(A))$ if and only if there is a function $x: S_m \rightarrow \mathbb{C}$ such that $\|x\| = 1$ and

$$\begin{aligned}
 r &= x^* \Pi(A) x \\
 &= \sum_{\mu, \tau \in S_m} \overline{x(\mu)} \left(\prod_{i=1}^m a_{\mu(i), \tau(i)} \right) x(\tau) \\
 &= \sum_{\mu, \tau \in S_m} \overline{x(\mu)} x(\tau) \left(\prod_{i=1}^m a_{i, \tau \mu^{-1}(i)} \right) \\
 &= \sum_{\sigma \in S_m} \left(\sum_{\mu \in S_m} \overline{x(\mu)} x(\sigma \mu) \right) \prod_{i=1}^m a_{i, \sigma(i)} \\
 &= \sum_{\sigma \in S_m} c(\sigma) \prod_{i=1}^m a_{i, \sigma(i)} \\
 &= d_c(A),
 \end{aligned}$$

where (Lemma 7.9) $c \in C_m^+$ is defined by

$$c(\sigma) = \sum_{\mu \in S_m} \overline{x(\mu)} x(\sigma \mu), \quad \sigma \in S_m. \quad (7.10)$$

□

In addition to proving Theorem 7.15, Lemma 7.16 has the following interesting consequence.

LEMMA 7.17 *Suppose G is a subgroup of S_m . Let $\sigma \rightarrow B(\sigma) = (b_{ij}(\sigma))$ be an irreducible unitary representation of G . Fix i and extend b_{ii} to S_m by defining $b_{ii}(\sigma) = 0$, $\sigma \notin G$. Then $d_{b_{ii}}(A)$ lies in the field of values of $\Pi(A)$.*

Proof By Example 7.13, b_{ii} (extended) is an element of C_m^+ , and $b_{ii}(e) = 1$. $\square$

Observe that Lemma 7.17, together with Theorem 7.15 and Theorem 2.34 gives another proof of Corollary 7.14 and, therefore, of Schur's Inequality.

By the Toeplitz-Hausdorff Theorem,⁵ the field of values of any square complex matrix is a convex set. It follows, therefore, from Lemma 7.17 that any convex combination of $d_{b_{ii}}(A)$ is an element of $F(\Pi(A))$. In fact, as we now see, $F(\Pi(A))$ is equal to the "convex hull" of $d_{b_{ii}}(A)$, $1 \leq i \leq \deg(B)$, as B ranges over the irreducible unitary representations of S_m .

THEOREM 7.18 *Suppose $c \in C_m^+$ satisfies $c(e) = 1$. Then there exists a complete set of inequivalent, irreducible, unitary representations of S_m , $\sigma \rightarrow R_k(\sigma) = (r_{ij}^k(\sigma))$, $1 \leq k \leq p$, such that c is a convex combination of r_{ii}^k , $1 \leq i \leq n_k$, $1 \leq k \leq p$, where n_k is the degree of R_k .*

Proof If $c \in C_m^+$, then (Lemma 7.9) there exists a function $b: S_m \rightarrow \mathbb{C}$ such that

$$c(\sigma) = \sum_{\tau \in S_m} \overline{b(\tau)} b(\sigma\tau), \quad \sigma \in S_m.$$

Let $\sigma \rightarrow B_k(\sigma) = (b_{ij}^k(\sigma))$, $1 \leq k \leq p$, be a complete set of inequivalent irreducible representations of S_m which, by Theorem 4.70, we may take to be unitary. Denote the degree of B_k by n_k . Then, by Theorems 4.23 and 4.35, the $n_1^2 + n_2^2 + \cdots + n_p^2$ functions b_{ij}^k , $1 \leq i, j \leq n_k$, $1 \leq k \leq p$, are a basis for the vector space of complex valued functions of S_m . Therefore, there exist constants x_{ij}^k such that

$$b(\sigma) = \sum_{k=1}^p \sum_{i,j=1}^{n_k} x_{ij}^k b_{ij}^k(\sigma),$$

$\sigma \in S_m$. Hence,

$$\begin{aligned} c(\sigma) &= \sum_{\tau \in S_m} \left(\sum_{k=1}^p \sum_{i,j=1}^{n_k} \overline{x_{ij}^k} b_{ji}^k(\tau^{-1}) \right) \left(\sum_{r=1}^p \sum_{s,t=1}^{n_r} x_{st}^r b_{st}^r(\sigma\tau) \right) \\ &= \sum_{k,r=1}^p \sum_{i,j=1}^{n_k} \sum_{s,t=1}^{n_r} \overline{x_{ij}^k} x_{st}^r \left(\sum_{\tau \in S_m} b_{ji}^k(\tau^{-1}) b_{st}^r(\sigma\tau) \right) \\ &= \sum_{k=1}^p (m!/n_k) \sum_{i,j,s=1}^{n_k} x_{sj}^k \overline{x_{ij}^k} b_{si}^k(\sigma) \\ &= \sum_{k=1}^p \sum_{s,i=1}^{n_k} (X_k X_k^*)_{si} b_{si}^k(\sigma), \quad \sigma \in S_m, \end{aligned}$$

⁵See [Toeplitz (1918)] and/or [Hausdorff (1919)].

by the Schur Relations, where X_k is the n_k -by- n_k matrix whose (i, j) -entry is $\sqrt{m!/n_k}x_{ij}$. Replacing i with j and s with i ,

$$c(\sigma) = \sum_{k=1}^p \sum_{i,j=1}^{n_k} c_{ij}^k b_{ij}^k(\sigma), \quad \sigma \in S_m,$$

where $C_k = (c_{ij}^k) \geq 0$. Hence, there exists a unitary matrix $U_k = (u_{ij}^k)$ such that

$$c_{ij}^k = \sum_{t=1}^{n_k} u_{it}^k \lambda_t^k \bar{u}_{jt}^k,$$

where $\lambda_t^k \geq 0$, $1 \leq t \leq n_k$, are the eigenvalues of C_k . Substituting this expression into the previous equation yields

$$\begin{aligned} c(\sigma) &= \sum_{k=1}^p \sum_{t=1}^{n_k} \lambda_t^k \sum_{i,j=1}^{n_k} u_{it}^k b_{ij}^k(\sigma) \bar{u}_{jt}^k \\ &= \sum_{k=1}^p \sum_{t=1}^{n_k} \lambda_t^k (\bar{U}_k^* B_k(\sigma) \bar{U}_k)_{tt} \\ &= \sum_{k=1}^p \sum_{t=1}^{n_k} \lambda_t^k r_{tt}^k(\sigma), \quad \sigma \in S_m, \end{aligned}$$

where $\bar{U}_k^* B_k(\sigma) \bar{U}_k = R_k(\sigma) = (r_{ij}^k(\sigma))$. Because $c(e) = 1 = r_{ii}^k(e)$,

$$\sum_{k=1}^p \sum_{t=1}^{n_k} \lambda_t^k = 1.$$

□

Returning to Equations (7.2)–(7.4), one might wonder whether $h(A) \geq \bar{d}_\chi(A)$, $A \geq 0$. In fact,

$$\begin{aligned} \sum_{\chi \in I(G)} (\chi(e)^2 / o(G)) \bar{d}_\chi(A) &= \sum_{\sigma \in G} \frac{1}{o(G)} \left(\sum_{\chi \in I(G)} \chi(e) \chi(\sigma) \right) \prod_{t=1}^m a_{t\sigma(t)} \\ &= h(A), \end{aligned} \tag{7.11}$$

by the Orthogonality Relations of the Second Kind. Because

$$\sum_{\chi \in I(G)} \chi(e)^2 = o(G),$$

Equation (7.11) expresses $h(A)$ as a convex combination of $\bar{d}_\chi(A)$ as χ ranges over $I(G)$. Rather than an upper bound, $h(A)$ is a weighted average of $\bar{d}_\chi(A)$, $\chi \in I(G)$.

In 1963, M. Marcus published a Hadamard Theorem for Permanents,⁶ namely,

$$\text{per}(A) \geq h(A), \quad A \in \mathcal{H}_m.$$

The corresponding analog of Fischer's Inequality was proved three years later by E.H. Lieb:

$$\text{per}(A) \geq \text{per}(A_{11}) \text{per}(A_{22}), \quad A \in \mathcal{H}_m, \quad (7.12)$$

where A is partitioned as in Equation (7.1). These results led to the following conjecture, first published in [Lieb (1966)].⁷

PERMANENTAL DOMINANCE CONJECTURE 7.19 *Let χ be an irreducible character of the subgroup G of S_m . Then*

$$\text{per}(A) \geq \bar{d}_\chi(A), \quad A \in \mathcal{H}_m. \quad (7.13)$$

Among the first to suggest a permanental analog of Equation (7.9) was Ralph Freese.

CONJECTURE 7.20 *Suppose χ is an irreducible character of the subgroup G of S_m . Let $\sigma \rightarrow B(\sigma) = (b_{ij}(\sigma))$ be a unitary representation of G affording χ . Then, for $1 \leq i \leq \chi(e)$,*

$$\text{per}(A) \geq d_{b_{ii}}(A), \quad A \in \mathcal{H}_m. \quad (7.14)$$

Because it is obtained by averaging the inequalities in (7.14), the permanental dominance conjecture would follow from Conjecture 7.20. In view of Lemmas 7.16 and 7.17, Theorem 7.18, Theorem 2.34, and Exercise 9 (below), the following is equivalent to Conjecture 7.20:

SOULES'S CONJECTURE 7.21⁸ *If $A \geq 0$, then $\text{per}(A)$ is the largest eigenvalue of the Schur Power matrix, $\Pi(A)$.*

Following the publication of [Merris & Watkins (1985)] and [James & Liebeck (1987)], a substantial body of evidence supporting the permanental dominance conjecture has accumulated. In addition to results indicating that the permanent is "on top", a variety of other relations have been discovered. For example, suppose

⁶Marcus's Inequality has since been strengthened, e.g., in [Bapat (1991)] as corrected by [Chan (1993b)].

⁷Also see [Marcus & Minc (1965b)] and [Minc (1978) & (1983)].

⁸Conjecture 7.21 first appeared in [Soules (1966)]. (Also see [Soules (1983) & (1994)].)

$\pi = [2^r, 1^s] \vdash m, s \geq 2$. If $G = S_m$ and $\chi = \chi_\pi$ then [Grone, Merris & Watkins (1986)]

$$h(A) \geq \bar{d}_\chi(A), \quad A \in \mathcal{H}_m, \quad (7.15)$$

which, in view of the Hadamard Theorem for Permanents, implies $\text{per}(A) \geq \bar{d}_{[2^r, 1^s]}(A)$, $A \in \mathcal{H}_m$. From the perspective of Equation (7.11), the normalized immanants $\bar{d}_{[2^r, 1^s]}$, $s \geq 2$, are "below average".

Denote by $\bar{d}_r$ the normalized immanant corresponding to the single-hook character, $\chi_r = \chi_{[r, 1^{m-r}]}$. Then $\bar{d}_1 = \det$ and $\bar{d}_m = \text{per}$. In (1988), Peter Heyfron showed that these single-hook immanants form a chain from the determinant to the permanent.

HEYFRON'S THEOREM 7.22 *If $A \in \mathcal{H}_m$, then*

$$\det(A) = \bar{d}_1(A) \leq \bar{d}_2(A) \leq \cdots \leq \bar{d}_m(A) = \text{per}(A), \quad A \in \mathcal{H}_m. \quad (7.16)$$

It is natural to wonder where, in Heyfron's chain of inequalities, Hadamard's function might be found. That question was addressed in [Heyfron (1992)]:

THEOREM 7.23 *If $r \geq 1$, $m \geq r + 2$, and $m - 1 \geq (r - 1)^2$, then*

$$h(A) \geq \bar{d}_r(A), \quad A \in \mathcal{H}_m.$$

In 1992, Tom Pate obtained the following sweeping generalization of several previous results.

PATE'S THEOREM 7.24 *Let $\pi = [\pi_1, \pi_2, \dots, \pi_t] \vdash m$. Suppose $\pi_s > \max\{\pi_{s+1}, 1\}$. Let*

$$\rho = [\pi_1, \pi_2, \dots, \pi_{s-1}, \pi_s - 1, \pi_{s+1}, \dots, \pi_t, 1].$$

Then $\bar{d}_\pi(A) \geq \bar{d}_\rho(A)$, $A \in \mathcal{H}_m$.

In Pate's Theorem, the Ferrers diagram $F(\rho)$ is obtained from $F(\pi)$ by removing the last box from column π_s and placing it at the end of column 1. If $\pi = [7, 5^2, 3^2]$, for example, then $t = 5$. Legitimate values for s are $s = 1$ corresponding to $\rho = [6, 5^2, 3^2, 1]$; $s = 3$ corresponding to $\rho = [7, 5, 4, 3^2, 1]$; and $s = 5$ corresponding to $\rho = [7, 5^2, 3, 2, 1]$.

Among a growing collection of more specialized results are these: If $r \geq s + 1$, then [Heyfron (1991)] $\bar{d}_{[r+2, 2^{r-1}]}(A) \geq \bar{d}_{[r, 2^s]}(A)$, $A \in \mathcal{H}_m$. If $s > 1$ and $\pi_s = 2$, then [Pate (1994b)] $\bar{d}_\nu(A) \geq \bar{d}_\pi(A)$, $A \in \mathcal{H}_m$, where $\pi = [\pi_1, \pi_2, \dots, \pi_s, 1^t]$ and $\nu = [\pi_1 + 2, \pi_2, \dots, \pi_{s-1}, 1^t]$. If $L(\pi) = s$ and $2t \geq \pi_s > t$, then [Pate (1996)] $\bar{d}_\pi(A) \geq \bar{d}_\rho(A)$, $A \in \mathcal{H}_m$, where $\rho = [\pi_1, \pi_2, \dots, \pi_{s-1}, t, \pi_s - t]$. Finally, if p, q, r, s , and t are nonnegative integers, then [Pate (1996)]

$$\bar{d}_{[p, q, r, 2^s, 1^t]}(A) \leq \text{per}(A), \quad A \in \mathcal{H}_m. \quad (7.17)$$

Using these results and some *ad hoc* arguments, for example,

$$\bar{d}_{[3^4]}(A) \leq \text{per}(A), \quad A \in \mathcal{H}_{12}, \quad (7.18)$$

the permanent dominance conjecture has been confirmed for every irreducible character of S_m , $m < 14$ [Pate (1996)].

CONJECTURE 7.25 *Suppose the conjugate partition $\pi^* \neq \pi \vdash m$. If π majorizes π^* , then $d_\pi(A) \geq d_{\pi^*}(A)$, $A \in \mathcal{H}_m$.*

Because $\chi_\pi(e) = \chi_{\pi^*}(e)$, this conjecture could just as well have been stated for normalized immanants. It would imply $\bar{d}_{[(n+1)^*]}(A) \geq \bar{d}_{[\pi^{(n+1)}]}(A)$, $A \in \mathcal{H}_{n(n+1)}$, which, together with (7.17), would prove (7.18), eliminating the need for *ad hoc* arguments.

If $\pi \neq \pi^*$, then the restriction of χ_π to the alternating group remains irreducible. Because (Theorem 4.47) $\chi_{\pi^*} = \varepsilon \chi_\pi$, the generalized matrix function afforded by $G = A_m$ and $\chi = \chi_\pi$ is the mean of the immanants d_π and d_{π^*} . Consequently, in cases where the permanent dominance conjecture has been confirmed for d_π and d_{π^*} , it is valid for the alternating group and χ_π as well.

There has been very little progress on the permanent dominance conjecture when $G \neq S_m$ or A_m . Lieb's Inequality (7.12) establishes the conjecture when χ is the principal character of a Young subgroup. Another confirming instance is the following: Suppose $A = (A_{ij}) \in \mathbb{C}_{kn, kn}$ is partitioned into k^2 n -by- n submatrices A_{ij} . Let $B = (b_{ij})$ be the k -by- k matrix defined by $b_{ij} = \text{per}(A_{ij})$. Then $\text{per}(B) = d_\chi(A)$, where χ is the principal character of an appropriately chosen subgroup of S_{kn} . It was proved in [Pate (1982)] that, for each k , there exists a positive integer n_k such that if $n \geq n_k$, then

$$\text{per}(A) \geq \text{per}(\text{per}(A_{ij})), \quad (7.19)$$

for all real, symmetric, positive semidefinite, nk -by- nk matrices A .

We turn now to the intimate relationship between generalized matrix functions and decomposable symmetrized tensors.

THEOREM 7.26 *Let V be an m -dimensional inner product space. Suppose $u_1, u_2, \dots, u_m$ and $v_1, v_2, \dots, v_m$ are vectors in V . Let G be a subgroup of S_m and suppose $\chi \in I(G)$. If $A = (a_{ij})$ is the matrix defined by $a_{ij} = (u_i, v_j)$ then, with respect to the induced inner product in $V_\chi(G)$,*

$$d_\chi(A) = \frac{o(G)}{\chi(e)} (u_1 * u_2 * \dots * u_m, v_1 * v_2 * \dots * v_m).$$

Proof Because $T(G, \chi)$ is an orthogonal projection,

$$\begin{aligned}
 & (u_1 * u_2 * \cdots * u_m, v_1 * v_2 * \cdots * v_m) \\
 &= (u_1 \otimes u_2 \otimes \cdots \otimes u_m, T(G, \chi)(v_1 \otimes v_2 \otimes \cdots \otimes v_m)) \\
 &= \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \overline{\chi(\sigma)} \prod_{i=1}^m (u_i, v_{\sigma^{-1}(i)}). \\
 &= \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma) \prod_{i=1}^m (u_i, v_{\sigma(i)}). \\
 &= \frac{\chi(e)}{o(G)} d_\chi((u_i, v_j)).
 \end{aligned}$$

□

COROLLARY 7.27 *Let V be an m -dimensional inner product space. If $A = (a_{ij}) \in \mathcal{H}_m$, then there exist $v_1, v_2, \dots, v_m \in V$ such that $a_{ij} = (v_i, v_j)$, and $d_\chi(A) = o(G) \|v_1 * v_2 * \cdots * v_m\|^2 / \chi(e)$.*

Proof The existence of vectors $v_1, v_2, \dots, v_m \in V$ such that $a_{ij} = (v_i, v_j)$ was established in Exercise 41, Chapter 2. The rest follows from Theorem 7.26. □

Together with Watkins's Theorem, Corollary 7.27 gives another proof of Schur's Inequality.

COROLLARY 7.28 *Let $E = \{e_1, e_2, \dots, e_m\}$ be an orthonormal basis of the inner product space V . Suppose $T \in L(V, V)$. If A^t is the matrix representation of T with respect to E , then*

$$d_\chi(A) = \frac{o(G)}{\chi(e)} (K(T)(e_1 * e_2 * \cdots * e_m), e_1 * e_2 * \cdots * e_m),$$

where $K(T) \in L(V_\chi(G), V_\chi(G))$ is the induced transformation.

Proof Because the (i, j) -entry of A is $a_{ij} = (T(e_i), e_j)$, the result follows from Theorem 7.26 by setting $v_i = e_i$ and $u_i = T(e_i)$, $1 \leq i \leq m$. □

THEOREM 7.29 *Suppose $\chi \in I(G)$. If $A, B \in \mathbb{C}_{m,m}$, then*

$$d_\chi(A^* A) d_\chi(B^* B) \geq |d_\chi(B^* A)|^2. \quad (7.20)$$

Proof Let $E = \{e_1, e_2, \dots, e_m\}$ be an orthonormal basis of the inner product space V . Then there exist $S, T \in L(V, V)$ such that the matrix representation of S with respect to E is $[S]_E = A^t$ and $[T]_E = B^t$. Thus, $[ST]_E = [S]_E[T]_E = A^t B^t = (BA)^t$. By Corollary 7.28,

$$\begin{aligned} \frac{\chi(e)}{o(G)} d_\chi(BA) &= (K(ST)(e_1 * e_2 * \dots * e_m), e_1 * e_2 * \dots * e_m) \\ &= (K(S)K(T)(e_1 * e_2 * \dots * e_m), e_1 * e_2 * \dots * e_m) \\ &= (K(T)(e_1 * e_2 * \dots * e_m), K(S^*)(e_1 * e_2 * \dots * e_m)). \end{aligned}$$

Applying the Cauchy-Schwarz Inequality, we obtain

$$\begin{aligned} \left| \frac{\chi(e)}{o(G)} d_\chi(BA) \right|^2 &\leq \|K(T)(e_1 * e_2 * \dots * e_m)\|^2 \|K(S^*)(e_1 * e_2 * \dots * e_m)\|^2 \\ &= (K(T^*T)(e_1 * e_2 * \dots * e_m), e_1 * e_2 * \dots * e_m) \\ &\quad \times (K(SS^*)(e_1 * e_2 * \dots * e_m), e_1 * e_2 * \dots * e_m) \\ &= \frac{\chi(e)^2}{o(G)^2} d_\chi(BB^*) d_\chi(A^*A), \end{aligned}$$

which is equivalent to Inequality (7.20). $\square$

COROLLARY 7.30 Suppose $\chi \in I(G)$. If $A \in \mathbf{C}_{m,m}$, then

$$\chi(e) d_\chi(A^*A) \geq |d_\chi(A)|^2.$$

Proof Let $B = I_m$ in Equation (7.20). $\square$

Corollary 7.30, together with Watkins's Theorem, gives yet another proof of Schur's Inequality.

COROLLARY 7.31 Let G be a subgroup of S_m . Suppose $\chi \in I(G)$. If $A, B \in \mathcal{H}_m$, then

$$d_\chi(A + B) \geq d_\chi(A) + d_\chi(B). \quad (7.21)$$

Proof Let V be an inner product space of dimension m . If E is a fixed but arbitrary orthonormal basis of V , there exist unique linear operators $S, T \in L(V, V)$ such that $[S]_E = A^t$ and $[T]_E = B^t$. By Theorem 6.75(v), $K(S + T) \geq K(S) + K(T)$. Hence, the result is a consequence of Corollary 7.28. $\square$

Because d_χ is linear in χ , results like (7.21) remain valid for arbitrary characters of G .

COROLLARY 7.32 *Let χ be an irreducible character of the subgroup G of S_m . Suppose $A, B \in \mathcal{H}_m$.*

- (i) *If $A \geq B$, then $d_\chi(A) \geq d_\chi(B)$; and*
 (ii) *If $A \geq B$ and $d_\chi(A) = d_\chi(B) \neq 0$, then $A = B$.*

Proof By definition, $A \geq B$ means there is a $C \in \mathcal{H}_m$ such that $A = B + C$. By Corollary 7.31, $d_\chi(A) = d_\chi(B + C) \geq d_\chi(B) + d_\chi(C)$. Because (Corollary 7.30) $d_\chi(C) \geq 0$, part (i) is established. Part (ii) was proved in [Chan & Lim (1981)]. $\square$

THEOREM 7.33⁹ *Let χ be an irreducible character of the subgroup G of S_m . If $A, B \in \mathcal{H}_m$, then*

$$d_\chi((A + B)^{1/m}) \geq d_\chi(A^{1/m}) + d_\chi(B^{1/m}). \quad (7.22)$$

Proof Let V be an inner product space of dimension m . If E is a fixed but arbitrary orthonormal basis of V , there exist unique linear operators $S, T \in L(V, V)$ such that $[S]_E = A'$ and $[T]_E = B'$. Setting $n = m$, $S_i = S$, and $T_i = T$, $1 \leq i \leq m$, we have, from Theorem 5.40 and the definition of induced transformations, that

$$K([\theta S + (1 - \theta)T]^{1/m}) \geq \theta K(S^{1/m}) + (1 - \theta)K(T^{1/m}),$$

whenever $0 \leq \theta \leq 1$. Therefore (Corollary 7.28),

$$d_\chi([\theta A + (1 - \theta)B]^{1/m}) \geq \theta d_\chi(A^{1/m}) + (1 - \theta)d_\chi(B^{1/m}). \quad (7.23)$$

The result follows by setting $\theta = 1/2$. $\square$

CAUCHY-BINET THEOREM 7.34 *Let $A, B \in \mathbb{C}_{n,n}$. Suppose χ is an irreducible character of the subgroup G of S_m . If $\alpha, \beta \in \Omega = \{\gamma \in \Gamma_{m,n} : (\chi, 1)_{G_\gamma} \neq 0\}$, then*

$$d_\chi((AB)[\alpha|\beta]) = \frac{\chi(e)}{o(G)} \sum_{\gamma \in \Omega} d_\chi(A[\alpha|\gamma]) d_\chi(B[\gamma|\beta]). \quad (7.24)$$

Before proving Theorem 7.34, let's look at a special case. If $n = m$, $\alpha = \beta = (1, 2, \dots, m)$, $G = S_m$, and $\chi = \varepsilon$. Then Equation (7.24) becomes

$$\det(AB) = \frac{1}{m!} \sum_{\sigma \in S_m} \det(A[(1, 2, \dots, m)|\sigma]) \det(B[\sigma|(1, 2, \dots, m)]), \quad (7.25)$$

⁹When $G=S_m$ and $\chi=\varepsilon$, so that $d_\chi=\det$, Theorem 7.33 is the Minkowski Determinant Theorem [Minkowski (1905)]. When $d_\chi=\text{per}$, it was proved in [Ando (1981)] and, when A and B commute, in [Marcus & Pierce (1968)]. The version given here is from [Merris (1979)].

where $A[(1, 2, \dots, m)|\sigma] = (a_{i\sigma(j)})$, the matrix whose (i, j) -entry is the $(i, \sigma(j))$ -entry of A . In particular, $\det(A[(1, 2, \dots, m)|\sigma]) = \varepsilon(\sigma) \det(A)$. Similarly, $\det(B[\sigma|(1, 2, \dots, m)]) = \varepsilon(\sigma) \det(B)$. Thus, Equation (7.25) may be written

$$\begin{aligned} \det(AB) &= \frac{1}{m!} \sum_{\sigma \in S_m} \det(A) \det(B) \\ &= \det(A) \det(B). \end{aligned} \quad (7.26)$$

It will follow from the representation theory of the general linear group, $GL(m, \mathbb{C})$, that the determinant is the only multiplicative generalized matrix function.¹⁰ Thus, the natural generalization of Equation (7.26) is not $d_\chi(AB) = d_\chi(A)d_\chi(B)$; it is the Cauchy-Binet Theorem.¹¹

The proof of Theorem 7.34 depends on the following straight-forward extension of Corollary 7.28.

LEMMA 7.35 Suppose χ is an irreducible character of the subgroup G of S_m . Let $E = \{e_1, e_2, \dots, e_n\}$ be an orthonormal basis of the inner product space V . Let $T \in L(V, V)$ be the unique linear operator such that $[T]_E = A^t$. If $\alpha, \beta \in \Gamma_{m,n}$, then

$$d_\chi(A[\alpha|\beta]) = \frac{o(G)}{\chi(e)} \left(K(T)(e_\alpha^*), e_\beta^* \right).$$

Proof (of Theorem 7.34): Let $E = \{e_1, e_2, \dots, e_n\}$ be an orthonormal basis of the inner product space V . Let S and T be linear operators on V whose matrix representations with respect to E are A^t and B^t , respectively. Then

$$\begin{aligned} (\chi(e)/o(G))d_\chi((AB)[\alpha|\beta]) &= \left(K(TS)(e_\alpha^*), e_\beta^* \right) \\ &= \left(K(S)(e_\alpha^*), K(T^*)(e_\beta^*) \right) \\ &= \sum_{\gamma \in \Gamma_{m,n}} \left(K(S)(e_\alpha^*), e_\gamma^\otimes \right) \left(e_\gamma^\otimes, K(T^*)(e_\beta^*) \right), \end{aligned} \quad (7.27)$$

by Lemma 7.35 and Parseval's Identity. Because $T(G, \chi)$ is idempotent, hermitian, and commutes with both $S^{m\otimes}$ and $T^{m\otimes}$, we may replace $e_\gamma^\otimes$ with e_γ^* . Because

¹⁰See Exercise 2, Chapter 8.

¹¹The equation $d_\chi(AB) = d_\chi(A)d_\chi(B)$ has received some attention, e.g., in [Beasley (1969)], [Beasley & Cummings (1972), (1973), (1978), (1982), & (1992)], and [Wang (1974)].

$e_\gamma^* = 0$, $\gamma \notin \Omega$, another application of Lemma 7.35 yields

$$(\chi(e)/o(G))d_\chi((AB)[\alpha|\beta]) = \sum_{\gamma \in \Omega} (\chi(e)/o(G))^2 d_\chi(A[\alpha|\gamma]) \overline{d_\chi(B^*[\beta|\gamma])}.$$

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

$$d_\chi((AB)[\alpha|\beta]) = \frac{\chi(e)}{o(G)} \sum_{\gamma \in \Omega} d_\chi(A[\alpha|\gamma]) d_\chi(B[\gamma|\beta]).$$

□

Our next result is a useful variation on the Cauchy-Binet Theorem, valid when $\chi(e) = 1$.

COROLLARY 7.36 *Let χ be a linear character of the subgroup G of S_m . If $A, B \in \mathbb{C}_{n,n}$, then*

$$d_\chi((AB)[\alpha|\beta]) = \sum_{\gamma \in \Delta} \frac{1}{o(G_\gamma)} d_\chi(A[\alpha|\gamma]) d_\chi(B[\gamma|\beta]), \quad (7.28)$$

for all $\alpha, \beta \in \overline{\Delta}$.

Proof Replacing $e_\gamma^\otimes$ with e_γ^* on the right-hand side of Equation (7.27) produces

$$\begin{aligned} & \frac{1}{o(G)} d_\chi((AB)[\alpha|\beta]) \\ &= \sum_{\gamma \in \Gamma_{m,n}} \left(K(S)(e_\alpha^*), e_\gamma^* \right) \left(e_\gamma^*, K(T^*)(e_\beta^*) \right) \\ &= \sum_{\gamma \in \Delta} \frac{1}{o(G_\gamma)} \sum_{\sigma \in G} \left(K(S)(e_\alpha^*), e_{\gamma\sigma}^* \right) \left(e_{\gamma\sigma}^*, K(T^*)(e_\beta^*) \right) \\ &= \sum_{\gamma \in \Delta} \frac{1}{o(G_\gamma)} \sum_{\sigma \in G} \overline{\chi(\sigma)} \chi(\sigma) \left(K(S)(e_\alpha^*), e_\gamma^* \right) \left(e_\gamma^*, K(T^*)(e_\beta^*) \right) \\ &= o(G) \sum_{\gamma \in \Delta} \frac{1}{o(G_\gamma)} \left(K(S)(e_\alpha^*), e_\gamma^* \right) \left(e_\gamma^*, K(T^*)(e_\beta^*) \right), \end{aligned}$$

by Equation (6.7) and the fact that $|\chi(\sigma)| = 1$, $\sigma \in G$. Applying Lemma 7.35 (with $\chi(e) = 1$), we obtain the result. □

Having exploited Theorem 7.26 to obtain a variety of results for generalized matrix functions, we now reverse course and use generalized matrix functions as a tool to study symmetry classes of tensors. Suppose V is a vector space. Let χ be an irreducible character of the subgroup G of S_m . Suppose $u_1, u_2, \dots, u_m \in V$ are fixed but arbitrary vectors, not all zero. Choose a function $\alpha \in Q_{r,m}$ such that $\{u_{\alpha(k)} : 1 \leq k \leq r\}$ is a basis of $U = \langle u_i : 1 \leq i \leq m \rangle$. Let $A = (a_{ij}) \in \mathbb{C}_{m,r}$ be the unique matrix satisfying

$$u_i = \sum_{j=1}^r a_{ij} u_{\alpha(j)}, \quad 1 \leq i \leq m. \quad (7.29)$$

Then

$$\begin{aligned} & u_1 * u_2 * \cdots * u_m \\ &= \left(\sum_{j=1}^r a_{1j} u_{\alpha(j)} \right) * \left(\sum_{j=1}^r a_{2j} u_{\alpha(j)} \right) * \cdots * \left(\sum_{j=1}^r a_{mj} u_{\alpha(j)} \right) \\ &= \sum_{\gamma \in \Gamma_{m,r}} \left(\prod_{i=1}^m a_{i\gamma(i)} \right) u_{\alpha\gamma(1)} * u_{\alpha\gamma(2)} * \cdots * u_{\alpha\gamma(m)} \\ &= \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma^{-1}) \sum_{\gamma \in \Gamma_{m,r}} \left(\prod_{i=1}^m a_{i\gamma(i)} \right) u_{\alpha\gamma\sigma(1)} \otimes u_{\alpha\gamma\sigma(2)} \otimes \cdots \otimes u_{\alpha\gamma\sigma(m)} \\ &= \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma^{-1}) \sum_{\gamma \in \Gamma_{m,r}} \left(\prod_{i=1}^m a_{i\gamma\sigma^{-1}(i)} \right) u_{\alpha\gamma(1)} \otimes u_{\alpha\gamma(2)} \otimes \cdots \otimes u_{\alpha\gamma(m)} \\ &= \frac{\chi(e)}{o(G)} \sum_{\gamma \in \Gamma_{m,r}} d_{\chi}(A[\bullet|\gamma]) u_{\alpha\gamma}^{\otimes} \end{aligned}$$

where $A[\bullet|\gamma] = A[(1, 2, \dots, m)|\gamma]$. Because $\{u_{\alpha\gamma}^{\otimes} : \gamma \in \Gamma_{m,r}\}$ is a basis of $U^{\otimes m}$, we deduce that $u_1 * u_2 * \cdots * u_m = 0$ if and only if $d_{\chi}(A[\bullet|\gamma]) = 0$ for all $\gamma \in \Gamma_{m,r}$.

If $v_1, v_2, \dots, v_m \in U$, there exists a unique matrix $B = (b_{ij}) \in \mathbb{C}_{m,r}$ such that

$$v_i = \sum_{j=1}^r b_{ij} u_{\alpha(j)}, \quad 1 \leq i \leq m. \quad (7.30)$$

By the same arguments,

$$v_1 * v_2 * \cdots * v_m = \frac{\chi(e)}{o(G)} \sum_{\gamma \in \Gamma_{m,r}} d_\chi(B[\bullet|\gamma]) u_{\alpha\gamma}^\otimes.$$

Because (Lemma 6.15) $u_1 * u_2 * \cdots * u_m = v_1 * v_2 * \cdots * v_m \neq 0$ only if $\langle u_1, u_2, \dots, u_m \rangle = \langle v_1, v_2, \dots, v_m \rangle$, it follows that $u_1 * u_2 * \cdots * u_m = v_1 * v_2 * \cdots * v_m \neq 0$ if and only if $d_\chi(A[\bullet|\gamma]) = d_\chi(B[\bullet|\gamma])$, $\gamma \in \Gamma_{m,r}$. These observations are summarized in the following.

LEMMA 7.37 *Let V be a vector space. Suppose G is a permutation group of degree m and $\chi \in I(G)$. Let $u_1, u_2, \dots, u_m \in V$ be fixed but arbitrary vectors, not all zero. Choose $\alpha \in Q_{r,m}$ so that $\{u_{\alpha(k)} : 1 \leq k \leq r\}$ is a basis of $\langle u_i : 1 \leq i \leq m \rangle$. Let $A \in \mathbb{C}_{m,r}$ be the matrix defined in Equation (7.29). If $v_1, v_2, \dots, v_m \in V$, then*

$$u_1 * u_2 * \cdots * u_m = v_1 * v_2 * \cdots * v_m$$

*if and only if $d_\chi(A[\bullet|\gamma]) = d_\chi(B[\bullet|\gamma])$, $\gamma \in \Gamma_{m,r}$, where B is the matrix defined in Equation (7.30). In particular, $u_1 * u_2 * \cdots * u_m = 0$ if and only if $d_\chi(A[\bullet|\gamma]) = 0$, for all $\gamma \in \Gamma_{m,r}$.*

Lemma 7.37 does not so much solve the problem of equality of decomposable symmetrized tensors as state it in another form. Equivalent to Lemma 7.37, the next result seems to illuminate the situation from a somewhat different perspective.¹²

THEOREM 7.38 *Let V be a vector space. Suppose G is a permutation group of degree m and $\chi \in I(G)$. Let $u_1, u_2, \dots, u_m \in V$ be fixed but arbitrary vectors, not all zero. Choose $\alpha \in Q_{r,m}$ so that $\{u_{\alpha(k)} : 1 \leq k \leq r\}$ is a basis of $\langle u_i : 1 \leq i \leq m \rangle$. Let $A \in \mathbb{C}_{m,r}$ be the matrix defined in Equation (7.29). If $v_1, v_2, \dots, v_m \in V$, then*

$$u_1 * u_2 * \cdots * u_m = v_1 * v_2 * \cdots * v_m$$

*if and only if $d_\chi(AX) = d_\chi(BX)$, for all $X \in \mathbb{C}_{r,m}$, where B is the matrix defined in Equation (7.30). In particular, $u_1 * u_2 * \cdots * u_m = 0$ if and only if $d_\chi(AX) = 0$ for all $X \in \mathbb{C}_{r,m}$.*

¹²Lemma 7.37 and Theorem 7.38 are adaptations of more general results from [Marcus & Chollet (1983)] and [Oliveira, Santana & Dias da Silva (1983)]. Related work appears in [Marcus & Chollet (1980)], [Oliveira & Dias da Silva (1983) & (1985–86)], [Dias da Silva & Coelho (1990)], and [Duffner (1995)].

Proof If $X = (x_{ij}) \in \mathbf{C}_{r,m}$, then

$$\begin{aligned}
 d_X(AX) &= \sum_{\sigma \in G} \chi(\sigma) \prod_{i=1}^m \left(\sum_{t=1}^r a_{it} x_{t\sigma(i)} \right) \\
 &= \sum_{\sigma \in G} \chi(\sigma) \prod_{i=1}^m \left(\sum_{t=1}^r a_{\sigma^{-1}(i)t} x_{ti} \right) \\
 &= \sum_{\sigma \in G} \chi(\sigma) \sum_{\gamma \in \Gamma_{m,r}} \prod_{i=1}^m a_{\sigma^{-1}(i)\gamma(i)} \prod_{i=1}^m x_{\gamma(i)i} \\
 &= \sum_{\gamma \in \Gamma_{m,r}} \left(\sum_{\sigma \in G} \chi(\sigma) \prod_{i=1}^m a_{i\gamma\sigma(i)} \right) \prod_{i=1}^m x_{\gamma(i)i} \\
 &= \sum_{\gamma \in \Gamma_{m,r}} d_X(A[\bullet|\gamma]) \prod_{i=1}^m x_{\gamma(i)i}. \tag{7.31}
 \end{aligned}$$

If $u_1 * u_2 * \cdots * u_m = 0$ then, by Lemma 7.37 and Equation (7.31), $d_X(AX) = 0$ for all $X \in \mathbf{C}_{r,m}$. Conversely, suppose $d_X(AX) = 0$ for all $X \in \mathbf{C}_{r,m}$. Let $\beta \in \Gamma_{m,r}$ be fixed but arbitrary. Define $Y \in \mathbf{C}_{r,m}$ to be the matrix whose only nonzero entry in column i is a 1 in row $\beta(i)$, $1 \leq i \leq m$. Then, from Equation (7.31), $0 = d_X(AY) = d_X(A[\bullet|\beta])$, and $0 = u_1 * u_2 * \cdots * u_m$ by Lemma 7.37. The rest is similar. $\square$

Suppose $B = \{e_1, e_2, \dots, e_n\}$ is an orthonormal basis of the inner product space V . If χ is a linear character of the subgroup G of S_m then (Corollary 6.32 and Equation (6.12)), $\{[G : G_\alpha]^{1/2} e_\alpha^* : \alpha \in \bar{\Delta}\}$ is an orthonormal basis of $V_\chi(G)$.

DEFINITION 7.39 Let $B = \{e_1, e_2, \dots, e_n\}$ be an ordered orthonormal basis of the inner product space V . Suppose G is a subgroup of S_m and χ is a linear character of G . Denote by $B^\#$ the lexicographically ordered orthonormal basis $\{[G : G_\alpha]^{1/2} e_\alpha^* : \alpha \in \bar{\Delta}\}$ of $V_\chi(G)$.

THEOREM 7.40 Let χ be a linear character of the subgroup G of S_m . Suppose $B = \{e_1, e_2, \dots, e_n\}$ is an ordered orthonormal basis of the inner product space V . Let $T \in L(V, V)$. If $A = (a_{ij})$ is the matrix representation of T with respect to B , then the (α, β) -entry of the matrix representation of $K(T)$ with respect to $B^\#$ is

$$(o(G_\alpha) o(G_\beta))^{-1/2} \overline{d_{\bar{\chi}}(A[\alpha|\beta])},$$

where $\bar{\chi}(\sigma) = \overline{\chi(\sigma)}$, $\sigma \in G$.

Proof By definition,

$$T(e_j) = \sum_{i=1}^n a_{ij} e_i, \quad 1 \leq i \leq n.$$

Therefore,

$$\begin{aligned} & K(T)([G : G_\beta]^{1/2} e_\beta^*) \\ &= [G : G_\beta]^{1/2} (T(e_{\beta(1)}) * T(e_{\beta(2)}) * \cdots * T(e_{\beta(m)})) \\ &= [G : G_\beta]^{1/2} \left(\sum_{i=1}^n a_{i\beta(1)} e_i \right) * \left(\sum_{i=1}^n a_{i\beta(2)} e_i \right) * \cdots * \left(\sum_{i=1}^n a_{i\beta(m)} e_i \right) \\ &= [G : G_\beta]^{1/2} \sum_{\alpha \in \Gamma_{m,n}} \left(\prod_{i=1}^m a_{\alpha(i)\beta(i)} \right) e_\alpha^* \\ &= [G : G_\beta]^{1/2} \sum_{\alpha \in \bar{\Delta}} \frac{1}{o(G_\alpha)} \sum_{\sigma \in G} \left(\prod_{i=1}^m a_{\alpha\sigma^{-1}(i)\beta(i)} \right) e_{\alpha\sigma^{-1}}^* \\ &= [G : G_\beta]^{1/2} \sum_{\alpha \in \bar{\Delta}} \frac{1}{o(G_\alpha)} \left(\sum_{\sigma \in G} \chi(\sigma^{-1}) \prod_{i=1}^m a_{\alpha(i)\beta\sigma(i)} \right) e_\alpha^* \\ &= \sum_{\alpha \in \bar{\Delta}} ([o(G_\alpha)o(G_\beta)]^{-1/2} d_{\bar{\chi}}(A[\alpha|\beta])) [G : G_\alpha]^{1/2} e_\alpha^*. \end{aligned}$$

□

DEFINITION 7.41 Suppose $A \in \mathbb{C}_{n,n}$. Let $\bar{\Delta} \subset \Gamma_{m,n}$ be the index set associated with the linear character χ of the subgroup G of S_m . The rows and columns of the corresponding induced matrix, $K(A)$, are indexed by $\bar{\Delta}$, ordered lexicographically. The (α, β) -entry of $K(A)$

$$(o(G_\alpha)o(G_\beta))^{-1/2} d_{\bar{\chi}}(A[\alpha|\beta]).$$

If A is the matrix representation of $T \in L(V, V)$ with respect to the orthonormal basis B of V , then $K(A)$ is the matrix representation of $K(T) \in L(V_\chi(G), V_\chi(G))$ with respect to B^* .

COROLLARY 7.42 Let χ be a linear character of the subgroup G of S_m . If A and $B \in \mathbb{C}_{n,n}$, then $K(AB) = K(A)K(B)$.

Proof In view of Theorem 7.40 and Definition 7.41, this is just the matrix version of Theorem 6.75(i). $\square$

Comparing the (α, β) -entries of $K(AB)$ and $K(A)K(B)$ yields another proof of Corollary 7.36.

DEFINITION 7.43 Let $G = S_m$. If $\chi = \varepsilon$, $K(A)$ is called the m -th **compound** of A and is denoted $C_m(A)$. If χ is the principal character, $K(A)$ is called the m -th **induced power** of A , denoted $P_m(A)$.

Suppose $A \in \mathbb{C}_{n,n}$. If $G = S_m$ and $\chi = \varepsilon$, then $\bar{\Delta} = Q_{m,n}$. Because $G_\alpha = \{e\}$, $\alpha \in Q_{m,n}$, and $\bar{\varepsilon} = \varepsilon$, $C_m(A)$ is the $\binom{n}{m}$ -square matrix whose (α, β) -entry is $\det(A[\alpha|\beta])$. When $G = S_m$ and χ is the principal character, $\bar{\Delta} = G_{m,n}$, the set consisting of all $C(n+m-1, m)$ nondecreasing sequences/functions.¹³ If $\alpha \in G_{m,n}$, then G_α is isomorphic to the Young subgroup $S_{\mu(\alpha)}$, and $o(G_\alpha)$ is the product of the factorials of the multiplicities of the distinct integers appearing in α .

EXAMPLE 7.44 Let χ be the principal character of $G = S_2$. If $n = m = 2$, then $\Delta = \bar{\Delta} = G_{2,2} = \{(1, 1), (1, 2), (2, 2)\}$. If $\alpha = (1, 1)$ or $(2, 2)$, then $o(G_\alpha) = 2$; if $\alpha = (1, 2)$, then $o(G_\alpha) = 1$. Suppose

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is a generic 2-by-2 matrix. Then

$$\begin{aligned} P_2(A) &= \left(\frac{\text{per}(A[\alpha|\beta])}{\sqrt{o(G_\alpha)o(G_\beta)}} \right) \\ &= \begin{pmatrix} a^2 & \sqrt{2}ab & b^2 \\ \sqrt{2}ac & ad+bc & \sqrt{2}bd \\ c^2 & \sqrt{2}cd & d^2 \end{pmatrix}. \end{aligned} \quad (7.32)$$

Suppose $m = 2$, $n = 4$, $G = S_2$, and $\chi = \varepsilon$. Then $\bar{\Delta} = Q_{2,4} = \{(1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (3, 4)\}$, and $o(G_\alpha) = 1$ for all $\alpha \in \bar{\Delta}$. If

$$A = \begin{pmatrix} 4 & 1 & 3 & 3 \\ 1 & 2 & 1 & 1 \\ 5 & 3 & 4 & 5 \\ -6 & -4 & -4 & -5 \end{pmatrix},$$

¹³Note that $G_{m,n}$ is a set of sequences, while G and G_α are groups.

then

$$C_2(A) = \begin{pmatrix} 7 & 1 & 1 & -5 & -5 & 0 \\ 7 & 1 & 5 & -5 & -4 & 3 \\ -10 & 2 & -2 & 8 & 7 & -3 \\ -7 & -1 & 0 & 5 & 7 & 1 \\ 8 & 2 & 1 & -4 & -6 & -1 \\ -2 & 4 & 5 & 4 & 5 & 0 \end{pmatrix}. \quad (7.33)$$

□

DEFINITION 7.45 Suppose $A \in \mathbb{C}_{n,n}$. The principal submatrix of $P_m(A)$ corresponding to the sequences in $\mathcal{Q}_{m,n}$ is the m -th **permanental compound** of A . The $(n-1)$ -st permanental compound is the **permanental adjoint** of A , denoted $\text{p adj}(A)$.

The permanental adjoint is an n -by- n matrix, ostensibly indexed by the sequences $(1, 2, \dots, n-1)$, $(1, 2, \dots, n-2, n)$, and so on, ordered lexicographically. In fact, however, it is more conventional to index the rows and columns of $\text{p adj}(A)$ with the integers $1, 2, \dots, n$. Thus, $\text{p adj}(A)$ is the n -by- n matrix whose (i, j) -entry is $\text{per}(A(i|j))$.¹⁴

JACOBI'S IDENTITY 7.46 Let $A \in \mathbb{C}_{n,n}$. Suppose $m \leq n$. If A is invertible, then

$$\det(A) \det(A^{-1}[\alpha|\beta]) = (-1)^{r(\alpha)+r(\beta)} \det(A(\beta|\alpha)), \quad (7.34)$$

$\alpha, \beta \in \mathcal{Q}_{m,n}$, where $r(\alpha) = \alpha(1) + \alpha(2) + \dots + \alpha(m)$.

Proof Let $J_m(A)$ be the $\binom{n}{m}$ -by- $\binom{n}{m}$ matrix whose (α, β) -entry is $(-1)^{r(\alpha)+r(\beta)} \det(A(\beta|\alpha))$, $\alpha, \beta \in \mathcal{Q}_{m,n}$. (Up to permutation similarity, $J_1(A)$ is the classical adjoint (or adjugate) of A .) By the Laplace Expansion Theorem, the (α, β) -entry of $C_m(A)J_m(A)$ is $\det(A)$, when $\beta = \alpha$, and the determinant of a matrix with two equal rows when $\beta \neq \alpha$. Therefore,

$$\frac{1}{\det(A)} J_m(A) = C_m(A)^{-1}.$$

By Corollary 7.42, $C_m(A)^{-1} = C_m(A^{-1})$. Because inverses are unique, $\det(A)C_m(A^{-1}) = J_m(A)$. □

EXAMPLE 7.47 Suppose $n = 3$ and $m = 2$. If $A = (a_{ij}) \in \mathbb{C}_{3,3}$ is invertible, then

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} \det(A(1|1)) & -\det(A(2|1)) & \det(A(3|1)) \\ -\det(A(1|2)) & \det(A(2|2)) & -\det(A(3|2)) \\ \det(A(1|3)) & -\det(A(2|3)) & \det(A(3|3)) \end{pmatrix}.$$

¹⁴The permanental adjoint occurs, e.g., in [Ando (1981)], [Cohen (1992)], [Grone, Johnson, Sa, & Wolkowicz (1986)], [Lal (1992)] and [Marcus & Merris (1973)].

If $\alpha = (1, 2)$ and $\beta = (1, 3)$, then

$$A^{-1}[\alpha|\beta] = \frac{1}{\det(A)} \begin{pmatrix} \det(A(1|1)) & \det(A(3|1)) \\ -\det(A(1|2)) & -\det(A(3|2)) \end{pmatrix}$$

and

$$\begin{aligned} \det(A) \det(A^{-1}[\alpha|\beta]) &= (-\det(A(1|1)) \det(A(3|2)) \\ &\quad + \det(A(3|1)) \det(A(1|2))) / \det(A). \end{aligned}$$

On the other hand, from Equation (7.34),

$$\begin{aligned} \det(A) \det(A^{-1}[\alpha|\beta]) &= -\det(A(\beta|\alpha)) \\ &= -a_{23}, \end{aligned}$$

because $r(\alpha) + r(\beta) = (1 + 2) + (1 + 3) = 7$. Therefore,

$$-a_{23} \det(A) = -\det(A(1|1)) \det(A(3|2)) + \det(A(3|1)) \det(A(1|2)). \quad (7.35)$$

□

The careful reader will have observed that $K(A)$ has not been defined in general, but only for linear characters. This is because the description of $\hat{\Delta}$ given in Equation (6.19) does not lend itself to the kind of approach used to prove Theorem 7.40. The difficulties will become more apparent as we discuss the eigenvalues of $K(T)$.

Let $B = \{e_1, e_2, \dots, e_n\}$ be an ordered basis of the vector space V , and χ a fixed but arbitrary irreducible character of the subgroup G of S_m . Let α be the first element in the lexicographically ordered $\bar{\Delta}$ set. With $\alpha_1 = \alpha$, choose a set $\{\alpha_1, \alpha_2, \dots, \alpha_{s_\alpha}\}$ from $\{\alpha\sigma : \sigma \in G\}$ such that $\{e_{\alpha_i}^* : 1 \leq i \leq s_\alpha\}$ is a basis for $\langle e_\alpha^* : \alpha \in G \rangle$. Execute this procedure for each $\gamma \in \bar{\Delta}$. If $\{\alpha, \beta, \dots\}$ is the lexicographically ordered $\bar{\Delta}$ set, take

$$\hat{\Delta} = \{\alpha_1, \alpha_2, \dots, \alpha_{s_\alpha}, \beta_1, \beta_2, \dots, \beta_{s_\beta}, \dots\} \quad (7.36)$$

to be ordered as indicated. By Theorem 6.31,

$$B^* = \{e_\gamma^* : \gamma \in \hat{\Delta}\} \quad (7.37)$$

is an ordered basis of $V_\chi(G)$. (Because of the latitude in choosing $\hat{\Delta}$ when $\chi(e) > 1$, B^* may not be uniquely determined by m, n, G , and χ .)

LEMMA 7.48 *Let χ be an irreducible character of the subgroup G of S_m . Suppose $B = \{e_1, e_2, \dots, e_n\}$ is an ordered basis of the vector space V . Let B^* be the basis of $V_\chi(G)$ assembled as in Equation (7.37). If B is an upper triangular basis for $T \in L(V, V)$ (that is, a basis with respect to which the matrix representation of T is upper triangular), then B^* is an upper triangular basis for $K(T)$.*

Proof By hypothesis, $T(e_1) = \lambda_1 e_1$, and $T(e_i) = \lambda_i e_i + u_i$, where $u_i \in \langle e_1, e_2, \dots, e_{i-1} \rangle$, $1 < i \leq n$. Set $u_1 = 0$. Suppose ω is a fixed but arbitrary element of $\hat{\Delta}$. Let $\tau \in G$ be such that $\omega\tau \in \bar{\Delta}$. Then

$$\begin{aligned} K(T)(e_\omega^*) &= T(e_{\omega(1)}) * \cdots * T(e_{\omega(m)}) \\ &= (\lambda_{\omega(1)} e_{\omega(1)} + u_{\omega(1)}) * \cdots * (\lambda_{\omega(m)} e_{\omega(m)} + u_{\omega(m)}) \\ &= \left(\prod_{i=1}^m \lambda_{\omega(i)} \right) e_\omega^* + \sum c_\gamma e_\gamma^*, \end{aligned} \quad (7.38)$$

where the sum is over those $\gamma \in \Omega$ such that $\gamma(i) \leq \omega(i)$, $1 \leq i \leq m$, with at least one strict inequality. Fix γ and suppose it is equivalent (mod G) to $\gamma' \in \bar{\Delta}$. Because $\gamma\tau(i) \leq \omega\tau(i)$ for all i , with at least one strict inequality, γ' comes strictly before $\omega\tau$ in lexicographic order. Therefore e_γ^* is a linear combination of tensors that come strictly before e_ω^* in the ordered basis B^* . $\square$

THEOREM 7.49 *Let χ be an irreducible character of the subgroup G of S_m . Let V be a vector space of dimension n . Suppose $T \in L(V, V)$ has eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ (multiplicities included). Then the eigenvalues of $K(T)$ are*

$$\prod_{i=1}^m \lambda_{\omega(i)}, \quad \omega \in \hat{\Delta}.$$

Proof By Schur's Triangularization Theorem, there exists an upper triangular basis for T . Thus, the eigenvalues of $K(T)$ are visible in Equation (7.38). $\square$

It follows from Theorem 7.49 that the trace of $K(T)$ is

$$\begin{aligned} \sum_{\alpha \in \hat{\Delta}} \prod_{i=1}^m \lambda_{\alpha(i)} &= \chi(e) \sum_{\alpha \in \bar{\Delta}} (\chi, 1)_{G_\alpha} \prod_{i=1}^m \lambda_{\alpha(i)} \\ &= \chi(e) \sum_{\alpha \in \bar{\Delta}} (\chi, 1)_{G_\alpha} \prod_{i=1}^m \lambda_{\alpha(i)}. \end{aligned} \quad (7.39)$$

If $G = S_m$, so that $\chi = \chi_\pi$ for some partition $\pi \vdash m$, then G_α is conjugate to the Young subgroup S_μ , where $\mu = \mu(\alpha)$ is the multiplicity partition of α , and

$(\chi, 1)_{S_\mu}$ is the Kostka coefficient $K_{\pi, \mu}$. In this case, Equation (7.39) becomes

$$\text{tr}(K(T)) = \chi(e) \sum_{\alpha \in G_{m,n}} K_{\pi, \mu(\alpha)} \prod_{i=1}^m \lambda_{\alpha(i)}, \quad (7.40)$$

the product of the character degree and the “Schur polynomial”¹⁵ associated with π . If $G = S_m$ and $\pi = [1^m]$, so that $\chi_\pi = \varepsilon$, then $\bar{\Delta} = Q_{m,n}$. Because $G_\alpha = \{e\}$, $\alpha \in Q_{m,n}$, it follows from Equation (7.39) that

$$\begin{aligned} \text{tr}(C_m(A)) &= \sum_{\alpha \in Q_{m,n}} \prod_{i=1}^m \lambda_{\alpha(i)} \\ &= E_m(\lambda_1, \lambda_2, \dots, \lambda_n), \end{aligned} \quad (7.41)$$

the m -th elementary symmetric function of the eigenvalues of A . If $G = S_m$ and $\pi = [m]$, so that χ_π is the principal character, then $\bar{\Delta} = \Delta = G_{m,n}$, and

$$\begin{aligned} \text{tr}(P_m(A)) &= \sum_{\alpha \in G_{m,n}} \prod_{i=1}^m \lambda_{\alpha(i)} \\ &= H_m(\lambda_1, \lambda_2, \dots, \lambda_n), \end{aligned} \quad (7.42)$$

the m -th homogeneous symmetric function of the eigenvalues of A .

EXAMPLE 7.50 Returning to Example 7.44, suppose λ_1 and λ_2 are the eigenvalues of the 2-by-2 matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Then the eigenvalues of $P_2(A)$ are λ_1^2 , $\lambda_1\lambda_2$, and λ_2^2 . Therefore, from Equation (7.32),

$$\begin{aligned} a^2 + ad + bc + d^2 &= \text{tr}(P_2(A)) \\ &= H_2(\lambda_1, \lambda_2) \\ &= \lambda_1^2 + \lambda_1\lambda_2 + \lambda_2^2. \end{aligned}$$

This identity is easy to confirm: $\lambda_1\lambda_2 = \det(A) = ad - bc$ and $\lambda_1^2 + \lambda_2^2 = \text{tr}(A^2) = a^2 + 2bc + d^2$.

¹⁵See Equation (6.44).

The eigenvalues of

$$A = \begin{pmatrix} 4 & 1 & 3 & 3 \\ 1 & 2 & 1 & 1 \\ 5 & 3 & 4 & 5 \\ -6 & -4 & -4 & -5 \end{pmatrix}$$

are $-1, 1, 2$, and 3 . Therefore, $\text{tr}(C_2(A)) = E_2(-1, 1, 2, 3) = 5$. (Compare with the value obtained by summing the main diagonal entries of Equation (7.33).) $\square$

LEMMA 7.51 Suppose χ is an irreducible character of the subgroup G of S_m . Let V be a vector space of dimension n . Suppose $T \in L(V, V)$ has eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ (multiplicities included). Then $\text{tr}(K(T))$ is a symmetric function of $\lambda_1, \lambda_2, \dots, \lambda_n$.

Recall (Definition 6.66) that $m_t(\alpha)$ is the multiplicity of the integer t in the sequence $\alpha \in \Gamma_{m,n}$. In this notation,

$$\prod_{t=1}^m \lambda_{\alpha(t)} = \prod_{t=1}^n \lambda_t^{m_t(\alpha)}.$$

Let G be a subgroup of S_m , and suppose $\chi \in I(G)$. Define

$$e_t(\chi) = \sum_{\alpha \in \hat{\Delta}} m_t(\alpha),$$

$1 \leq t \leq n$. It follows from Lemma 7.51 that $e_1(\chi) = e_2(\chi) = \dots = e_n(\chi)$. Denote their common value by $e(\chi)$. Because $m_1(\alpha) + m_2(\alpha) + \dots + m_n(\alpha) = m$, $\alpha \in \hat{\Delta}$, it must be that $ne(\chi) = mo(\hat{\Delta})$.

SYLVESTER-FRANKE THEOREM 7.52 Suppose χ is an irreducible character of the subgroup G of S_m . Let V be an n -dimensional vector space. If $T \in L(V, V)$, then $\det(K(T)) = \det(T)^{e(\chi)}$, where $e(\chi) = mo(\hat{\Delta})/n$.

Proof Denote the eigenvalues of T by $\lambda_1, \lambda_2, \dots, \lambda_n$ (multiplicities included). It follows from Theorem 7.49 that

$$\begin{aligned} \det(K(T)) &= \prod_{\alpha \in \hat{\Delta}} \left(\prod_{t=1}^m \lambda_{\alpha(t)} \right) \\ &= \prod_{t=1}^n \lambda_t^{e_t(\chi)} \\ &= \det(A)^{e(\chi)}. \end{aligned}$$

$\square$

EXAMPLE 7.53 Let χ be the principal character of $G = S_2$. If $n = m = 2$, then $\hat{\Delta} = \bar{\Delta} = \Delta = G_{2,2} = \{(1, 1), (1, 2), (2, 2)\}$, and $e_1(\chi) = 2 + 1 + 0 = 3 = 0 + 1 + 2 = e_2(\chi)$. Thus, $e(\chi) = 3 = mo(\hat{\Delta})/n$. If

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

then $\det(A) = ad - bc$ and, by Theorem 7.52, $\det(P_2(A)) = (ad - bc)^3$. On the other hand, from Equation (7.32),

$$\begin{aligned} \det(P_2(A)) &= a^2(ad + bc)d^2 + 4ab^2c^2d - 4a^2bcd^2 - b^2(ad + bc)c^2 \\ &= a^3d^3 - 3a^2bcd^2 + 3ab^2c^2d - b^3c^3 \\ &= (ad - bc)^3. \end{aligned}$$

□

Suppose $T \in L(V, V)$. Let A be the matrix representation of T with respect to the ordered basis $B = \{e_1, e_2, \dots, e_n\}$ of V . If χ is the principal character of S_m , then $P_m(A)$ is the matrix representation of $K(T)$ with respect to the lexicographically ordered basis $B^* = \{[m!/o(G_\alpha)]^{1/2}e_\alpha^* : \alpha \in G_{m,n}\}$. The complicating factor $[m!/o(G_\alpha)]^{1/2}$ is necessary to preserve orthonormality, guaranteeing that $K(A)$ will inherit certain desirable properties from A . In some situations, however, the complicating factor is unnecessary, or even undesirable. (For example, B may not be orthonormal, or one may wish to consider generalizations to fields that do not contain square roots.) In these cases, it is sometimes useful to work with the alternative induced basis $B^* = \{e_\alpha^* : \alpha \in G_{m,n}\}$.

DEFINITION 7.54 Suppose B is a basis of the vector space V . Let $T \in L(V, V)$ be fixed but arbitrary. Suppose A is the matrix representation of T with respect to B . If χ is the principal character of S_m , denote by $A^{[m]}$ the matrix representation of $K(T)$ with respect to B^* .

Of course, $A^{[m]}$ is similar (in fact, diagonally similar) to $P_m(A)$. To obtain an explicit description of $A^{[m]}$, multiply both sides of

$$\begin{aligned} K(T) \left([G : o(G_\beta)]^{1/2} e_\beta^* \right) \\ = \sum_{\alpha \in G_{m,n}} \left([o(G_\alpha)o(G_\beta)]^{-1/2} \text{per}(A[\alpha|\beta]) \right) [G : G_\alpha]^{1/2} e_\alpha^* \end{aligned}$$

by $[G : o(G_\beta)]^{-1/2}$ to obtain

$$K(T)(e_\beta^*) = \sum_{\alpha \in G_{m,n}} o(G_\alpha)^{-1} \text{per}(A[\alpha|\beta]) e_\alpha^*. \quad (7.43)$$

This proves the following:

LEMMA 7.55 *Let $A \in \mathbb{C}_{n,n}$. Then $A^{[m]}$ is the $C(n+m-1, m)$ -square matrix whose rows and columns are indexed by $G_{m,n}$ ordered lexicographically. The (α, β) -entry of $A^{[m]}$ is $\text{per}(A[\alpha|\beta])/\nu(\alpha)$, where $\nu(\alpha) = o(S_{\mu(\alpha)})$ is the product of the factorials of the multiplicities of the distinct integers appearing in α .*

Being similar, $P_m(A)$ and $A^{[m]}$ have the same eigenvalues and, hence, the same trace.

EXAMPLE 7.56 If

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is a generic 2-by-2 matrix, then

$$A^{[2]} = \begin{pmatrix} a^2 & ab & b^2 \\ 2ac & ad+bc & 2bd \\ c^2 & cd & d^2 \end{pmatrix}. \quad (7.44)$$

Notice that $D^{-1}A^{[2]}D = P_2(A)$, where $D = \text{diag}(1, \sqrt{2}, 1)$, and $P_2(A)$ is given by Equation (7.32). $\square$

This variation on the theme of induced power matrices is an important tool in classical nineteenth century invariant theory.¹⁶

An Excursion Into Invariant Theory

Let $V = \mathbb{C}_1[x_1, x_2, \dots, x_n]$, the vector space of homogeneous polynomials of degree 1 in the independent indeterminates $x_1, x_2, \dots, x_n$. Then

$$V = \left\{ \sum_{i=1}^n c_i x_i : c_i \in \mathbb{C}, \quad 1 \leq i \leq n \right\}.$$

Suppose $\sigma \rightarrow A(\sigma) = (a_{ij}(\sigma))$, $\sigma \in G$, is a representation of degree n of the finite group G . Then $A(\sigma)$ determines a unique linear operator on V defined by

$$A(\sigma) \circ x_j = \sum_{i=1}^n a_{ij}(\sigma) x_i, \quad (7.45)$$

¹⁶The resurrection of invariant theory, due largely to Gian-Carlo Rota and his coauthors (see, e.g., [Doubilet, Rota & Stein (1974)], [Kung & Rota (1984)], [Barnabei, Brini & Rota (1985)], and [Rota & Stein (1986)] has received added momentum, e.g., from applications in coding theory [Sloane (1979)] and connections to Cohen-Macaulay algebras [Stanley (1979)]. For a nice overview, see [Gardner (1980)].

and linear extension. If $p = \sum c_j x_j \in V$, let $[p] \in \mathbb{C}_{n,1}$ be the column n -tuple whose i -th entry is c_i . Then

$$\begin{aligned} A(\sigma) \circ p &= \sum_{j=1}^n c_j A(\sigma) \circ x_j \\ &= \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij}(\sigma) c_j \right) x_i. \end{aligned}$$

In other words,

$$[A(\sigma) \circ p] = A(\sigma)[p], \quad (7.46)$$

the product of $A(\sigma)$ and $[p]$. In particular,

$$\begin{aligned} [A(\tau) \circ (A(\sigma) \circ p)] &= (A(\tau)A(\sigma))[p] \\ &= [A(\tau\sigma) \circ p], \end{aligned}$$

for all $\tau, \sigma \in G$ and all $p \in V$.

This action of the group $A(G) = \{A(\sigma) : \sigma \in G\}$ can be extended to arbitrary polynomials by defining

$$\begin{aligned} (A(\sigma) \circ f)(x_1, x_2, \dots, x_n) \\ = f(A(\sigma) \circ x_1, A(\sigma) \circ x_2, \dots, A(\sigma) \circ x_n), \end{aligned} \quad (7.47)$$

$f \in \mathbb{C}[x_1, x_2, \dots, x_n]$. If $g, h \in \mathbb{C}[x_1, x_2, \dots, x_n]$ and $b, c \in \mathbb{C}$, then

$$\begin{aligned} (A(\sigma) \circ (bg + ch))(x_1, x_2, \dots, x_n) \\ &= (bg + ch)(A(\sigma) \circ x_1, A(\sigma) \circ x_2, \dots, A(\sigma) \circ x_n) \\ &= bg(A(\sigma) \circ x_1, A(\sigma) \circ x_2, \dots, A(\sigma) \circ x_n) \\ &\quad + ch(A(\sigma) \circ x_1, A(\sigma) \circ x_2, \dots, A(\sigma) \circ x_n) \\ &= b(A(\sigma) \circ g)(x_1, x_2, \dots, x_n) + c(A(\sigma) \circ h)(x_1, x_2, \dots, x_n), \end{aligned}$$

that is, $A(\sigma) \circ (bg + ch) = b(A(\sigma) \circ g) + c(A(\sigma) \circ h)$. Thus, $A(\sigma)$ acts linearly on $\mathbb{C}[x_1, x_2, \dots, x_n]$, $\sigma \in G$.

DEFINITION 7.57 Suppose A is a representation of degree n of the finite group G . Then $f \in \mathbb{C}[x_1, x_2, \dots, x_n]$ is a **polynomial invariant** of the matrix group $A(G) = \{A(\sigma) : \sigma \in G\}$ if $A(\sigma) \circ f = f$, for all $\sigma \in G$.

From the perspective of groups acting on sets, f is a polynomial invariant of $A(G)$ if and only if it is a fixed point of $A(\sigma)$, $\sigma \in G$. From the perspective of linear operators, the polynomial invariants of $A(G)$ comprise an intersection of eigenspaces of $A(\sigma)$, $\sigma \in G$. In particular, they comprise a subspace of $\mathbb{C}[x_1, x_2, \dots, x_n]$. In fact, more is true. If f and g are invariants of $A(G)$, their product, fg , is another. The subspace of polynomial invariants of $A(G)$ is a “subalgebra” of $\mathbb{C}[x_1, x_2, \dots, x_n]$.

EXAMPLE 7.58 Suppose A is a representation of degree n of the finite group G . For a fixed but arbitrary $f \in \mathbb{C}[x_1, x_2, \dots, x_n]$, define

$$\bar{f} = \frac{1}{o(G)} \sum_{\sigma \in G} A(\sigma) \circ f.$$

Then

$$\begin{aligned} A(\tau) \circ \bar{f} &= \frac{1}{o(G)} \sum_{\sigma \in G} A(\tau) \circ (A(\sigma) \circ f) \\ &= \frac{1}{o(G)} \sum_{\sigma \in G} A(\tau\sigma) \circ f \\ &= \frac{1}{o(G)} \sum_{\sigma \in G} A(\sigma) \circ f \\ &= \bar{f}, \end{aligned}$$

for all $\tau \in G$. Hence, $\bar{f}$ is a polynomial invariant of $A(G)$. □

EXAMPLE 7.59 Suppose $A(e) = I_2$,

$$A((123)) = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} \quad \text{and} \quad A((132)) = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}.$$

Then $\sigma \rightarrow A(\sigma)$ is a (reducible) representation of the alternating group $G = A_3$.

If $f(x_1, x_2) = x_1^2 + x_1x_2 \in \mathbb{C}[x_1, x_2]$, then

$$\begin{aligned}\bar{f}(x_1, x_2) &= \frac{1}{3} \sum_{\sigma \in A_3} A(\sigma) \circ (x_1^2 + x_1x_2) \\ &= \frac{1}{3} ([x_1^2 + x_1x_2] + [(-x_2)^2 + (-x_2)(x_1 - x_2)] \\ &\quad + [(-x_1 + x_2)^2 + (-x_1 + x_2)(-x_1)]) \\ &= \frac{1}{3} ([x_1^2 + x_1x_2] + [2x_2^2 - x_1x_2] + [2x_1^2 - 3x_1x_2 + x_2^2]) \\ &= x_1^2 - x_1x_2 + x_2^2.\end{aligned}$$

Notice that

$$\begin{aligned}[A((123)) \circ \bar{f}](x_1, x_2) &= (-x_2)^2 - (-x_2)(x_1 - x_2) + (x_1 - x_2)^2 \\ &= x_1^2 - x_1x_2 + x_2^2 \\ &= \bar{f}(x_1, x_2),\end{aligned}$$

and

$$\begin{aligned}[A((132)) \circ \bar{f}](x_1, x_2) &= (-x_1 + x_2)^2 - (-x_1 + x_2)(-x_1) + (-x_1)^2 \\ &= x_1^2 - x_1x_2 + x_2^2 \\ &= \bar{f}(x_1, x_2),\end{aligned}$$

confirming that $\bar{f}$ is a (degree 2, homogeneous) polynomial invariant of

$$A(G) = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}, \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix} \right\}. \quad (7.48)$$

□

Returning to the general discussion, suppose $f \in \mathbb{C}[x_1, x_2, \dots, x_n]$. Then

$$f = \sum_{m \geq 0} p_m,$$

where $p_m \in \mathbb{C}_m[x_1, x_2, \dots, x_n]$, the subspace of $\mathbb{C}[x_1, x_2, \dots, x_n]$ consisting of (0 together with) the homogeneous polynomials of (total) degree m , and

all but finitely many of the p_m are zero. It follows from Equation (7.47) that $A(\sigma) \circ p_m \in \mathbf{C}_m[x_1, x_2, \dots, x_n]$, $m \geq 0$. Therefore,

$$\begin{aligned} A(\sigma) \circ f &= \sum_{m \geq 0} A(\sigma) \circ p_m \\ &= f, \end{aligned}$$

if and only if $A(\sigma) \circ p_m = p_m$ for all $m \geq 0$. It suffices, therefore, to restrict our investigation to homogeneous polynomials of degree m .

If $p \in \mathbf{C}_m[x_1, x_2, \dots, x_n]$ then (Lemma 6.65) there exist constants c_β , $\beta \in G_{m,n}$, such that

$$p(x_1, x_2, \dots, x_n) = \sum_{\beta \in G_{m,n}} c_\beta \prod_{t=1}^m x_{\beta(t)}.$$

Replacing x_j with

$$\sum_{i=1}^n a_{ij}(\sigma) x_i, \quad 1 \leq j \leq n,$$

yields

$$\begin{aligned} &(A(\sigma) \circ p)(x_1, x_2, \dots, x_n) \\ &= \sum_{\beta \in G_{m,n}} c_\beta \prod_{t=1}^m \left(\sum_{i=1}^n a_{i\beta(t)}(\sigma) x_i \right) \\ &= \sum_{\beta \in G_{m,n}} c_\beta \sum_{\alpha \in \Gamma_{m,n}} \prod_{t=1}^m a_{\alpha(t)\beta(t)}(\sigma) x_{\alpha(t)} \\ &= \sum_{\beta \in G_{m,n}} c_\beta \sum_{\alpha \in G_{m,n}} \frac{1}{v(\alpha)} \sum_{\tau \in S_m} \left(\prod_{t=1}^m a_{\alpha\tau(t)\beta(t)}(\sigma) \right) \left(\prod_{t=1}^m x_{\alpha(t)} \right) \\ &= \sum_{\alpha \in G_{m,n}} \left[\sum_{\beta \in G_{m,n}} \left(\frac{1}{v(\alpha)} \sum_{\tau \in S_m} \prod_{t=1}^m a_{\alpha\tau(t)\beta(t)}(\sigma) \right) c_\beta \right] \prod_{t=1}^m x_{\alpha(t)} \\ &= \sum_{\alpha \in G_{m,n}} \left[\sum_{\beta \in G_{m,n}} \left(\frac{1}{v(\alpha)} \text{per } (A(\sigma)[\alpha|\beta]) \right) c_\beta \right] \prod_{t=1}^m x_{\alpha(t)}, \quad (7.49) \end{aligned}$$

where $v(\alpha) = o(S_{\mu(\alpha)})$.

In lexicographic order, suppose $G_{m,n} = \{\beta_1, \beta_2, \dots, \beta_N\}$, where $N = C(n + m - 1, m)$. If $[p] \in \mathbb{C}_{N,1}$ is the column vector

$$[p] = (c_{\beta_1}, c_{\beta_2}, \dots, c_{\beta_N})^t,$$

then Equation (7.49) can be expressed as

$$[A(\sigma) \circ p] = A(\sigma)^{[m]}[p], \quad (7.50)$$

where $A(\sigma)^{[m]}$ is the variation of the m -th induced power of $A(\sigma)$ described in Definition 7.54 and Lemma 7.55. (Notice that Equation (7.50) reduces to Equation (7.46) when $m = 1$.) It follows that p is a homogeneous polynomial invariant of $A(G)$ of degree m , if and only if

$$A(\sigma)^{[m]}[p] = [p], \quad \sigma \in G, \quad (7.51)$$

if and only if, for all $\sigma \in G$, $[p]$ is an eigenvector of $A(\sigma)^{[m]}$, corresponding to the eigenvalue $\lambda = 1$. In other words, the space of homogeneous polynomial invariants of degree m of $A(G)$ is the intersection of the eigenspaces, afforded by $\lambda = 1$, of $A(\sigma)^{[m]}$, $\sigma \in G$. Because $I_n^{[m]} = I_{C(n+m-1, m)}$ and

$$\begin{aligned} A(\tau)^{[m]}A(\sigma)^{[m]} &= (A(\tau)A(\sigma))^{[m]} \\ &= A(\tau\sigma)^{[m]}, \end{aligned}$$

$\sigma \rightarrow A(\sigma)^{[m]}$ is a representation of G . It follows that the dimension of the common eigenspace is $(1, \xi_m)_G$, where $\xi_m(\sigma) = \text{tr}(A(\sigma)^{[m]})$, $\sigma \in G$. These observations are summarized in the following.

THEOREM 7.60 *Suppose A is a representation of degree n of the finite group G . Denote by d_m the dimension of the space of homogeneous polynomial invariants of degree m of $A(G)$. Then*

$$d_m = \frac{1}{o(G)} \sum_{\sigma \in G} \text{tr}(A(\sigma)^{[m]}). \quad (7.52)$$

For a fixed G and A , let f_A be the generating function for these dimensions, that is,

$$f_A(z) = \sum_{m \geq 0} d_m z^m.$$

MOLIER'S THEOREM 7.61 *Suppose A is a representation of degree n of the finite group G . Then*

$$f_A(z) = \frac{1}{o(G)} \sum_{\sigma \in G} \frac{1}{\det(I_n - zA(\sigma))}.$$

Proof By Equation (7.52), it suffices to show that the trace of $A(\sigma)^{[m]}$ is equal to the coefficient of z^m in

$$[\det(I_n - zA(\sigma))]^{-1} = \left[\prod_{i=1}^n (1 - z\lambda_i) \right]^{-1},$$

where $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of $A(\sigma)$. Because $P_m(A(\sigma))$ and $A(\sigma)^{[m]}$ are similar they have the same trace, namely (Equation (7.42)),

$$\text{tr } (A(\sigma)^{[m]}) = \sum_{\beta \in G_{m,n}} \prod_{t=1}^m \lambda_{\beta(t)}. \quad (7.53)$$

On the other hand,

$$\begin{aligned} \prod_{i=1}^n (1 - z\lambda_i)^{-1} &= \prod_{i=1}^n \left(\sum_{m \geq 0} (z\lambda_i)^m \right) \\ &= \sum_{m \geq 0} \left(\sum_{\beta \in G_{m,n}} \prod_{t=1}^m \lambda_{\beta(t)} \right) z^m. \end{aligned} \quad (7.54)$$

□

EXAMPLE 7.62 Returning to Example 7.59, where $G = A_3$ and

$$A(G) = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}, \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix} \right\},$$

$\det(I_2 - zA(e)) = (1 - z)^2$, and $\det(I_2 - zA((123))) = \det(I_2 - zA((132))) = 1 + z + z^2$. The “Molien Series” for $A(G)$ is

$$\begin{aligned} f_A(z) &= \frac{1}{3} \left[\frac{1}{(1-z)^2} + \frac{2}{1+z+z^2} \right] \\ &= [(1 + 2z + 3z^2 + 4z^3 + 5z^4 + \dots) \\ &\quad + 2(1 - z + z^3 - z^4 + z^6 - z^7 + \dots)]/3 \\ &= 1 + z^2 + 2z^3 + z^4 + 2z^5 + 3z^6 + 2z^7 + \dots \end{aligned} \quad (7.55)$$

Because $d_2 = 1$, every homogeneous polynomial invariant of degree 2 must be a multiple of the one we found in Example 7.59, namely,

$$x_1^2 - x_1x_2 + x_2^2.$$

Moreover, there are no homogeneous polynomial invariants of degree 1. (This is easily seen directly from Equation (7.52):

$$\begin{aligned} d_1 &= \frac{1}{3} \sum_{\sigma \in G} \operatorname{tr}(A(\sigma)) \\ &= \frac{1}{3}(2 - 1 - 1) \\ &= 0. \end{aligned}$$

As the coefficient of z^3 in Equation (7.55) is 2, the space of homogeneous polynomial invariants of degree 3 has dimension 2. Let's try to find two linearly independent invariants of degree 3. Begin with some arbitrary homogeneous polynomial of degree 3, say $f_1(x_1, x_2) = x_1^3$. As in Example 7.59, average $A(\sigma) \circ f_1$ over $A(G)$ obtaining

$$\bar{f}_1(x_1, x_2) = x_1^2 x_2 - x_1 x_2^2.$$

If $f_2(x_1, x_2) = x_2^3$, then $\bar{f}_2 = -\bar{f}_1$, which is useless because we are seeking a linearly independent pair. (Evidently, if $g(x_1, x_2) = x_1^3 + x_2^3$, then $\bar{g} = 0$.) However, if $f_3(x_1, x_2) = -3x_1^2 x_2$, then

$$\bar{f}_3(x_1, x_2) = x_1^3 - 3x_1^2 x_2 + x_2^3.$$

Because $\bar{f}_1$ and $\bar{f}_3$ are linearly independent, we have found all homogeneous polynomial invariants of $A(G)$ of degree 3, namely, the nonzero polynomials in

$$\left\{ c_1 (x_1^2 x_2 - x_1 x_2^2) + c_2 (x_1^3 - 3x_1^2 x_2 + x_2^3) : c_1, c_2 \in \mathbb{C} \right\}.$$

What about degree 4? We already know that $x_1^2 - x_1 x_2 + x_2^2$ is a homogeneous polynomial invariant of degree 2. Because the space of all invariants is an algebra, $(x_1^2 - x_1 x_2 + x_2^2)^2$ is a homogeneous polynomial invariant of degree 4. Because the coefficient of z^4 in Equation (7.55) is 1, every homogeneous polynomial invariant of degree 4 must be a nonzero multiple of $(x_1^2 - x_1 x_2 + x_2^2)^2$. The degree 5 case is just as easy. From Equation (7.55), two linearly independent homogeneous invariants of degree 5 are required. From the solutions to the second and third degree problems, we obtain

$$(x_1^2 - x_1 x_2 + x_2^2)(x_1^2 x_2 - x_1 x_2^2)$$

and

$$(x_1^2 - x_1 x_2 + x_2^2)(x_1^3 - 3x_1^2 x_2 + x_2^3).$$

□

Applications to Graphs

Let $H = (V, E)$ be a graph with vertex set $V = \{1, 2, \dots, n\}$ and edge set $E = E(H) \subset V^{(2)}$, the two-element subsets of V . Recall (Definition 6.85) that the natural action of $\sigma \in S_n$ on $V^{(2)}$ is defined by $\tilde{\sigma}(\{i, j\}) = \{\sigma(i), \sigma(j)\}$, $\{i, j\} \in V^{(2)}$. In particular, $\sigma \rightarrow \tilde{\sigma}$ is an isomorphism¹⁷ from S_n onto the pair group, $S_n^{(2)} = \{\tilde{\sigma} : \sigma \in S_n\}$. Moreover (Equation (6.55)) $H = (V, E)$ is isomorphic to $K = (V, F)$ if and only if there is a $\sigma \in S_n$ such that $\tilde{\sigma}(E) = F$. We will abuse the language and write this condition as $\tilde{\sigma}(H) = K$.

Let $x_{\{1,2\}}, x_{\{1,3\}}, \dots, x_{\{n-1,n\}}$ be $C(n, 2)$ independent indeterminates over $\mathbb{C}$, indexed by the elements of $V^{(2)}$. Define the monomial

$$f_H(x_{\{1,2\}}, x_{\{1,3\}}, \dots, x_{\{n-1,n\}}) = \prod_{e \in E(H)} x_e. \quad (7.56)$$

Observe that f_H uniquely determines H ; if $H = (V, E)$ and $K = (V, F)$, then $f_H = f_K$ if and only if $H = K$. Another point to be made about f_H is that it is "square-free". It is a product of $o(E(H))$ different variables.

Define $\tilde{\sigma} \circ x_{\{i,j\}} = x_{\{\sigma(i), \sigma(j)\}}$. Then, as in the general discussion of polynomial invariants,

$$\begin{aligned} & (\tilde{\sigma} \circ f_H)(x_{\{1,2\}}, x_{\{1,3\}}, \dots, x_{\{n-1,n\}}) \\ &= f_H(\tilde{\sigma} \circ x_{\{1,2\}}, \tilde{\sigma} \circ x_{\{1,3\}}, \dots, \tilde{\sigma} \circ x_{\{n-1,n\}}) \\ &= f_H(x_{\{\sigma(1), \sigma(2)\}}, x_{\{\sigma(1), \sigma(3)\}}, \dots, x_{\{\sigma(n-1), \sigma(n)\}}) \\ &= f_K(x_{\{1,2\}}, x_{\{1,3\}}, \dots, x_{\{n-1,n\}}), \end{aligned}$$

where $K = \tilde{\sigma}(H)$. As in Example 7.58, define

$$\bar{f}_H = \frac{1}{n!} \sum_{\sigma \in S_n} \tilde{\sigma} \circ f_H. \quad (7.57)$$

Properties of $\bar{f}_H$ are summarized in the following.

LEMMA 7.63 *Let $H = (V, E)$ and $K = (V, F)$ be graphs on vertex set $V = \{1, 2, \dots, n\}$. Then H and K are isomorphic if and only if $\bar{f}_H = \bar{f}_K$. Moreover, if $H = H_1, H_2, \dots, H_k$ are the different graphs with vertex set V that are isomorphic to H , then*

$$\bar{f}_H = \frac{1}{k} \sum_{i=1}^k f_{H_i}.$$

¹⁷ One may view $\sigma \rightarrow \tilde{\sigma}, \sigma \in S_n$, as a faithful permutation representation of degree $C(n, 2)$.

Proof H and K are isomorphic if and only if there is a permutation $\tau \in S_n$ such that $\tilde{\tau}(H) = K$, in which case,

$$\begin{aligned}\bar{f}_K &= \frac{1}{n!} \sum_{\sigma \in S_n} \tilde{\sigma} \circ f_K \\ &= \frac{1}{n!} \sum_{\sigma \in S_n} \tilde{\sigma} \circ [\tilde{\tau} \circ f_H] \\ &= \frac{1}{n!} \sum_{\sigma \in S_n} [\tilde{\sigma} \circ \tilde{\tau}] \circ f_H \\ &= \frac{1}{n!} \sum_{\sigma \in S_n} \tilde{\sigma} \circ f_H \\ &= \bar{f}_H.\end{aligned}$$

Conversely, if $\bar{f}_H = \bar{f}_K$, then f_K must be among the summands comprising $\bar{f}_H$, that is, there exists a $\sigma \in S_n$ such that $\tilde{\sigma} \circ f_H = f_K$ and, hence, $\tilde{\sigma}(H) = K$.

To prove the “moreover” part, denote the automorphism group of H by $\mathcal{A}(H) = \{\sigma \in S_n : \tilde{\sigma}(H) = H\}$. In the summation on the right-hand side of Equation (7.57) f_H occurs $o(\mathcal{A}(H))$ times. If K is isomorphic to H , then $\{\sigma \in S_n : \tilde{\sigma}(H) = K\}$ is a coset of $\mathcal{A}(H)$. Therefore, f_K also occurs $o(\mathcal{A}(H))$ times, and $k = n!/o(\mathcal{A}(H))$. $\square$

LEMMA 7.64 *Let $H_1 = (V, E_1)$, $H_2 = (V, E_2)$, ..., $H_s = (V, E_s)$ be graphs with vertex set $V = \{1, 2, \dots, n\}$. Then $\{\bar{f}_{H_i} : 1 \leq i \leq s\}$ is linearly independent if and only if the graphs $H_1, H_2, \dots, H_s$ are pairwise nonisomorphic.*

Proof If $H_i \cong H_j$ then (Lemma 7.63) $\bar{f}_{H_i} = \bar{f}_{H_j}$. Conversely, assume $H_1, H_2, \dots, H_s$ are pairwise nonisomorphic. Suppose

$$\sum_{i=1}^s c_i \bar{f}_{H_i} = 0.$$

Because f_{H_i} occurs only in $\bar{f}_{H_i}$, its coefficient in this sum is a positive integer multiple of c_i . Taking successive partial derivatives with respect to x_e , as e ranges over E_i , yields $c_i = 0$. $\square$

For a fixed n and m , suppose $H_1, H_2, \dots, H_k$ is a system of distinct representatives for the nonisomorphic graphs having n vertices and m edges. Without loss of generality, we may assume these k graphs share the vertex set $V = \{1, 2, \dots, n\}$. Let U be the “free vector space” generated by $H_1, H_2, \dots, H_k$,

that is, $\{H_1, H_2, \dots, H_k\}$ is a basis of U . Denote by W the space of homogeneous polynomial invariants of degree m of $S_n^{(2)}$ in the indeterminates $x_{\{1,2\}}, x_{\{1,3\}}, \dots, x_{\{n-1,n\}}$. Define $T \in L(U, W)$ by $T(H_i) = \bar{f}_{H_i}$, $1 \leq i \leq k$, and linear extension. It follows from Lemma 7.64 that T is one-to-one. We claim that T is onto the subspace of square-free polynomials in W . To see this, let $q = q(x_{\{1,2\}}, x_{\{1,3\}}, \dots, x_{\{n-1,n\}})$ be a square-free homogeneous polynomial invariant of degree m of $S_n^{(2)}$. Suppose the monomial M occurs in q with coefficient c . Let $H = (V, E)$, $o(E) = m$, be the unique graph determined by M . Because q is an invariant of $S_n^{(2)}$, $c\bar{f}_H$ is among its summands. Therefore, $q - c\bar{f}_H$ contains strictly fewer monomials than q , and the result follows by induction.

The main outcome of these observations can be summarized as follows:

THEOREM 7.65¹⁸ *The number, $g(n, m)$, of nonisomorphic graphs having n vertices and m edges is equal to the dimension of the subspace of square-free homogeneous polynomial invariants of degree m of $S_n^{(2)}$ in the $C(n, 2)$ indeterminates $x_{\{1,2\}}, x_{\{1,3\}}, \dots, x_{\{n-1,n\}}$.*

We now investigate another kind of polynomial invariant of graphs. A matrix $M \in \mathbb{C}_{n,n}$ is a **monomial matrix** if it can be factored as $M = PC$, where C is an invertible diagonal matrix and P is a permutation matrix. It is not difficult to prove that the set of monomial matrices comprises a subgroup of $GL(n, \mathbb{C})$. It is somewhat harder to show that the monomial group is a maximal subgroup of $GL(n, \mathbb{C})$.¹⁹

LEMMA 7.66 *Suppose $\chi \in I(S_n)$. Let $A = (a_{ij})$ be a generic n -by- n matrix. If $M = PC \in \mathbb{C}_{n,n}$ is a monomial matrix, then $d_\chi(M^{-1}AM) = d_\chi(A)$.*

Proof Suppose $P = (\delta_{i,\tau(j)})$ and $C = \text{diag}(c_1, c_2, \dots, c_n)$. If $M = PC$, then

$$\begin{aligned} d_\chi(M^{-1}AM) &= \sum_{\sigma \in S_n} \chi(\sigma) \prod_{i=1}^n c_{\sigma(i)} a_{\tau(i), \tau\sigma(i)} / c_i \\ &= \sum_{\sigma \in S_n} \chi(\sigma) \left(\prod_{i=1}^n a_{i, \tau\sigma\tau^{-1}(i)} \right) \left(\prod_{i=1}^n c_{\sigma(i)} / c_i \right) \\ &= \sum_{\sigma \in S_n} \chi(\tau^{-1}\sigma\tau) \prod_{i=1}^n a_{i\sigma(i)} \\ &= d_\chi(A). \end{aligned}$$

□

¹⁸This result is taken from [Merris & Watkins (1984)].

¹⁹See [Friedland (1985)].

THEOREM 7.67 *Let H and K be graphs on n vertices with Laplacian matrices $L(H)$ and $L(K)$ respectively. If H and K are isomorphic then $d_\chi(xI_n - L(H)) = d_\chi(xI_n - L(K))$, for every $\chi \in I(S_n)$.*

Proof Because (Theorem 2.48) H and K are isomorphic if and only if $L(H)$ and $L(K)$ are permutation similar, the result follows from Lemma 7.66. $\square$

DEFINITION 7.68 If A is an n -by- n matrix and $\chi \in I(S_n)$, the χ -th immanantal polynomial²⁰ of A is $d_\chi(xI_n - A)$.

The immanantal polynomial corresponding to $\chi = \varepsilon$ is the characteristic polynomial, which is preserved under any similarity. Despite Lemma 7.66, the remaining immanantal polynomials are not similarity invariants. In a perverse way, this makes $d_\chi(xI_n - L(H))$ more attractive as a graph invariant. Generically preserved only under monomial similarities (when $\chi \neq \varepsilon$), these polynomials seem well suited to the condition that graphs H and K are isomorphic if and only if $L(H)$ and $L(K)$ are permutation similar. In fact, it is natural to wonder whether, taken all together, the Laplacian immanantal polynomials characterize graphs up to isomorphism. The answer is, they do not.

THEOREM 7.69 *Let t_n be the number of nonisomorphic trees on n vertices and s_n the number of such trees T for which there exists a nonisomorphic tree T' such that*

$$d_\chi(xI_n - L(T)) = d_\chi(xI_n - L(T')), \quad (7.58)$$

for every $\chi \in I(S_n)$. Then $\lim_{n \rightarrow \infty} s_n/t_n = 1$.

Theorem 7.69 appeared in [Botti & Merris (1993)]. The proof depends on the fact [Schwenk (1973)] that the probability of finding a fixed finite “limb” on a randomly chosen tree goes to 1 as the number of vertices goes to infinity. Because of this result, it suffices to exhibit a single pair of trees with certain nice properties. Such a pair can be found in [McKay (1977)].

DEFINITION 7.70 Two graphs, H and K , are *coimmanantal* if $d_\chi(xI_n - L(H)) = d_\chi(xI_n - L(K))$, for every $\chi \in I(S_n)$. A set of graphs is coimmanantal if its elements are pairwise coimmanantal.

It follows from Theorem 7.69 that the probability a randomly chosen tree on n vertices is part of a coimmanantal pair approaches 1 for large n . This raises the

²⁰ Immanantal polynomials have been studied in a variety of contexts. See, e.g., [Baxter (1978)], [Beasley & Brenner (1968)], [Brenner & Brualdi (1967)], [Engel (1973)], [Friedland (1972) & (1975)], [Gibson (1971), (1972) & (1978)], [Horad & Sossa (1995)], [Johnson, Merris & Pierce (1985–86)], [Kräuter (1987)], [Merris (1975b) & (1994b)], [Oliveira (1970), (1971) & (1972)], [Sá (1981)], and [Strok (1990)]. The first appearance of immanantal polynomials in the study of graphs seems to have been in [Turner (1968)].

question whether one can find arbitrarily large coimmanantal families of graphs. As the next result shows, there exist coimmanantal families whose numbers grow exponentially with n .

COROLLARY 7.71²¹ Suppose k is a fixed but arbitrary positive integer. Let $n = 17\lceil \log_2(k) \rceil + 1$, where $\lceil \cdot \rceil$ is the ceiling function. Then there exists a coimmanantal family of k trees on n vertices.

It is important not to misconstrue Theorem 7.69 and Corollary 7.71. While it is true that, by themselves, Laplacian immanantal polynomials do not distinguish all pairs of nonisomorphic graphs, this takes nothing away from the fact that they are genuine graph invariants.

There are some other perspectives from which to view Laplacian immanants of graphs. For example, given that the permanental dominance conjecture is so difficult, why not try to prove it for some interesting subsets of $\mathcal{H}_n$? The first result along these lines was obtained by [Chan & Lam (1996)] who proved that

$$\bar{d}_X(L(T)) \leq \text{per}(L(T)) \quad (7.59)$$

whenever T is a tree. Alternatively, an inequality among immanants that holds for every $A \in \mathcal{H}_n$ might well be subject to some improvement when restricted to a suitably chosen subset of $\mathcal{H}_n$. It is proved in [Brualdi & Goldwasser (1984)],²² for example, that the Hadamard Theorem for Permanents can be improved to

$$\text{per}(L(K)) \geq 2h(L(K)) \quad (7.60)$$

for bipartite K , and in [Chan & Lam (1996)] that, upon restriction to Laplacian matrices of trees, Heyfron's inequalities for single-hook immanants can be strengthened to

$$\bar{d}_{k-1}(L(T)) \leq \frac{k-2}{k-1} \bar{d}_k(L(T)), \quad 2 \leq k \leq n. \quad (7.61)$$

²¹Corollary 7.71 is from [Merris (1995)].

²²Also see [Bapat (1986)] and [Vrba (1986a&b)]. Related work can be found in [Balasubramaniam (1993)], [Borowiecki & Jozwiak (1982)], [Faria (1985) & (1996)], [Goldwasser (1986)], [Goulden & Jackson (1981) & (1992a)], [Grone & Merris (1988)], and [Merris (1986) & (1994b)].

Exercises

1. Prove that Hadamard's Theorem is a consequence of Fischer's Inequality.
2. Suppose $A = (a_{ij}) \in \mathbf{C}_{n,n}$. Prove that

$$|\det(A)|^2 \leq \prod_{j=1}^n \sum_{i=1}^n |a_{ij}|^2.$$

(Hint: Show that this is a variation of Hadamard's Inequality.)

3. Let G be a permutation group of degree m . If $\chi \in I(G)$, prove that
 - a. $d_\chi(A^*) = \overline{d_\chi(A)}$, $A \in \mathbf{C}_{m,m}$.
 - b. $d_\chi(A') = d_{\bar{\chi}}(A)$, $A \in \mathbf{C}_{m,m}$, where $\bar{\chi}(\sigma) = \overline{\chi(\sigma)}$, $\sigma \in G$.
4. Let $B = (b_{ij}) \in \mathcal{H}_m$ and $c \in \mathbf{C}_m$ be fixed but arbitrary. Define

$$c_B(\sigma) = c(\sigma) \prod_{i=1}^m b_{i\sigma(i)}, \quad \sigma \in S_m.$$

- a. Prove that $c_B \in \mathbf{C}_m$. (Hint: $d_{c_B}(A) = d_c(B \cdot A)$.)
 - b. Prove that $d_c(B \cdot A) \geq c(e)h(B)\det(A)$, $A \in \mathcal{H}_m$.
 - c. Prove Oppenheim's Inequality:²³ $\det(B \cdot A) \geq h(B)\det(A)$, $A, B \in \mathcal{H}_m$.
5. Prove Lemma 7.7.
 6. Let χ be an irreducible character of the subgroup G of S_m . Prove directly that $\chi \in \mathbf{C}_m^+$. (Hint: Mimic Example 7.13.)
 7. Let V be a vector space. A nonempty subset K of V is a cone if $cu + dv \in K$ for all $u, v \in K$ and for all nonnegative real numbers c and d .
 - a. Show that $\mathbf{C}_m$ is a cone.
 - b. Show that $\mathcal{H}_m$ is a cone.
 8. Denote the principal character of S_m by 1_m and the alternating character by ε . Define $c: S_m \rightarrow \mathbb{C}$ by $c = 1_m - \varepsilon$.
 - a. Prove that $c \in \mathbf{C}_m$. (Hint: $\text{per}(A) \geq \det(A)$, $A \in \mathcal{H}_m$.)
 - b. Prove that $\varepsilon^* M_c \varepsilon = -(n!)^2$.
 - c. Explain why $c \in \mathbf{C}_m \setminus \mathbf{C}_m^+$.
 - d. Show that $\{c \in \mathbf{C}_m : c(e) = 0\}$ coincides with the set of functions $c: S_m \rightarrow \mathbb{C}$ such that $d_c(A) \geq 0$ for all (not necessarily positive semidefinite) hermitian $A \in \mathbf{C}_{m,m}$.

²³See [Oppenheim (1930)]. The permanent analog of Oppenheim's Inequality, namely $\text{per}(B \cdot A) \leq h(B) \text{per}(A)$, $A, B \in \mathcal{H}_m$, is an unresolved conjecture. (See [Bapat & Sunder (1986)] and [Chollet (1982)].)

9. Prove that $\text{per}(A)$ is an eigenvalue of the Schur power matrix $\Pi(A)$.
10. Suppose A is a positive definite hermitian matrix, partitioned as in Equation (7.1), that is,

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

where A_{11} is p -by- p . Recall (Chapter 2, Exercise 54) that the Schur Complement of A_{11} in A is $A/A_{11} = A_{22} - A_{21}A_{11}^{-1}A_{12}$. It was shown in [Bunce (1991)] that $\Pi(A) \geq \Pi(A_{11} \oplus A/A_{11})$.

- a. Use this fact to prove that $d_c(A) \geq d_c(A_{11} \oplus A/A_{11})$, $1 < p < m$, for all m -by- m positive definite hermitian A , and all $c \in \mathcal{C}_m^+$.
- b. Can one draw the same conclusion for all $c \in \mathcal{C}_m$? (Hint: Let

$$A_{11} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad A_{12} = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \quad \text{and} \quad A_{22} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

Take c to be the function from Example 7.12.)

11. The matrix $A = (a_{ij}) \in \mathcal{H}_m$ is a **correlation matrix** if $a_{ii} = 1$, $1 \leq i \leq m$. Let $\chi, \xi \in I(G)$. Prove that if $\bar{d}_\chi(A) \geq \bar{d}_\xi(A)$ for every m -by- m correlation matrix, then $\bar{d}_\chi(A) \geq \bar{d}_\xi(A)$, $A \in \mathcal{H}_m$. (Thus, it would suffice to prove the permanental dominance conjecture for correlation matrices.)
12. It was shown in [Merris (1983)] that $(h(A^m))^{1/m} \geq \bar{d}_\chi(A)$, $A \in \mathcal{H}_m$.
- a. Use this result to prove the following inequality (that first appeared in [Marcus & Minc (1965a)]): If $\lambda_1, \lambda_2, \dots, \lambda_m$ are the eigenvalues of $A \in \mathcal{H}_m$ then

$$\frac{1}{m} \sum_{i=1}^m \lambda_i^m \geq \text{per}(A).$$

(Hint: Use the arithmetic-geometric mean inequality to show that $\text{tr}(A^m)/m \geq (h(A^m))^{1/m}$.)²⁴

- b. Suppose $\lambda_1 \geq \lambda_2 \geq 0$. Prove that there exists a matrix $A \in \mathcal{H}_2$ having eigenvalues λ_1 and λ_2 , and permanent $(\lambda_1^2 + \lambda_2^2)/2$. (Hint: Look for a real symmetric matrix having equal diagonal entries.)²⁵

²⁴It was conjectured in [Grone & Merris (1987)] that the permanent of a correlation matrix is not less than the average of the *squares* of its eigenvalues (multiplicities included).

²⁵When $m > 2$, the problem of maximizing $\text{per}(A)$ over those $A \in \mathcal{H}_m$ with a prescribed spectrum is not so easy. See, for example, [Grone, Johnson, Sa, & Wolkowicz (1986)].

13. (S. Pierce) Let $A = 3I_3 - J_3$, where J_3 is the 3-by-3 each of whose entries is 1.
- Prove that $A \in \mathcal{H}_3$.
 - If $G = A_3$ and χ is the principal character, show that $d_1(A) = 6$.
 - If $G = A_3$ and χ is one of the nonprincipal linear characters of G , show that $d_\chi(A) = 9$.
 - If χ is an irreducible character of the subgroup G of S_m , is it always true that $d_1(A) \geq \bar{d}_\chi(A)$, $A \in \mathcal{H}_m$?
14. Show that $\bar{d}_{[5,1]}(A) \geq \bar{d}_{[3,2,1]}(A)$, $A \in \mathcal{H}_6$.
15. Show that $\bar{d}_{[4,3]}(A) \geq \bar{d}_{[4,2,1]}(A)$, $A \in \mathcal{H}_7$.
16. Suppose $\pi = [\pi_1, \pi_2, \dots, \pi_k] \vdash m$. If $\pi_1 = r$, show that $\bar{d}_\pi(A) \geq \bar{d}_r(A)$, $A \in \mathcal{H}_m$, where $d_r = d_{[r, 1^{m-r}]}$. (Hint: Pate's Theorem.)
17. Suppose $\pi, \rho \vdash m$. Write $\pi \sqsubset \rho$ if, for all $A \in \mathcal{H}_m$, $d_\pi(A) > 0 \Rightarrow d_\rho(A) > 0$.
- Prove that $\pi \sqsubset \rho$ if and only if π is majorized by ρ . (Hint: Theorems 6.47 and 7.26.)²⁶
 - If $\bar{d}_\rho(A) \geq \bar{d}_\pi(A)$, $A \in \mathcal{H}_m$, prove that ρ majorizes π .
18. (S. Pierce) Let J_n be the n -by- n matrix each of whose entries is 1.
- If $A = J_2 \oplus J_2$, show that $\bar{d}_{[3,1]}(A) < \bar{d}_{[2^2]}(A)$.
 - If $A = J_2 \oplus J_2 \oplus J_1$, show that $\bar{d}_{[3,1^2]}(A) < \bar{d}_{[2^2,1]}(A)$.
 - Comment on the following conjecture: "If ρ majorizes π , then $\bar{d}_\rho(A) \geq \bar{d}_\pi(A)$, $A \geq 0$."
19. Suppose $\bar{d}_\chi(A) \geq \text{per}(A)$ for all m -square $A \geq 0$. Prove that $G = S_m$ and $\chi = 1$.
20. Suppose $A \in \mathcal{H}_m$ is fixed but arbitrary. Let G be a subgroup of S_m and χ be the principle character of G . Prove that $d_\chi(A) = 0$ if and only if A has a zero row. (Hint: Corollary 7.27.)
21. Suppose $\pi = [\pi_1, \pi_2, \dots, \pi_r] \vdash m$. Let $\chi = \chi_\pi \in I(S_m)$.
- If $A \in \mathbf{C}_{m,m}$ has more than π_1 equal rows, prove that the immanant $d_\chi(A) = 0$. (Hint: Corollary 7.27.)
 - What property of determinants is generalized in part (a)?
22. Suppose $\pi = [\pi_1, \pi_2, \dots, \pi_r] \vdash m$. Let $\chi = \chi_\pi \in I(S_m)$.
- If the rank of $A \in \mathbf{C}_{m,m}$ is less than r , prove that the immanant $d_\chi(A) = 0$. (Hint: Corollary 7.27.)
 - What property of determinants is generalized in part (a)?

²⁶"Singular sets" for generalized matrix functions are discussed in [Beasley & Cummings (1983)].

23. Suppose $A \in \mathcal{H}_m$. Then (Lemma 2.43), there exists a matrix $B \in \mathbf{C}_{m,m}$ such that $A = B^*B$.
- Prove that there exists an upper triangular matrix $C \in \mathbf{C}_{m,m}$ such that $A = C^*C$.
 - Use Theorem 7.29 to prove that $d_\chi(A) \geq \chi(e)|\det(C)|^2$.
 - Use part (b) and the fact that $|\det(C)|^2 = \det(C^*)\det(C)$ to give a proof of Schur's Inequality that does not depend on Watkins's Theorem.
24. Suppose $A \in \mathcal{H}_m$.
- Prove that $h(A) = \det(A)$ if and only if A has a zero row (and column) or A is diagonal.
 - Prove that $\text{per}(A) = \det(A)$ if and only if A has a zero row (and column) or A is diagonal.
25. [Williamson (1969)]²⁷ Let H be a subgroup of G , where G is a permutation group of degree m . Suppose χ is an irreducible character of G that remains irreducible when restricted to H .
- Prove that $T(G, \chi)T(H, \chi) = T(G, \chi)$. (Hint: Use the Schur Relations.)
 - Prove that $V_\chi(G) \subset V_\chi(H)$, for any vector space V .
 - Prove that $d_\chi^G(A)/o(G) \leq d_\chi^H(A)/o(H)$, $A \in \mathcal{H}_m$, where, d_χ^G is the generalized matrix function based on G and χ and d_χ^H is the generalized matrix function based on H and χ .
26. Let G be a subgroup of S_m . Suppose $\tau \in Z(G)$, the center of G . If $A = (a_{ij}) \in \mathbf{C}_{m,m}$, prove that

$$\prod_{i=1}^m a_{i\tau(i)} = \frac{1}{o(G)} \sum_{\chi \in I(G)} \chi(\tau^{-1}) d_\chi(A).$$

27. Suppose $A \in \mathcal{H}_m$ is partitioned into blocks,

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

where A_{11} is invertible. Prove that

$$\det(A_{22}) \geq \det((A/A_{11})) + \det(A_{12}^* A_{11}^{-1} A_{12}),$$

where (A/A_{11}) is the Schur complement of A_{11} in A . (See Exercise 10.)

²⁷ Stated originally for linear characters, Williamson's work has since been extended in several ways. (See, for example, [Merris (1976)].)

28. Suppose $A = (a_{ij}) \in \mathcal{H}_n$ is positive definite. Denote by A_t the leading t -by- t principal submatrix of A , and let u_t be the vector formed from the first $t - 1$ entries of row t of A , that is, $u_t = (a_{t,1}, a_{t,2}, \dots, a_{t,t-1})$
- Prove that $a_{tt} - u_t A_{t-1}^{-1} u_t^* > 0$, $t > 1$. (Hint: This is a Schur Complement problem. See Exercise 10.)
 - Prove that

$$a_{11} \prod_{t=2}^m (a_{tt} - u_t A_{t-1}^{-1} u_t^*) = \det(A).$$

- Compare and contrast part (b) with Hadamard's Inequality.
- If $\lambda = \lambda_1$ is the largest eigenvalue of A , prove that

$$h(A) - \det(A) \geq \frac{\det(A)}{\lambda^2} \sum_{i < j} |a_{ij}|.$$

(Hint: Show that $\det(A(t|t)) \geq \det(A)/\lambda$, and $u_t A_{t-1}^{-1} u_t^* \geq \|u_t\|/\lambda$.)

29. Let χ be an irreducible character of the subgroup G of S_m . Suppose $A \in \mathcal{H}_m$ has maximum eigenvalue λ_1 and minimum eigenvalue λ_m .
- Prove that $\lambda_1^m \geq \bar{d}_\chi(A) \geq \lambda_m^m$. (Hint: Show that $\lambda_1 I_m \geq A \geq \lambda_m I_m$.)
 - Show that the right-hand inequality in part (a) is inferior to Schur's Inequality.
 - When χ is the principal character of S_m , show that the left-hand inequality in part (a) is inferior to the inequality in Exercise 12(a).
30. Recall that an m -by- m matrix $A = (a_{ij})$ is doubly stochastic if $a_{ij} \geq 0$, for all i and j , and if each row and column of A sums to 1. If $A \in \mathcal{H}_m$ is doubly stochastic, prove that
- $A \geq \frac{1}{m} J_m$, the m -by- m matrix each of whose entries is $1/m$.
 - $\text{per}(A) \geq m!/m^m$.
31. Prove Lemma 7.35.
32. Let χ be an irreducible character of the subgroup G of S_m . Suppose $A \in \mathbb{C}_{n,n}$.
- Prove that $d_\chi(A[\alpha|\beta]) = 0$ if either α or β fails to belong to Ω .
 - What well known statement about determinants is generalized in part (a)?
33. Suppose $A, B \in \mathcal{H}_m$.
- Prove that $\det(A^{1/m}) = [\det(A)]^{1/m}$.
 - Prove that Theorem 7.33 is stronger than Corollary 7.31 when $G = S_m$ and $\chi = \varepsilon$. (Hint: $[\det(A)^{1/m} + \det(B)^{1/m}]^m = \det(A) + \det(B)$ plus a sum of terms of the form $\det(B^{s/2m} A^{t/m} B^{s/2m})$, where $s + t = m$.)

34. Prove that Theorem 7.33 is stronger than Corollary 7.31 when A and B commute.
35. Let χ be an irreducible character of the subgroup G of S_m . Fix n and take $\Gamma_{m,n}$ to be ordered lexicographically. For each $A \in \mathbf{C}_{n,n}$, let $R(A)$ be the n^m -by- n^m matrix whose (α, β) -entry is

$$(\chi(e)/o(G))d_\chi(A[\alpha|\beta]).$$

- a. Prove that $R(AB) = R(A)R(B)$.
 - b. Explain why $A \rightarrow R(A)$ is *not* a representation of $GL(n, \mathbf{C})$.
36. Let χ be a linear character of the subgroup G of S_m . If $V_\chi(G) \neq \{0\}$, prove that $A \rightarrow K(A)$ is a representation of $GL(n, \mathbf{C})$.
37. Let $m = n$, $G = S_n$, and $\chi = \varepsilon$. Show that $K(A)$ is the one-by-one matrix $(\det(A))$.
38. Show that Equation (7.28) is an immediate consequence of Corollary 7.42 and the definition of matrix multiplication.
39. Suppose χ is an irreducible character of the subgroup G of S_m . Let $A, B \in \mathcal{H}_n$, and suppose $A \geq B$.
- a. If χ is linear, prove that $K(A) \geq K(B)$.
 - b. Use part (a) to prove that $d_\chi(A[\alpha|\alpha]) \geq d_\chi(B[\alpha|\alpha])$, $\alpha \in \overline{\Delta}$, when χ is linear.
 - c. Prove that $d_\chi(A[\alpha|\alpha]) \geq d_\chi(B[\alpha|\alpha])$, $\alpha \in \Gamma_{m,n}$, whether χ is linear or not.
40. Suppose $A, B \in \mathcal{H}_n$. If $A \geq B$, prove that $\text{padj}(A) \geq \text{padj}(B)$.
41. The classical adjoint (or adjugate) of $A \in \mathbf{C}_{n,n}$ is the n -by- n matrix $\text{adj}(A)$ whose (i, j) -entry is $(-1)^{i+j} \det(A(j|i))$.
- a. Prove that $\text{adj}(A)$ and $C_{n-1}(A)$ have the same spectrum, that is, the same eigenvalues with the same multiplicities.
 - b. Suppose $n > 1$. Is $\text{adj}(A)$ similar to $C_{n-1}(A)$ for all $A \in \mathbf{C}_{n,n}$? Justify your answer.
42. Confirm by a direct computation that Equation (7.35) is valid for a generic 3-by-3 matrix $A = (a_{ij})$.
43. An *inversion* of $\sigma \in S_m$ is an ordered pair (i, j) such that $1 \leq i < j \leq m$ and $\sigma(i) > \sigma(j)$. Denote by $\text{inv}(\sigma)$ the number of inversions of σ . Let $q \in [-1, 1]$ be fixed but arbitrary and define $c: S_m \rightarrow \mathbf{C}$ by $c(\sigma) = q^{\text{inv}(\sigma)}$. It was proved in [Bozejko & Speicher (1991)] that $c \in \mathcal{C}_m$. Therefore, the

so-called “ q -permanent”,²⁸ defined by $\text{per}_q(A) = d_c(A)$, $A \in \mathbf{C}_{m,m}$, satisfies $\text{per}_q(A) \geq \det(A)$, $A \in \mathcal{H}_m$.

- a. Prove that $\text{per}_1 = \text{per}$, the permanent function.
 - b. Prove that $\text{per}_0 = h$, Hadamard’s function, provided 0^0 is defined to be 1.
 - c. Prove that $\text{per}_{-1} = \det$, the determinant function.
44. Finish the proof of Theorem 7.38.
45. Prove Lemma 7.51.
46. If $A, B \in \mathbf{C}_{n,n}$, prove that
- a. $\det(A + B) = \sum_{k=0}^n \sum_{\alpha, \beta \in Q_{k,n}} (-1)^{r(\alpha)+r(\beta)} \det(A[\alpha|\beta]) \det(B(\alpha|\beta))$.
 - b. $\text{per}(A + B) = \sum_{k=0}^n \sum_{\alpha, \beta \in Q_{k,n}} \text{per}(A[\alpha|\beta]) \text{per}(B(\alpha|\beta))$.
47. Suppose $A \in \mathcal{H}_m$ has eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$. Let G be a subgroup of S_m and suppose $\chi \in I(G)$.
- a. Prove that $\min_{\alpha \in \bar{\Delta}} \prod_{i=1}^m \lambda_{\alpha(i)} \leq \bar{d}_\chi(A) \leq \max_{\alpha \in \bar{\Delta}} \prod_{i=1}^m \lambda_{\alpha(i)}$.
 - b. Show that the left-hand inequality in part (a) is inferior to Schur’s Inequality.
 - c. If χ is the principal character of S_m , show that the right-hand inequality in part (a) is inferior to the inequality in Exercise 12(a).
48. Suppose $A, B \in \mathcal{H}_m$ are positive definite. If $0 \leq \theta \leq 1$ then (see, for example, [Bellman (1953)])

$$\det(\theta A + (1 - \theta)B) \geq \det(A)^\theta \det(B)^{1-\theta}.$$

Use this fact as the basis for another proof of Hadamard’s Inequality.

49. Suppose χ is a linear character of the subgroup G of S_m . Let $A \in \mathbf{C}_{n,n}$. Assume $\bar{\Delta} \subset \Gamma_{m,n}$ is not empty.
- a. If A is invertible, prove that $K(A)$ is invertible.
 - b. If A is normal, prove that $K(A)$ is normal.
 - c. If $A \geq 0$, prove that $K(A) \geq 0$.

Excursion Exercise

50. Continuing from Example 7.62,

²⁸See [Bapat (1992)] and [Bapat & Lal (1994)], where it is proven that $\text{per}_q(A) \geq \text{per}_q(A_{11}) \text{per}_q(A_{22})$, when $A \in \mathcal{H}_m$ is partitioned as in Equation (7.1) and $q \in [0, 1]$; and where it is conjectured that for any fixed non-diagonal matrix $A \in \mathcal{H}_m$, $q \rightarrow \text{per}_q(A)$ is a strictly increasing function of q in the interval $[-1, 1]$. (A weaker conjecture is that, for any $q \in [-1, 1]$, $\text{per}_q(A) \leq \text{per}(A)$, $A \in \mathcal{H}_m$.)

- a. find three linearly independent homogeneous polynomial invariants of $A(G)$ of degree 6.
- b. determine the dimension of the space of homogeneous polynomial invariants of degree 8.
- c. speculate about a generating set for the algebra of polynomial invariants.

Application Exercises

51. Find a pair of invertible 2-by-2 matrices A and B such that

- a. $\text{per}(AB) = \text{per}(A) \text{per}(B)$.
- b. $\text{per}(AB) \neq \text{per}(A) \text{per}(B)$.
- c. $\text{per}(ABA^{-1}) \neq \text{per}(B)$.

52. Find a 3-by-3 invertible matrix A such that

$$\text{per}(A) = 0 \neq \text{per}(A^{-1}).$$

53. The **permanental roots** of an n -by- n matrix A are the roots of its **permanental polynomial**, $p(x) = \text{per}(xI_n - A)$. Suppose $A \in \mathcal{H}_n$.

- a. Show that $p(x)$ is a monic polynomial of degree n .
- b. Show that $p(0) = (-1)^n \text{per}(A)$.
- c. Show that the sum of the permanental roots of A is equal to the sum of its eigenvalues.
- d. Prove that the real permanental roots of A lie in the closed interval $[\lambda_n, \lambda_1]$, where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ are the eigenvalues of A .
- e. Prove or disprove that all permanental roots of A are real.

54. The **star** on $n > 1$ vertices is the unique tree T_n having one vertex of degree $n - 1$ and $n - 1$ vertices of degree 1. The multiplicity of 1 as a root of $\text{per}(xI_n - L(T_n))$ is $n - 2$ [Faria (1985)].

- a. Find the other two roots if $n = 3$.
- b. Find the other two roots if $n = 4$.
- c. Find the other two roots as a function of n .
- d. Prove that the permanental roots of $L(T_n)$, $n \geq 7$, are all real.

55. If T is a tree on n vertices, then [Chan, Lam & Tang (1996)] $\text{per}(L(T_n)) \leq \text{per}(L(T)) \leq \text{per}(L(P_n))$, where T_n is the star (see Exercise 54) and P_n is the **path**, that is, the unique graph on $n > 1$ vertices having two vertices of degree 1 and $n - 2$ vertices of degree 2. Confirm this result when $n = 5$.

56. Let n be a fixed positive integer. Show that there exists a coimmanantal family of at least $k = 2^{(n-1)/17}$ trees on n vertices. (Hint: Corollary 7.71.)
57. Let $G = (V, E)$ be a graph on $n > 2$ vertices. A **hamiltonian path** in G is an arrangement of the n vertices of V into a sequence $(v_1, v_2, \dots, v_n)$, such that $\{v_i, v_{i+1}\} \in E$, $1 \leq i < n$. If $\{v_1, v_n\} \in E$ as well, then the hamiltonian path is a **hamiltonian cycle**. Prove that the number of hamiltonian cycles in G is given by the formula

$$\frac{1}{2n} \sum_{r=2}^n (-1)^r d_r(L(G)).$$

(Hint: Exercises 53 and 54, Chapter 4.)

CHAPTER 8

The Rational Representations of $GL(n, \mathbb{C})$

The **general linear group** (or “full” linear group) is the multiplicative group $GL(n, \mathbb{C})$ of invertible n -by- n complex matrices. The main thrust of the chapter concerns representations $R: GL(n, \mathbb{C}) \rightarrow GL(r, \mathbb{C})$ in which the elements of $R(A)$ are fixed rational functions¹ (ratios of polynomials) in the elements of A . Classical Schur polynomials emerge in the role of characters associated with these representations. These polynomials turn out to be powerful tools in the study of irreducible characters of symmetric groups and their associated immanants.

We begin with a deep algebraic result, commonly known as Weyl’s Principle of the Irrelevance of Algebraic Inequalities.²

THEOREM 8.1 *Suppose $x_1, x_2, \dots, x_m$ are independent indeterminates over the complex numbers. Let $f, g_1, g_2, \dots, g_p \in \mathbb{C}[x_1, x_2, \dots, x_m]$, where none of $g_1, g_2, \dots, g_p$ is the zero polynomial. Suppose*

$$f(c_1, c_2, \dots, c_m) = 0,$$

whenever $c_1, c_2, \dots, c_m \in \mathbb{C}$ are such that

$$g_t(c_1, c_2, \dots, c_m) \neq 0, \quad 1 \leq t \leq p.$$

Then f is the zero polynomial.

¹“Fixed” means the rational functions depends on R but not on A .

²The version presented here is but a special case of Weyl’s Principle.

Proof Define $h \in \mathbb{C}[x_1, x_2, \dots, x_m]$ by

$$h(x_1, x_2, \dots, x_m) = f(x_1, x_2, \dots, x_m) \prod_{t=1}^p g_t(x_1, x_2, \dots, x_m).$$

Then $h(c_1, c_2, \dots, c_m) = 0$ for all $c_1, c_2, \dots, c_m \in \mathbb{C}$. In other words, h is the zero polynomial. Because $\mathbb{C}[x_1, x_2, \dots, x_m]$ is an integral domain, it has no zero divisors. Thus, one of the factors of h must be zero. Because $g_t \neq 0$, $1 \leq t \leq p$, it must be that $f = 0$. $\square$

The next result, interesting in its own right, is of critical importance to subsequent developments.

THEOREM 8.2 *Let $R: GL(n, \mathbb{C}) \rightarrow GL(1, \mathbb{C})$ be a polynomial representation of $GL(n, \mathbb{C})$. Then there is an integer $k \geq 0$ such that $R(A) = \det(A)^k$, $A \in GL(n, \mathbb{C})$.*

Proof Because R is a homomorphism, $R(I_n) = 1$. (We will freely abuse the language by confusing $GL(1, \mathbb{C})$ with $\mathbb{C} \setminus \{0\}$.) Therefore, for any $A \in GL(n, \mathbb{C})$,

$$\begin{aligned} 1 &= R(AA^{-1}) \\ &= R(A)R\left(\frac{1}{\det(A)} \operatorname{adj}(A)\right). \end{aligned} \quad (8.1)$$

Because $R(A)$ is a fixed polynomial in the entries of A , there exists a positive integer, say q , such that multiplying both sides of Equation (8.1) by $\det(A)^q$ produces $\det(A)^q = R(A)g(A)$, where

$$g(A) = \det(A)^q R\left(\frac{1}{\det(A)} \operatorname{adj}(A)\right)$$

is a polynomial function of the entries of A . Moreover, neither g nor the integer q depends on A . In other words, if $X = (x_{ij})$ is an n -by- n matrix whose entries are n^2 independent indeterminates over $\mathbb{C}$, then $\det(X)^q - R(X)g(X) = 0$, as long as $\det(X) \neq 0$. It follows from Weyl's Principle that

$$\det(X)^q = R(X)g(X), \quad (8.2)$$

for all X , that is, Equation (8.2) is a polynomial identity. Because $\det(X)$ is an irreducible polynomial (see Lemma 8.3), it follows from Equation (8.2) that $R(X) = c \det(X)^k$ for some $c \in \mathbb{C}$ and some nonnegative integer k . Substituting I_n for X yields $c = 1$. $\square$

LEMMA 8.3 If $X = (x_{ij})$ is an n -by- n matrix whose entries are n^2 independent indeterminates over $\mathbb{C}$, then $\det(X)$ is an irreducible element of the unique factorization domain $\mathbb{C}[x_{11}, x_{12}, \dots, x_{nn}]$.

The next result allows us to focus on polynomial representations.

THEOREM 8.4 Let $R: GL(n, \mathbb{C}) \rightarrow GL(r, \mathbb{C})$ be a rational representation. Then there exists an integer k and a polynomial representation $S: GL(n, \mathbb{C}) \rightarrow GL(r, \mathbb{C})$ such that $R(A) = \det(A)^k S(A)$, $A \in GL(n, \mathbb{C})$.

Proof Let $X = (x_{ij})$ be an n -by- n matrix whose entries are n^2 independent indeterminates over $\mathbb{C}$. Denote the (s, t) -entry of $R(X)$ by $R_{st}(X)$. Suppose $q(X)$ is a least common multiple of the denominators of the rational functions $R_{st}(X)$, $1 \leq s, t \leq r$. (Then $q(X)$ is unique up to multiplication by a nonzero complex number.) Let $p(X)$ be a greatest common divisor of the r^2 polynomials $q(X)R_{st}(X)$. Then we may write

$$R(X) = \frac{p(X)}{q(X)} S(X), \quad (8.3)$$

where the (s, t) -entry of $S(X)$ is a polynomial in $\mathbb{C}[x_{11}, x_{12}, \dots, x_{nn}]$, and a greatest common divisor of the entries of $S(X)$ is 1. (In case $r = 1$, we may take $S(X) = 1$.) Without loss of generality, we may assume $p(X)$ and $q(X)$ are relatively prime. Replacing $p(X)$ with $p(X)/p(I_n)$, $q(X)$ with $q(X)/q(I_n)$, and $S(X)$ with $p(I_n)S(X)/q(I_n)$ allows us to assume that $p(I_n) = q(I_n) = 1$. Because $R(I_n) = I_r$, it follows that $S(I_n) = I_r$.

In view of Theorem 8.2, it remains to show that p , q and S are representations of $GL(n, \mathbb{C})$. For any $A, B \in GL(n, \mathbb{C})$, $R(AB) = R(A)R(B)$. Substituting Equation (8.3) into this identity yields

$$p(AB)q(A)q(B)S(AB) - p(A)p(B)q(AB)S(A)S(B) = 0. \quad (8.4)$$

This matrix equation is equivalent to r^2 equations in which a polynomial function in the $2n^2$ entries of A and B is equal to zero, provided (only) that $\det(A) \neq 0 \neq \det(B)$. Therefore, from Weyl's Principle, each of the r^2 polynomials is identically zero. That is, Equation (8.4) is an identity for all n -by- n matrices. In particular, if $B \in GL(n, \mathbb{C})$ is fixed but arbitrary, and X is an n -by- n matrix of indeterminates, then

$$p(XB)q(X)q(B)S(XB)S(B)^{-1} = p(X)p(B)q(XB)S(X). \quad (8.5)$$

Because $q(X)$ is a factor of the left-hand side of this equation, it divides each entry of the r -by- r matrix on the right-hand side. Since $q(X)$ and $p(X)$ are relatively prime, $q(X)$ divides each entry of the matrix $q(XB)S(X)$. Because the entries

of $S(X)$ are relatively prime, it must be that $q(X)$ divides $q(XB)$. Because the degree of the polynomial $q(XB)$ cannot be larger than the degree of $q(X)$, there is a complex number $c(B)$ such that

$$q(XB) = q(X)c(B). \quad (8.6)$$

Substituting $X = I_n$ we have, because $q(I_n) = 1$, $c(B) = q(B)$. It follows that $q: GL(n, \mathbb{C}) \rightarrow GL(1, \mathbb{C})$ is a representation. In particular (Theorem 8.2), $q(X)$ is a nonnegative integral power of $\det(X)$.

Setting $A = X$ and $B = X^{-1}$ in the polynomial identity (8.4) we have (because $q(X^{-1}) = q(X)^{-1}$, $p(I_n) = 1$, and $S(I_n) = I_r$)

$$I_r = p(X)p(X^{-1})S(X)S(X^{-1}). \quad (8.7)$$

Because $X^{-1} = \det(X)^{-1} \operatorname{adj}(X)$ and the entries of $p(X^{-1})$ and $S(X^{-1})$ are polynomials in the entries of X^{-1} , there exists a nonnegative integer m such that

$$Q(X) = \det(X)^m p(X^{-1})S(X)S(X^{-1})$$

has polynomial entries. Therefore, from Equation (8.7),

$$p(X)Q(X) = \det(X)^m I_r.$$

It follows from this equation and Lemma 8.3 that $p(X)$ is a nonnegative integral power of $\det(X)$. Returning to Equation (8.3), we see that

$$R(X) = \det(X)^k S(X),$$

for some integer k . Therefore,

$$\begin{aligned} S(AB) &= \det(AB)^{-k} R(AB) \\ &= [\det(A)^{-k} R(A)][\det(B)^{-k} R(B)] \\ &= S(A)S(B), \end{aligned}$$

for all $A, B \in GL(n, \mathbb{C})$, that is, $S: GL(n, \mathbb{C}) \rightarrow GL(r, \mathbb{C})$ is a representation. $\square$

THEOREM 8.5 Any polynomial representation $P: GL(n, \mathbb{C}) \rightarrow GL(r, \mathbb{C})$ is equivalent to a direct sum of homogeneous polynomial representations.

Proof Let y and z be independent indeterminates over the complex numbers. Because P is a polynomial representation, there is a nonnegative integer k , and $k + 1$ matrices $C_0, C_1, \dots, C_k \in \mathbb{C}_{r,r}$ such that

$$P(yI_n) = C_0 + yC_1 + \dots + y^k C_k.$$

Similarly, $P(zI_n) = \sum z^i C_i$. Consequently,

$$\begin{aligned} \sum_{i=0}^k (yz)^i C_i &= P(yzI_n) \\ &= P(yI_n)P(zI_n) \\ &= \left(\sum_{i=0}^k y^i C_i \right) \left(\sum_{j=0}^k z^j C_j \right) \\ &= \sum_{i=0}^k (yz)^i C_i^2 + \sum_{i=0}^k \sum_{j \neq i}^k y^i z^j C_i C_j. \end{aligned} \quad (8.8)$$

Comparing coefficients in Equation (8.8), reveals that $C_0, C_1, \dots, C_k$ are annihilating idempotents. Because $P(I_n) = \sum C_i = I_r$, it follows that there is a fixed $U \in GL(r, \mathbb{C})$ such that $U^{-1}C_i U = B_i$, $0 \leq i \leq k$, where

$$B_i = 0_{r_0} \oplus \dots \oplus 0_{r_{i-1}} \oplus I_{r_i} \oplus 0_{r_{i+1}} \oplus \dots \oplus 0_{r_k},$$

$r_i = \text{rank}(C_i)$ and 0_t is the t -by- t zero matrix. (If $C_i = 0$, then $r_i = 0$ and $B_i = 0$.) Suppose that q of the C_i 's are nonzero, say $C_{m_1}, C_{m_2}, \dots, C_{m_q}$.

Let $X = (x_{ij})$ be an n -by- n matrix such that y together with the x_{ij} are $n^2 + 1$ independent indeterminates over $\mathbb{C}$. Partition the matrix $U^{-1}P(X)U = (P_{st}(X))$, where $P_{st}(X)$ is r_{m_s} -by- r_{m_t} , $1 \leq s, t \leq q$. Because $U^{-1}P(yI_n)U$ commutes with $U^{-1}P(X)U$,

$$y^{m_s} P_{st}(X) = y^{m_t} P_{st}(X),$$

$1 \leq s, t \leq q$. It follows that $P_{st}(X) = 0$, $s \neq t$. In other words,

$$U^{-1}P(X)U = P_{11}(X) \oplus P_{22}(X) \oplus \dots \oplus P_{qq}(X).$$

Therefore, $P_{ii} : GL(n, \mathbb{C}) \rightarrow GL(r_i, \mathbb{C})$ is a representation. Moreover, because the entries of $U^{-1}P(X)U$ are linear combinations of the elements of $P(X)$, P_{ii} is a polynomial representation of $GL(n, \mathbb{C})$. Finally,

$$\begin{aligned} P_{ii}(yX) &= P_{ii}(yI_n X) \\ &= P_{ii}(yI_n)P_{ii}(X). \end{aligned} \quad (8.9)$$

Because $P_{ii}(yI_n) = y^{m_i} I_{r_i}$, $1 \leq i \leq q$, we obtain $P_{ii}(yX) = y^{m_i} P_{ii}(X)$. In other words, the entries of $P_{ii}(X)$ are homogeneous polynomials of degree m_i in the entries of X . $\square$

EXAMPLE 8.6 Let $X = (x_{ij})$ be an n -by- n matrix of indeterminates. Recall (Definition 5.41) that the m -th Kronecker power of X is an n^m -by- n^m matrix whose rows and columns are indexed by $\Gamma_{m,n}$. The (α, β) -entry of $X^{\otimes m}$ is

$$(X^{\otimes m})_{\alpha, \beta} = \prod_{i=1}^m x_{\alpha(i)\beta(i)}.$$

Observe that every monomial of (total) degree m in the n^2 variables $x_{11}, x_{12}, \dots, x_{nn}$ occurs (at least once) as an entry of $X^{\otimes m}$. $\square$

Suppose $P: GL(n, \mathbb{C}) \rightarrow GL(n^m, \mathbb{C})$ is the representation defined by $P(A) = A^{\otimes m}$. Because $P(A)$ is an n^m -by- n^m matrix, the *degree* of the representation is n^m . At the same time, the entries of $P(A)$ are monomials of *degree* m in the entries of A . The two different uses of the word “degree” might be confusing. Let us agree to understand the phrase “ P is a homogeneous polynomial representation of degree m ” to mean that m is the degree of the polynomial functions and not the degree of the representation.

Because $GL(n, \mathbb{C})$ is an infinite group, Maschke’s Theorem cannot be applied to the representation $A \rightarrow P(A)$. Nevertheless (as we will see presently), P is fully reducible. Moreover, for any fixed $U \in GL(n^m, \mathbb{C})$, the entries of $U^{-1}P(A)U$ are *linear* functions (homogeneous polynomial functions of degree 1) in the entries of $A^{\otimes m}$ and, therefore, homogeneous polynomials of degree m in the entries of A . It is useful to look at this situation from another perspective.

LEMMA 8.7 Let m be a fixed but arbitrary positive integer. Consider $\mathcal{G} = \{A^{\otimes m} : A \in GL(n, \mathbb{C})\}$. Then $\mathcal{G}$ is a subgroup of $GL(n^m, \mathbb{C})$, and the identity mapping, $A^{\otimes m} \rightarrow A^{\otimes m}$, $A^{\otimes m} \in \mathcal{G}$, is a representation of $\mathcal{G}$. If $F: \mathcal{G} \rightarrow GL(r, \mathbb{C})$ is a homogeneous polynomial representation of $\mathcal{G}$ of degree 1 in the entries of $A^{\otimes m}$, then F is fully reducible, and each irreducible constituent of F is equivalent to a constituent of the reduction of $A^{\otimes m} \rightarrow A^{\otimes m}$.

That the irreducible constituents of F should be equivalent to constituents of the m -th Kronecker power representation, already plausible from Example 8.6, can be proved using a dimension argument [Macdonald (1995), p. 162]. The remainder of Lemma 8.7 is a special case of a more general result from the representation theory of semi-simple algebras, a proof of which can be found, for example, in [Green (1980), pp. 29–30] or [Marcus (1975), pp. 386–395].

Let $R: GL(n, \mathbb{C}) \rightarrow GL(r, \mathbb{C})$ be a homogeneous polynomial representation of degree m . Let $\mathcal{G} = \{A^{\otimes m} : A \in GL(n, \mathbb{C})\}$, and define a function $R_m: \mathcal{G} \rightarrow GL(r, \mathbb{C})$ by

$$R_m(A^{\otimes m}) = R(A). \quad (8.10)$$

Then

$$\begin{aligned}
 R_m(A^{\otimes m} B^{\otimes m}) &= R_m((AB)^{\otimes m}) \\
 &= R(AB) \\
 &= R(A)R(B) \\
 &= R_m(A^{\otimes m})R_m(B^{\otimes m}).
 \end{aligned} \tag{8.11}$$

Therefore, R_m is a representation of $\mathcal{G}$. Moreover, the entries of $R_m(A^{\otimes m})$ are linear functions of the entries of $A^{\otimes m}$. It follows from Lemma 8.7 that R_m is fully reducible and its irreducible constituents are equivalent to irreducible constituents of the representation $A^{\otimes m} \rightarrow A^{\otimes m}$. This means that the (arbitrary) representation we started with, namely $R: GL(n, \mathbb{C}) \rightarrow GL(r, \mathbb{C})$, is fully reducible, and its irreducible constituents are equivalent to irreducible constituents of $A \rightarrow A^{\otimes m}$, the m -th Kronecker power representation of $GL(n, \mathbb{C})$.

Let's summarize. Suppose $R: GL(n, \mathbb{C}) \rightarrow GL(r, \mathbb{C})$ is a rational representation of the general linear group $GL(n, \mathbb{C})$. By Theorems 8.4 and 8.5, there exists an integer k and homogeneous polynomial representations P_i , $1 \leq i \leq q$, such that R is equivalent to a representation R' defined by

$$R'(A) = \det(A)^k \begin{pmatrix} P_1(A) & 0 & \cdots & 0 \\ 0 & P_2(A) & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & P_q(A) \end{pmatrix}, \tag{8.12}$$

$A \in GL(n, \mathbb{C})$. Moreover, as a consequence of Lemma 8.7, for each i , $A \rightarrow P_i(A)$ is fully reducible and each of its irreducible constituents is equivalent to an irreducible constituent of some Kronecker power representation. Therefore, to completely comprehend the rational representations of the general linear group, it only remains to reduce its Kronecker power representations. (Taken together, the Kronecker power representations play a role that is analogous to the role of the regular representation in the theory for finite groups.) In determining the irreducible constituents of the m -th Kronecker power representation, it will simplify matters to replace $GL(n, \mathbb{C})$ with the isomorphic group $GL(V)$ of invertible linear operators on a vector space V of dimension n . From the operator perspective, constituents of $A \rightarrow A^{\otimes m}$ correspond to subspaces of $V^{\otimes m}$ that are invariant under $T^{\otimes m}$, $T \in GL(V)$, and irreducible constituents correspond to minimal invariant subspaces.

Recall that we have already discovered some invariant subspaces of $V^{\otimes m}$. If χ is an irreducible character of the subgroup G of S_m , then $V_\chi(G)$ is an invariant subspace of $T^{\otimes m} = T \otimes T \otimes \cdots \otimes T$ (m -times). Hence (abusing the language), $T \rightarrow K(T)$ is a (not necessarily irreducible) constituent of $T \rightarrow T^{\otimes m}$, $T \in GL(V)$. Because (Corollary 6.6)

$$V^{\otimes m} = \bigoplus_{\chi \in I(G)} V_\chi(G), \tag{8.13}$$

the direct sum of the symmetry classes of tensors afforded by G and its irreducible characters, $T \rightarrow T^{\otimes m}$ is a "direct sum" of "representations" of the form $T \rightarrow K(T)$ as χ ranges over $I(G)$. Thus, it suffices to find the minimal invariant subspaces of $V_\chi(G)$, that is, to "reduce" $T \rightarrow K(T)$, $T \in GL(V)$. We begin this final step by introducing a refinement of the projection operator $T(G, \chi)$.

DEFINITION 8.8 Let χ be an irreducible character of the subgroup G of S_m . Suppose $\sigma \rightarrow A(\sigma) = (a_{ij}(\sigma))$ is a representation of G that affords χ . For a fixed but arbitrary vector space V , denote by $T_i(G, A) \in L(V^{\otimes m}, V^{\otimes m})$, $1 \leq i \leq \chi(e)$, the operator defined by

$$T_i(G, A) = \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} a_{ii}(\sigma) P(\sigma).$$

THEOREM 8.9 Let χ be an irreducible character of the subgroup G of S_m . Suppose $\sigma \rightarrow A(\sigma) = (a_{ij}(\sigma))$ is a representation of G that affords χ . If V is a vector space, then

(i) $T_i(G, A)T_j(G, A) = \delta_{ij}T_i(G, A)$, $1 \leq i, j \leq \chi(e)$; and

(ii) $\sum_{i=1}^{\chi(e)} T_i(G, A) = T(G, \chi)$.

If V is an inner product space and A is a unitary representation, then

(iii) $T_i(G, A)$ is hermitian.

Proof We prove part (i):

$$\begin{aligned} T_i(G, A)T_j(G, A) &= \left(\frac{\chi(e)}{o(G)} \right)^2 \left(\sum_{\sigma \in G} a_{ii}(\sigma) P(\sigma) \right) \left(\sum_{\tau \in G} a_{jj}(\tau) P(\tau) \right) \\ &= \left(\frac{\chi(e)}{o(G)} \right)^2 \sum_{\sigma, \tau \in G} a_{ii}(\sigma) a_{jj}(\tau) P(\sigma\tau) \\ &= \left(\frac{\chi(e)}{o(G)} \right)^2 \sum_{\tau \in G} \left(\sum_{\sigma \in G} a_{ii}(\sigma) a_{jj}(\sigma^{-1}\tau) \right) P(\tau) \\ &= \frac{\chi(e)}{o(G)} \sum_{\tau \in G} \delta_{ij} a_{ii}(\tau) P(\tau) \\ &= \delta_{ij} T_i(G, A), \end{aligned}$$

by Theorem 4.21. □

EXAMPLE 8.10 Consider the representation of S_3 from Example 4.3, namely,

$$\begin{aligned} A(e_3) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, & A((12)) &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ A((23)) &= \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix}, & A((123)) &= \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}, \\ A((132)) &= \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}, & \text{and } A((13)) &= \begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix}. \end{aligned}$$

Then

$$T_1(S_3, A) = \frac{1}{3}(P(e_3) - P((23)) - P((132)) + P((13))), \quad (8.14)$$

and

$$T_2(S_3, A) = \frac{1}{3}(P(e_3) + P((23)) - P((123)) - P((13))). \quad (8.15)$$

□

DEFINITION 8.11 Let χ be an irreducible character of the subgroup G of S_m . Suppose $\sigma \rightarrow A(\sigma) = (a_{ij}(\sigma))$ is a representation of G that affords χ . For a fixed but arbitrary vector space V , denote the image of $T_i(G, A)$ by $V_A^i(G)$.

COROLLARY 8.12 Let χ be an irreducible character of the subgroup G of S_m . Suppose $\sigma \rightarrow A(\sigma) = (a_{ij}(\sigma))$ is a representation of G that affords χ . If V is a vector space, then

$$V_\chi(G) = \bigoplus_{i=1}^{\chi(e)} V_A^i(G). \quad (8.16)$$

Moreover, if A is unitary and V is an inner product space, then the direct sum in Equation (8.16) is orthogonal.

Proof The result is an immediate consequence of Theorem 8.9 and the definitions. □

Because $T^{\otimes m}$ commutes with $P(\sigma)$, $\sigma \in S_m$, it commutes with $T_i(G, A)$. Therefore, $V_A^i(G)$ is an invariant subspace of $T^{\otimes m}$ and, hence, of $K(T)$. Denote the restriction of $K(T)$ to $V_A^i(G)$ by $K_A^i(T)$, $T \in L(V, V)$.

THEOREM 8.13 Let χ be an irreducible character of the subgroup G of S_m . Suppose $\sigma \rightarrow A(\sigma) = (a_{ij}(\sigma))$ is a representation of G affording χ . Let V be a vector space. If $V_\chi(G) \neq \{0\}$, then K_A^i and K_A^j are equivalent representations of $GL(V)$, $1 \leq i \leq j \leq \chi(e)$. If $G = S_m$, then $T \rightarrow K_A^i(T)$ is an irreducible representation of $GL(V)$.

If $G = S_m$ and $\chi = \chi_\pi$ where $\pi = [\pi_1, \pi_2, \dots, \pi_r] \vdash m$, then (Corollary 6.38) $V_\chi(S_m) \neq \{0\}$ if and only if $r \leq \dim(V)$.

Proof Because (Lemma 6.2) $P(\sigma)$ commutes with $T(G, \chi)$, $V_\chi(G)$ is an invariant subspace of $P(\sigma)$, $\sigma \in G$. Denote the restriction of $P(\sigma)$ to $V_\chi(G)$, by $P_\chi(\sigma)$, $\sigma \in G$. Then, $\sigma \rightarrow P_\chi(\sigma)$ is a representation of G , and

$$\frac{\chi(e)}{o(G)} \sum_{\sigma \in G} \chi(\sigma) P_\chi(\sigma) = I_{V_\chi(G)},$$

is the identity operator on $V_\chi(G)$.

Suppose ξ is an irreducible character of G different from χ . Then

$$Z = \frac{\xi(e)}{o(G)} \sum_{\sigma \in G} \xi(\sigma) P_\chi(\sigma)$$

is a linear operator on $V_\chi(G)$. Observe that

$$\begin{aligned} Z &= Z \circ I_{V_\chi(G)} \\ &= \left(\frac{\xi(e)}{o(G)} \sum_{\sigma \in G} \xi(\sigma) P_\chi(\sigma) \right) \left(\frac{\chi(e)}{o(G)} \sum_{\tau \in G} \chi(\tau) P_\chi(\tau) \right) \\ &= \frac{\xi(e)\chi(e)}{o(G)^2} \sum_{\mu \in G} \left(\sum_{\sigma \in G} \xi(\sigma) \chi(\sigma^{-1}\mu) \right) P_\chi(\mu) \\ &= 0, \end{aligned}$$

by Theorem 4.26. Therefore, $0 = \text{tr}(Z) = \xi(e)(\bar{\xi}, \eta)_G$, where η is the character afforded by $\sigma \rightarrow P_\chi(\sigma)$, $\sigma \in G$. Evidently, the restriction of η to G contains no irreducible character of G different from $\bar{\chi}$. It follows that there exists a basis B of $V_\chi(G)$ such that the matrix representation of $P_\chi(\sigma)$ with respect to B is the direct sum of (the contragredient representation) $C(\sigma) = A(\sigma^{-1})^t$ with itself $N = (\bar{\chi}, \eta)_G$ times. In other words, with respect to B , the matrix representation of $P_\chi(\sigma)$ is

$$[P_\chi(\sigma)]_B = I_N \otimes C(\sigma), \quad \sigma \in G. \quad (8.17)$$

If we abuse the language by confusing $T_i(G, A)$ with its restriction to $V_\chi(G)$, then, by Equation (8.17) and the Schur Relations,

$$\begin{aligned} [T_i(G, A)]_B &= \frac{\chi(e)}{o(G)} \sum_{\sigma \in G} a_{ii}(\sigma) (I_N \otimes C(\sigma)) \\ &= I_N \otimes \left(\frac{\chi(e)}{o(G)} \sum_{\sigma \in G} a_{ii}(\sigma) A(\sigma^{-1})^t \right) \\ &= I_N \otimes E_{ii}, \end{aligned} \quad (8.18)$$

where E_{ii} is the $\chi(e)$ -by- $\chi(e)$ matrix whose only nonzero entry is a 1 in position (i, i) .

Let $T \in L(V, V)$ be a (not necessarily invertible) linear operator on V . Partition the matrix $[K(T)]_B = (K_{st})$ into N^2 blocks, K_{st} , of size $\chi(e)$ -by- $\chi(e)$. Because $[K(T)]_B$ commutes with $[P_\chi(\sigma)]_B = I_N \otimes C(\sigma)$, and because C is irreducible, Corollary 4.17 implies that K_{st} is a multiple of $I_{\chi(e)}$, $1 \leq s, t \leq N$. In other words,

$$[K(T)]_B = B(T) \otimes I_{\chi(e)} \quad (8.19)$$

for some N -by- N matrix $B(T)$. By Exercise 21, Chapter 5, there exists a permutation matrix Q such that

$$Q^t(B \otimes C)Q = C \otimes B$$

for all $B \in \mathbb{C}_{N,N}$, and all $C \in \mathbb{C}_{\chi(e), \chi(e)}$. Similarity by Q merely permutes the elements of the ordered basis B into a new ordered basis B' . Thus, from Equation (8.18),

$$[T_i(G, A)]_{B'} = E_{ii} \otimes I_N.$$

It follows that the first N elements of B' form a basis B_1 of $V_A^1(G)$, the second N elements form a basis B_2 of $V_A^2(G)$, and so on. Applying this observation to

$$[K(T)]_{B'} = I_{\chi(e)} \otimes B(T),$$

we deduce that, with respect to B_i , the matrix representation of $K_A^i(T)$ is $B(T)$, $1 \leq i \leq \chi(e)$. Therefore, K_A^i and K_A^j are equivalent.

We now consider the case in which $G = S_m$. Recall (Definition 6.72) that a linear operator on $V^{\otimes m}$ is bisymmetric if it commutes with $P(\sigma)$, $\sigma \in S_m$. By Corollary 6.73, the space of bisymmetric operators is $\langle T^{\otimes m} : T \in L(V, V) \rangle$. The local version of the result is that $\langle K(T) : T \in L(V, V) \rangle$ is the set of linear operators that commute with $\{P_\chi(\sigma) : \sigma \in S_m\}$. In matrix terms,

$$\{[K(T)]_B = B(T) \otimes I_{\chi(e)} : T \in L(V, V)\}$$

is the set of matrices that commute with

$$\{[P_\chi(\sigma)]_B = I_N \otimes C(\sigma) : \sigma \in S_m\}.$$

However, by Corollary 4.17, the matrices that commute with

$$\{I_N \otimes C(\sigma) : \sigma \in S_m\}$$

are those of the form $C \otimes I_{\chi(e)}$, $C \in \mathbb{C}_{N,N}$. It follows that $\mathbb{C}_{N,N} = \langle B(T) : T \in L(V, V) \rangle$. Therefore, $\{B(T) : T \in L(V, V)\}$ is irreducible. What we need to prove is that $\{B(T) : T \in GL(V)\}$ is irreducible. This will require one more application of Weyl's Principle. Fix an arbitrary basis of V . Denote by $[T]$ the matrix representation of $T \in L(V, V)$ with respect to this basis. Then there exists a fixed invertible matrix M such that $B(T)$ is a principal submatrix of $M^{-1}([T]^{\otimes m})M$, $T \in L(V, V)$. In particular, the elements of $B(T)$ are homogeneous polynomials of degree m in the entries of $[T]$. Let $U \in GL(n, \mathbb{C})$ be fixed but arbitrary. Denote the $(1, N)$ -entry of $U^{-1}B(T)U$ by $p([T])$, $T \in L(V, V)$. Suppose $p([T]) = 0$ for all $T \in GL(V)$. Then the polynomial function $p([T]) = 0$ whenever $\det([T]) \neq 0$. By Weyl's Principle, $p([T]) = 0$ for all $T \in L(V, V)$, contradicting the fact that $\{B(T) : T \in L(V, V)\}$ is irreducible. Because U was arbitrary, it follows that $\{B(T) : T \in GL(V)\}$ is irreducible. $\square$

DEFINITION 8.14 Let χ be an irreducible character of the subgroup G of S_m . Suppose $\sigma \rightarrow A(\sigma) = (a_{ij}(\sigma))$ is a representation of G affording χ . Let V be a vector space. Denote the restriction of $T^{\otimes m}$ to $V_A^1(G)$ by $B_\chi^G(T)$, $T \in L(V, V)$. The abbreviation $B_\chi^m(T)$ will be used in place of $B_\chi^m(T)$.

The matrix $B(T)$ that occurs, for example, in Equation (8.19) is a matrix representation of the linear transformation $B_\chi^G(T)$ introduced in Definition 8.14.

It follows from Theorem 8.13 that, as long as $V_\chi(G) \neq \{0\}$, the representation $T \rightarrow K(T)$, $T \in GL(V)$, is equivalent to the direct sum of $B_\chi^G(T)$ with itself $\chi(e)$ times. If $G = S_m$, this direct sum constitutes a reduction of $T \rightarrow K(T)$. (Again, we are guilty of using matrix language in the context of linear operators.) Given representations of S_m corresponding to each of its irreducible characters, Theorem 8.13 completes the reduction of $T \rightarrow T^{\otimes m}$, $T \in GL(V)$, and, therefore, the description of the rational representations of $GL(n, \mathbb{C})$. (Explicit constructions can be found, for example, in [Boerner (1970)], [Dias da Silva (1981)], [Hamermesh (1962)], [Hunter (1983)], [James & Kerber 1981)], [Littlewood (1958)], [Marcus (1975)] and/or [Sagan (1991)].) More remarkable than any of the details is the elegant relationship that has emerged between the representations of the finite symmetric groups and the homogeneous polynomial representations of the infinite group $GL(n, \mathbb{C})$.

Theorem 8.13 raises new questions even as it answers old ones. For example, in order to avoid the awkward step of having to replace $GL(n, \mathbb{C})$ with $GL(V)$, we need to know more about induced bases of $V_\chi(G)$ and $V_A^i(G)$ when $\chi(e) > 1$. Another question left unanswered by Theorem 8.13 concerns the reduction of $T \rightarrow B_\chi^G(T)$, $T \in GL(V)$, when $G \neq S_m$. In fact, something can be said about this situation.

THEOREM 8.15 Let ξ be an irreducible character of the subgroup G of S_m . Suppose V is a vector space of dimension n . Then the representation $T \rightarrow B_\xi^G(T)$, $T \in GL(V)$, is equivalent to a direct sum of irreducible constituents B_χ^m , where

χ ranges over those characters χ_π of S_m corresponding to partitions $\pi \vdash m$ of length $L(\pi) \leq n$. If $L(\pi) \leq n$, then the multiplicity of $B_{\chi_\pi}^m$ as an irreducible constituent of B_ξ^G is $(\xi, \chi_\pi)_G$. That is, the number of occurrences of $B_{\chi_\pi}^m$ in B_ξ^G is equal to the number of occurrences of ξ in the restriction of χ_π to G .

Proof Let $\sigma \rightarrow A(\sigma) = (a_{ij}(\sigma))$ be an irreducible representation of G that affords ξ . Suppose $\sigma \rightarrow R(\sigma) = (r_{ij}(\sigma))$ is an irreducible representation of S_m that affords χ . Without loss of generality, we may assume that the restriction of R to G is fully reduced, and that any components of the restriction that afford ξ are equal to A . In these terms, the problem is to express the vector space $V_A^j(G)$ as a direct sum of $V_R^j(S_m)$. Observe that

$$\begin{aligned} T_i(G, A)T_j(S_m, R) &= \frac{\xi(e)\chi(e)}{o(G)m!} \sum_{\tau \in S_m} \left(\sum_{\sigma \in G} a_{ii}(\sigma)r_{jj}(\sigma^{-1}\tau) \right) P(\tau) \\ &= \begin{cases} T_j(S_m, R), & \text{if } r_{jj|G} = a_{ii} \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

As j runs from 1 to $\chi(e)$, the restriction of r_{jj} to G will equal a_{ii} exactly $(\xi, \chi)_G$ times. $\square$

It can happen that $B_\xi^G = B_\chi^m$, in other words, that B_ξ^G is an irreducible representation of $GL(V)$ even when G is a proper subgroup of S_m . When $m \leq n$ and $\xi(e) = 1$, a complete list of such pairs (G, ξ) was obtained in [Djoković & Malzan (1975)].

6	4	1
4	2	
3	1	
1		

FIGURE 8.1

Suppose $\pi = [\pi_1, \pi_2, \dots, \pi_r] \vdash m$. Recall that to each ordered pair (i, j) , $1 \leq i \leq r$, $1 \leq j \leq \pi_i$, there corresponds a box, B_{ij} , in the Ferrers diagram of $F(\pi)$. Box B_{ij} determines a unique hook in $F(\pi)$ consisting of B_{ij} , all the boxes in row i of $F(\pi)$ to the right of B_{ij} , and all boxes in column j of $F(\pi)$ below it. The number of boxes in the hook determined by B_{ij} is its hook length, $h_{ij}(\pi) = 1 + (\pi_i - i) + (\pi_j^* - j)$. By the Frame-Robinson-Thrall hook length formula, the degree of the irreducible character χ_π of S_m is $\chi_\pi(e) = m! / \prod h_{ij}(\pi)$.

If $\pi = [3, 2^2, 1]$, Figure 8.1 illustrates $F([3, 2^2, 1])$, its boxes filled with their hook lengths. In this case,

$$\begin{aligned}\chi_\pi(e) &= \frac{8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{6 \times 4 \times 1 \times 4 \times 2 \times 3 \times 1 \times 1} \\ &= 70.\end{aligned}$$

The next result involves filling box B_{ij} of $F(\pi)$, not with numbers but with monomials. Figure 8.2 illustrates $F([3, 2^2, 1])$ with $x - i + j$ filling box (i, j) .

x	$x + 1$	$x + 2$
$x - 1$	x	
$x - 2$	$x - 1$	
$x - 3$		

FIGURE 8.2

DEFINITION 8.16 If $\pi = [\pi_1, \pi_2, \dots, \pi_r] \vdash m$, let

$$f_\pi(x) = \prod_{i=1}^r \prod_{j=1}^{\pi_i} (x - i + j).$$

Not to be confused with $f(\pi)$, the length of the diagonal of $F(\pi)$, $f_\pi(x)$ is a polynomial of degree m .

EXAMPLE 8.17 If $\pi = [3, 2^2, 1]$ then, from Figure 8.2,

$$f_\pi(x) = (x - 3)(x - 2)(x - 1)^2 x^2 (x + 1)(x + 2).$$

□

THEOREM 8.18³ Suppose $\pi = [\pi_1, \pi_2, \dots, \pi_r] \vdash m$. Let $A \rightarrow A(\sigma) = (a_{ij}(\sigma))$ be an irreducible representation of S_m that affords the character $\chi = \chi_\pi$. Suppose

³This "Frame-Robinson-Thrall type" formula emerges from the relationship between the representations of S_m and the homogeneous polynomial representations of degree m of $GL(n, \mathbb{C})$. Proofs can be found, e.g., in [Boerner (1970)], [James & Kerber (1981)], [King (1970)], [Littlewood (1958)], or [Murtaza and Rashid (1973)]. Values for $\dim(V_\lambda^1(S_m))$ were tabulated by P.H. Butler in his Appendix to [Wybourne (1970)].

V is a vector space of dimension n . Then the dimension of $V_A^1(S_m)$ (and, hence, of $V_A^k(S_m)$, $1 \leq k \leq \chi(e)$) is given by the formula

$$\begin{aligned} \dim(V_A^1(S_m)) &= \chi_\pi(e) f_\pi(n)/m! \\ &= \prod_{i=1}^r \prod_{j=1}^{\pi_i} (n - i + j) / h_{ij}(\pi). \end{aligned} \quad (8.20)$$

It follows from Definition 8.14 and Theorem 8.18 that the degree of the irreducible representation B_χ^m of $GL(V)$ corresponding to $\chi = \chi_\pi$ is given by Equation (8.20).

COROLLARY 8.19 Suppose V is a vector space of dimension n . Let ξ be an irreducible character of the subgroup G of S_m . Then

$$\dim(V_\xi(G)) = \sum_{\pi \vdash m} \xi(e) \chi_\pi(e) (\xi, \chi_\pi)_G f_\pi(n)/m!. \quad (8.21)$$

Proof Let A be a representation of G affording ξ . By Corollary 8.12,

$$V_\xi(G) = \bigoplus_{i=1}^{\xi(e)} V_A^i(G).$$

By Theorem 8.13, each of the $\xi(e)$ direct summands has the same dimension, namely, the degree of B_ξ^G . By Theorem 8.15, this dimension is equal to the sum over $\pi \vdash m$ of $(\xi, \chi_\pi)_G$ times the degree of $B_{\chi_\pi}^m$ ($L(\pi) \leq n$ if and only if $f_\pi(n) \neq 0$). This degree is given by Equation (8.20). $\square$

EXAMPLE 8.20 Suppose V is a vector space of dimension n . When $G = S_m$ and $\xi = \chi_\rho$, Equation (8.21) becomes

$$\dim(V_{\chi_\rho}(S_m)) = \chi_\rho(e)^2 f_\rho(n)/m!. \quad (8.22)$$

In accordance with Equation (8.16), this is $\chi_\rho(e)$ times the dimension of the minimal invariant subspace given in Equation (8.20).

Suppose, for example, that $\rho = [1^m]$. Then, $K_A^i(T) = K(T) = C_m(T)$, the m -th compound. The degree of the m -th compound is $\dim(\wedge^m V) = C(n, m)$. Equation (8.22) gives

$$\begin{aligned} \dim(V_\varepsilon(S_m)) &= f_{[1^m]}(n)/m! \\ &= n(n-1) \times \cdots \times (n-m+1)/m! \\ &= C(n, m). \end{aligned}$$

If $\rho = [m]$, then $K_A^i(T) = K(T) = P_m(T)$, the m -th induced power, the degree of which is $C(m + n - 1, m)$. Equation (8.22) gives

$$\begin{aligned} \dim(V_1(S_m)) &= f_{[m]}(n)/m! \\ &= n(n+1) \times \cdots \times (n+m-1)/m! \\ &= C(n+m-1, m). \end{aligned}$$

□

Other results pertaining to dimensions of symmetry classes of tensors can be found in [Chan (1978a&b) & (1979a&b)], [Chan & Lim (1980)], [Chang (1976)], [Cummings (1976)], [Cummings & Robinson (1976)], [Marcus & Chollet (1982)], and [Westwick (1970)].

We turn next to the character of the representation B_χ^G . Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of $T \in GL(V)$ (multiplicities included). Then (Equation (7.39))

$$\begin{aligned} \text{tr}(K(T)) &= \sum_{\alpha \in \bar{\Delta}} \prod_{i=1}^m \lambda_{\alpha(i)} \\ &= \chi(e) \sum_{\alpha \in \bar{\Delta}} (\chi, 1)_{G_*} \prod_{i=1}^m \lambda_{\alpha(i)} \\ &= \chi(e) \sum_{\alpha \in \bar{\Delta}} (\chi, 1)_{G_*} \prod_{i=1}^m \lambda_{\alpha(i)}. \end{aligned} \quad (8.23)$$

By Theorem 8.13 (and Definition 8.14), $T \rightarrow K(T)$ is equivalent to the direct sum of $\chi(e)$ copies of $B_\chi^G(T)$. Thus, from Equation (8.23),

$$\text{tr}(B_\chi^G(T)) = \sum_{\alpha \in \bar{\Delta}} (\chi, 1)_{G_*} \prod_{i=1}^m \lambda_{\alpha(i)}.$$

If $G = S_m$, then $\chi = \chi_\pi$ for some $\pi \vdash m$, and this identity becomes

$$\text{tr}(B_\chi^m(T)) = \sum_{\alpha \in \bar{G}_{m,n}} K_{\pi, \mu(\alpha)} \prod_{i=1}^m \lambda_{\alpha(i)}, \quad (8.24)$$

where $\mu(\alpha)$ is the multiplicity partition of α , and the Kostka coefficient $K_{\pi, \mu(\alpha)} = (\chi_\pi, 1)_{\mu(\alpha)}$, the number of occurrences of the principal character in the restriction of χ_π to the Young subgroup $S_{\mu(\alpha)}$.

DEFINITION 8.21 Let $x_1, x_2, \dots, x_n$ be independent indeterminates over $\mathbb{C}$. For each $\pi \vdash m$, the corresponding Schur polynomial is defined by

$$s_\pi(x_1, x_2, \dots, x_n) = \sum_{\alpha \in G_{m,n}} K_{\pi, \mu(\alpha)} \prod_{i=1}^m x_{\alpha(i)}. \quad (8.25)$$

THEOREM 8.22 Let $\chi = \chi_\pi$ be the irreducible character of S_m corresponding to $\pi \vdash m$. Suppose V is a vector space of dimension $n \geq L(\pi)$. If $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of $T \in L(V, V)$, then $\text{tr}(K(T)) = \chi(e)s_\pi(\lambda_1, \lambda_2, \dots, \lambda_n)$.

Theorem 8.22 amounts to putting old wine in a new bottle.

EXAMPLE 8.23 From Example 8.20, $B_\varepsilon^m(T) = C_m(T)$ and $B_1^m(T) = P_m(T)$. From Equation (7.41), $\text{tr}(C_m(T)) = E_m(x_1, x_2, \dots, x_n)$, the m -th elementary symmetric function. Evidently,

$$\begin{aligned} s_{[1^m]}(x_1, x_2, \dots, x_n) &= E_m(x_1, x_2, \dots, x_n) \\ &= M_{[1^m]}(x_1, x_2, \dots, x_n), \end{aligned}$$

a monomial symmetric function.⁴ From Equation (7.42), $\text{tr}(P_m(T)) = H_m(x_1, x_2, \dots, x_n)$, the m -th homogeneous symmetric function. Therefore,

$$\begin{aligned} s_{[m]}(x_1, x_2, \dots, x_n) &= H_m(x_1, x_2, \dots, x_n) \\ &= \sum_{\pi \vdash m} M_\pi(x_1, x_2, \dots, x_n). \end{aligned} \quad \square$$

THEOREM 8.24 Let $x_1, x_2, \dots, x_n$ be independent indeterminates over $\mathbb{C}$. If $\pi \vdash m$, then the Schur polynomial

$$s_\pi(x_1, x_2, \dots, x_n) = \sum_{\rho \vdash m} K_{\pi, \rho} M_\rho(x_1, x_2, \dots, x_n), \quad (8.26)$$

where $M_\rho(x_1, x_2, \dots, x_n)$ is the monomial symmetric function corresponding to ρ .

Proof

$$\sum_{\alpha \in G_{m,n}} K_{\pi, \mu(\alpha)} \prod_{i=1}^m x_{\alpha(i)} = \sum_{\rho \vdash m} K_{\pi, \rho} \sum_{(\rho)} \prod_{i=1}^m x_{\alpha(i)},$$

⁴See Definition 1.15.

where the final summation on the right-hand side is over those $\alpha \in G_{m,n}$ whose multiplicity partition $\mu(\alpha) = \rho$, that is,

$$\sum_{(\rho)} \prod_{i=1}^m x_{\alpha(i)} = \sum x_1^{k_1} x_2^{k_2} \dots x_n^{k_n},$$

the sum over all different rearrangements, $(k_1, k_2, \dots, k_n)$, of the n -tuple $(m_1, m_2, \dots, m_r, 0, 0, \dots, 0)$ obtained by appending $n - r$ zeros to the end of ρ . $\square$

Already evident from the definition (and/or Lemma 7.51), Equation (8.26) shows that $s_\pi(x_1, x_2, \dots, x_n)$ is symmetric for all $\pi \vdash m$.

EXAMPLE 8.25 Suppose $m = 3$ and $\pi = [2, 1]$. Then $K_{[2,1],[3]} = 0$ because $[2, 1]$ does not majorize $[3]$, $K_{[2,1],[2,1]} = 1$ by Corollary 4.54, and $K_{[2,1],[1^3]} = 2 = \chi_{[2,1]}(e)$. Therefore,

$$s_{[2,1]}(x_1, x_2, \dots, x_n) = M_{[2,1]}(x_1, x_2, \dots, x_n) + 2M_{[1^3]}(x_1, x_2, \dots, x_n).$$

From $M_{[2,1]}(x, y) = x^2y + xy^2$ and $M_{[1^3]}(x, y) = E_3(x, y) = 0$, we obtain the explicit formula $s_{[2,1]}(x, y) = x^2y + xy^2$. When $n = 3$,

$$s_{[2,1]}(x, y, z) = x^2y + x^2z + xy^2 + xz^2 + y^2z + yz^2 + 2xyz. \quad (8.27)$$

If V is a vector space of dimension 3 and $\chi = \chi_{[2,1]}$, the $\overline{\Delta}$ set for $V_\chi(S_3)$ can be read off from Equation (8.27):

$$\overline{\Delta} = \{(1, 1, 2), (1, 1, 3), (1, 2, 2), (1, 2, 3), (1, 3, 3), (2, 2, 3), (2, 3, 3)\}.$$

From Figure 8.3, $f_{[2,1]}(x) = x^3 - x$, so (Theorem 8.18) the degree of B_χ^m is $\chi_{[2,1]}(e)f_{[2,1]}(3)/3! = (2 \times 24)/6 = 8$, a value confirmed by Equation (8.27):

$$\begin{aligned} \text{tr}(B_\chi^m(I_3)) &= s_{[2,1]}(1, 1, 1) \\ &= 8. \end{aligned}$$

$\square$

x	$x + 1$
$x - 1$	

FIGURE 8.3

Because $M_\rho(x_1, x_2, \dots, x_n) = 0$ whenever $L(\rho) > n$, Theorem 8.24 allows us to be more casual about the number of variables. In particular, it does little harm to express the Schur polynomials as

$$s_\pi = \sum_{\rho \vdash m} K_{\pi, \rho} M_\rho, \quad \pi \vdash m. \quad (8.28)$$

From Equations (6.44)–(6.45), we have the following alternative expression for s_π :

$$s_\pi = \frac{1}{m!} \sum_{\sigma \in S_m} \chi_\pi(\sigma) \prod_{i=1}^m P_i^{c_i(\sigma)}, \quad (8.29)$$

where, recall, $P_t = M_{[t]}(x_1, x_2, \dots, x_n) = x_1^t + x_2^t + \dots + x_n^t$ is the t -th power sum. If $\pi = [1^m]$, then $\chi_\pi = \varepsilon$ and $s_\pi = E_m$. Evidently,

$$E_m = \frac{1}{m!} \sum_{\sigma \in S_m} \varepsilon(\sigma) \prod_{i=1}^m P_i^{c_i(\sigma)}. \quad (8.30)$$

From Newton's identities (Exercise 17, Chapter 1), Equation (8.30) can be expressed in the form

$$E_m = \frac{1}{m!} \det(L_m), \quad (8.31)$$

where

$$L_m = \begin{pmatrix} P_1 & 1 & 0 & 0 & \dots & 0 \\ P_2 & P_1 & 2 & 0 & \dots & 0 \\ P_3 & P_2 & P_1 & 3 & \dots & 0 \\ & & & & \dots & \\ P_{m-1} & P_{m-2} & P_{m-3} & P_{m-4} & \dots & m-1 \\ P_m & P_{m-1} & P_{m-2} & P_{m-3} & \dots & P_1 \end{pmatrix}. \quad (8.32)$$

Equations (8.29)–(8.32) lead to an identity for Schur polynomials in terms of immanants.

THEOREM 8.26 *Suppose $\pi \vdash m$. Let $\chi = \chi_\pi$ be the corresponding irreducible character of S_m . Then*

$$s_\pi = \frac{1}{m!} d_\chi(L_m). \quad (8.33)$$

EXAMPLE 8.27 Suppose $\pi = [2, 1] \vdash 3$. Let $\chi = \chi_\pi$. Then $\chi(\sigma) = F(\sigma) - 1$, $\sigma \in S_3$, where $F(\sigma)$ is the number of fixed points of σ . By Equation (8.33),

$$\begin{aligned} s_\pi &= \frac{1}{6} d_\chi \begin{pmatrix} P_1 & 1 & 0 \\ P_2 & P_1 & 2 \\ P_3 & P_2 & P_1 \end{pmatrix} \\ &= [2P_1^3 - 2P_3]/6. \end{aligned}$$

By definition, $P_3 = M_{[3]}$. From the Multinomial Theorem,⁵ $P_1^3 = M_{[3]} + 3M_{[2,1]} + 6M_{[1^3]}$. Therefore, $s_{[2,1]} = M_{[2,1]} + 2M_{[1^3]}$, confirming the first part of Example 8.25. $\square$

It follows from Equation (8.30) and the Fundamental Theorem of Symmetric Functions that any polynomial, symmetric in the variables $x_1, x_2, \dots, x_n$, is a polynomial in the power sum functions

$$\begin{aligned} P_t &= P_t(x_1, x_2, \dots, x_n) \\ &= x_1^t + x_2^t + \dots + x_n^t. \end{aligned}$$

THEOREM 8.28 Let $x_1, x_2, \dots, x_n$ be n independent indeterminates over the complex numbers. Set $s_\pi = s_\pi(x_1, x_2, \dots, x_n)$, $\pi \vdash m$. Then $\{s_\pi : \pi \vdash m, L(\pi) \leq n\}$ is a basis of $SC_m[x_1, x_2, \dots, x_n]$, the symmetric homogeneous polynomials of degree m in $x_1, x_2, \dots, x_n$.

Proof Recall (Theorem 1.27) that $\dim(SC_m[x_1, x_2, \dots, x_n]) = o(\{\pi \vdash m : L(\pi) \leq n\})$. Because $s_\pi(x_1, x_2, \dots, x_n) = 0$, $L(\pi) > n$, it suffices to show that $\{s_\pi : \pi \vdash m\}$ spans $SC_m[x_1, x_2, \dots, x_n]$.

Multiply Equation (8.29) by $\chi_\pi(\tau)$ and sum on π to obtain

$$\sum_{\pi \vdash m} \chi_\pi(\tau) s_\pi = \sum_{\sigma \in S_m} \left(\frac{1}{m!} \sum_{\pi \vdash m} \chi_\pi(\tau) \chi_\pi(\sigma) \right) \prod_{i=1}^m P_i^{c_i(\sigma)}.$$

Since $\chi_\pi(\tau^{-1}) = \chi_\pi(\tau)$, $\tau \in S_m$, it follows from the Orthogonality Relations of the Second Kind that

$$\sum_{\pi \vdash m} \chi_\pi(\tau) s_\pi = \prod_{i=1}^m P_i^{c_i(\tau)}. \quad (8.34)$$

⁵See Theorem 1.13.

Because any homogeneous symmetric polynomial of degree m is a polynomial in the power sums, it is a linear combination of terms like the one on the right-hand side of Equation (8.34), $\tau \in S_m$, and, therefore, a linear combination of the elements of $\{s_\pi : \pi \vdash m\}$. $\square$

Consider the inner product on $SC_m[x_1, x_2, \dots, x_m]$ with respect to which $\{s_\pi : \pi \vdash m\}$ is orthonormal. If $\tau \in S_m$ is fixed but arbitrary then, from Equation (8.34),

$$\chi_\pi(\tau) = \left(\prod_{i=1}^m P_i^{c_i(\tau)}, s_\pi \right).$$

Suppose $x_1, x_2, \dots, x_n$ and z are $n + 1$ independent indeterminates over $\mathbb{C}$. Let $E_m = E_m(x_1, x_2, \dots, x_n)$, $1 \leq m \leq n$. Define the "generating function" $g(z) = 1 + E_1z + E_2z^2 + E_3z^3 + \dots + E_nz^n$.

LEMMA 8.29 *The generating function for the elementary symmetric functions is given by the formula*

$$g(z) = (1 + x_1z)(1 + x_2z) \dots (1 + x_nz). \quad (8.35)$$

Proof To evaluate the right-hand side of Equation (8.35), choose one element from each of the n sets of brackets, and multiply them together. Do this in all 2^n possible ways and add the resulting products. The coefficient of z^m in this expression is the sum of the products of the x 's taken m at a time, namely, $E_m(x_1, x_2, \dots, x_n)$. $\square$

A similar result holds for the m -th homogeneous symmetric function

$$H_m(x_1, x_2, \dots, x_n) = \sum_{\alpha \in G_{m,n}} \prod_{i=1}^m x_{\alpha(i)}.$$

Define

$$h(z) = 1 + H_1z + H_2z^2 + H_3z^3 + \dots,$$

where $H_m = H_m(x_1, x_2, \dots, x_n)$. (Because $E_m(x_1, x_2, \dots, x_n) = 0$ for all $m > n$, $g(z)$ is a polynomial; $h(z)$, on the other hand, is an infinite series.)

LEMMA 8.30 *Suppose $x_1, x_2, \dots, x_n$ and z are independent indeterminates over $\mathbb{C}$. Then*

$$\begin{aligned} h(z) &= \frac{1}{(1 - x_1z)(1 - x_2z) \dots (1 - x_nz)} \\ &= \prod_{i=1}^n \frac{1}{(1 - x_i z)}. \end{aligned} \quad (8.36)$$

Proof The coefficient of z^m in the product

$$(1 + x_1 z + x_1^2 z^2 + \dots)(1 + x_2 z + x_2^2 z^2 + \dots) \dots (1 + x_n z + x_n^2 z^2 + \dots)$$

is a sum of terms of the form $x_1^{m_1} x_2^{m_2} \dots x_n^{m_n}$, one for each of the $C(m+n-1, m)$ nonnegative integer solutions to the equation $m_1 + m_2 + \dots + m_n = m$. $\square$

Comparing Equations (8.35) and (8.36), one sees that

$$g(-z)h(z) = 1,$$

in other words, setting $E_0 = H_0 = 1$,

$$\sum_{r=0}^m (-1)^r E_r H_{m-r} = 0, \quad m \geq 1. \quad (8.37)$$

EXAMPLE 8.31 Written out, Equations (8.37) become

$$H_1 - E_1 = 0,$$

$$H_2 - E_1 H_1 + E_2 = 0,$$

$$H_3 - E_1 H_2 + E_2 H_1 - E_3 = 0,$$

and so on.⁶ This means, for example, that

$$E_1 = H_1,$$

$$E_2 = H_1^2 - H_2,$$

and

$$E_3 = H_1^3 - 2H_1 H_2 + H_3.$$

Let's confirm the last of these equations when $n = 3$. Substituting $x_1 = a$, $x_2 = b$, and $x_3 = c$,

$$\begin{aligned} H_1(a, b, c)^3 &= (a + b + c)^3 \\ &= M_{[3]}(a, b, c) + 3M_{[2,1]}(a, b, c) + 6M_{[1^3]}(a, b, c), \\ H_1(a, b, c)H_2(a, b, c) &= (a + b + c)(a^2 + b^2 + c^2 + ab + ac + bc) \\ &= M_{[3]}(a, b, c) + 2M_{[2,1]}(a, b, c) + 3M_{[1^3]}(a, b, c), \end{aligned}$$

⁶Compare with Newton's Identities.

and

$$H_3(a, b, c) = M_{[3]}(a, b, c) + M_{[2,1]}(a, b, c) + M_{[1^3]}(a, b, c).$$

Therefore,

$$\begin{aligned} H_1^3 - 2H_1H_2 + H_3 &= M_{[1^3]} \\ &= E_3. \end{aligned}$$

□

THEOREM 8.32 Let $x_1, x_2, \dots, x_n$ be independent indeterminates over $\mathbb{C}$. Set $H_r = H_r(x_1, x_2, \dots, x_n)$ and $E_r = E_r(x_1, x_2, \dots, x_n)$ when $r \geq 1$, $H_0 = E_0 = 1$, and $H_r = E_r = 0$ when $r < 0$. Suppose H is a k -by- k matrix whose (i, j) -entry is H_{i-j} . Then H is invertible, and its inverse is the k -by- k matrix E whose (i, j) -entry is $(-1)^{i+j} E_{i-j}$.

Proof The (i, j) -entry of the product EH is

$$\sum_{t=j}^i (-1)^{i+t} E_{i-t} H_{t-j} = \sum_{r=0}^{i-j} (-1)^r E_r H_{i-j-r},$$

if $i \geq j$, and 0 otherwise. Because $E_0 = H_0 = 1$, the result follows from Equation (8.37). □

COROLLARY 8.33 Let $\pi = [\pi_1, \pi_2, \dots, \pi_r]$ be a partition of m . Suppose $x_1, x_2, \dots, x_n$ are independent indeterminates over $\mathbb{C}$. Set $H_r = H_r(x_1, x_2, \dots, x_n)$ and $E_r = E_r(x_1, x_2, \dots, x_n)$ when $r \geq 1$, $H_0 = E_0 = 1$, and $H_r = E_r = 0$ when $r < 0$. Then

$$\det (H_{\pi_i - i + j}) = \det (E_{\pi_i^* - i + j}). \quad (8.38)$$

The determinant on the left side of Equation (8.38) involves an r -by- r matrix and the one on the right a π_1 -by- π_1 matrix.

EXAMPLE 8.34 Suppose $m = 3$. If $\pi = [1^3]$, then $r = 3$, $\pi_1 = 1$, $\pi^* = [3]$, and Equation (8.38) becomes

$$\det \begin{pmatrix} H_1 & H_2 & H_3 \\ H_0 & H_1 & H_2 \\ 0 & H_0 & H_1 \end{pmatrix} = E_3,$$

that is, $E_3 = H_1^3 - 2H_1H_2 + H_3$, one of the identities from Example 8.31. □

Proof (of Corollary 8.33): Let $s = L(\pi^*) = \pi_1$ and set $k = r + s$. Let H be the k -by- k matrix from Theorem 8.32, whose (i, j) -entry is H_{i-j} . If

$$\alpha = (s + 1, s + 2, \dots, s + r)$$

and

$$\beta = (s - \pi_1 + 1, s - \pi_2 + 2, \dots, s - \pi_r + r),$$

then $\alpha, \beta \in Q_{r, r+s}$, and

$$H[\alpha|\beta]^t = (H_{\pi_i - i + j}).$$

By Theorem 8.32, $H = E^{-1}$. Because $\det(E) = 1$, Jacobi's Identity (Theorem 7.46) yields

$$\begin{aligned} \det(H[\alpha|\beta]^t) &= \det(H[\alpha|\beta]) \\ &= (-1)^{r(\alpha) + r(\beta)} \det(E(\beta|\alpha)), \end{aligned}$$

where $r(\beta) = \beta_1 + \beta_2 + \dots + \beta_r = rs - m + r(r + 1)/2 = r(\alpha) - m$, that is,

$$\det(H[\alpha|\beta]) = (-1)^m \det(E[\beta^c|\alpha^c]). \quad (8.39)$$

The sequence in $Q_{s, r+s}$ complementary to α is $\alpha^c = (1, 2, \dots, s)$. By Lemma 1.6, $\beta^c = (\pi_s^* + 1, \pi_{s-1}^* + 2, \dots, \pi_1^* + s)$. After multiplying row $s + 1 - i$ of $E(\beta|\alpha) = E[\beta^c|\alpha^c]$ by $(-1)^{\pi_i^*}$, $1 \leq i \leq s$, the resulting s -by- s matrix is

$$X = \begin{pmatrix} E_{\pi_1^*} & \dots & \mp E_{\pi_1^* - s + 2} & \pm E_{\pi_1^* - s + 1} \\ & \dots & & \\ \mp E_{\pi_2^* + s - 2} & \dots & E_{\pi_2^*} & -E_{\pi_2^* - 1} \\ \pm E_{\pi_1^* + s - 1} & \dots & -E_{\pi_1^* + 1} & E_{\pi_1^*} \end{pmatrix}.$$

Because $(-1)^{\pi_1^*}(-1)^{\pi_2^*} \dots (-1)^{\pi_s^*} = (-1)^m$, we see from Equation (8.39) that $\det(H[\alpha|\beta]) = \det(X)$. Because reversing the order of the rows and columns of X has no effect on its determinant, $\det(H[\alpha|\beta]) = \det(Y)$, where

$$Y = \begin{pmatrix} E_{\pi_1^*} & -E_{\pi_1^* + 1} & \dots & \pm E_{\pi_1^* + s - 1} \\ -E_{\pi_2^* - 1} & E_{\pi_2^*} & \dots & \mp E_{\pi_2^* + s - 2} \\ & & \dots & \\ \pm E_{\pi_1^* - s + 1} & \mp E_{\pi_1^* - s + 2} & \dots & E_{\pi_1^*} \end{pmatrix}.$$

Finally, if $D = \text{diag}(1, -1, 1, -1, \dots, (-1)^s)$, then $D^{-1}YD = (E_{\pi_i^* - i + j})$. $\square$

THEOREM 8.35 *The common value of the two determinants in Equation (8.38) is the Schur polynomial, $s_\pi(x_1, x_2, \dots, x_n)$.*⁷

EXAMPLE 8.36 Suppose $m = 3$. Let $\pi = [2, 1] = \pi^*$. Then

$$\begin{aligned}\det(H_{\pi_i - i + j}) &= \det \begin{pmatrix} H_2 & H_3 \\ H_0 & H_1 \end{pmatrix} \\ &= H_2 H_1 - H_3.\end{aligned}\tag{8.40}$$

Therefore, by Theorem 8.35,

$$\begin{aligned}s_\pi(x, y, z) &= (x^2 + y^2 + z^2 + xy + xz + yz)(x + y + z) - H_3(x, y, z) \\ &= M_{[3]}(x, y, z) + 2M_{[2,1]}(x, y, z) + 3M_{[1^3]}(x, y, z) \\ &\quad - M_{[3]}(x, y, z) - M_{[2,1]}(x, y, z) - M_{[1^3]}(x, y, z) \\ &= M_{[2,1]}(x, y, z) + 2M_{[1^3]}(x, y, z) \\ &= x^2y + x^2z + xy^2 + xz^2 + y^2z + yz^2 + 2xyz,\end{aligned}$$

exactly the value obtained in Equation (8.27). Similarly,

$$\begin{aligned}\det(E_{\pi_i^* - i + j}) &= \det \begin{pmatrix} E_2 & E_3 \\ E_0 & E_1 \end{pmatrix} \\ &= E_2 E_1 - E_3,\end{aligned}$$

so

$$\begin{aligned}s_\pi(x, y, z) &= (xy + xz + yz)(x + y + z) - xyz \\ &= x^2y + x^2z + xy^2 + xz^2 + y^2z + yz^2 + 2xyz,\end{aligned}$$

the same. □

⁷Theorem 8.35 is a variation of a result from [Jacobi (1841)]. (Also see [Trudi (1864)].) A recent combinatorial proof can be found in [Egécioğlu & Remmel (1990)]. The most complete source of information about Schur polynomials is [Macdonald (1995)]. Other useful references are [Doubilet, Fox & Rota (1980)], [Garsia & Remmel (1981)], [Littlewood (1958)], [Read (1968)], [Stanley (1971)], and [Thomas (1976a&b)]. The "Jacobi-Trudi matrix" corresponding to $\pi, \rho \vdash m$ is

$$H(\pi, \rho) = (H_{\pi_i - i - \rho_j - j});$$

the "skew Schur" polynomial is defined by $s_{\pi/\rho} = \det(H(\pi, \rho))$. (Because $s_\pi = s_{\pi/\emptyset}$ when ρ is empty, skew Schur polynomials are generalizations of Schur polynomials.) Conjectures of [Goulden & Jackson (1992a)] and [Stembridge (1992)] have led to some interesting work on immanants of Jacobi-Trudi matrices. (See, for example, [Greene (1992)], [Haiman (1993)], and [Stanley & Stembridge (1993)].)

DEFINITION 8.37 Let $x_1, x_2, \dots, x_n$ be independent indeterminates over $\mathbb{C}$. Set $H_r = H_r(x_1, x_2, \dots, x_n)$, $r \geq 1$. For each $\rho = [\rho_1, \rho_2, \dots, \rho_k] \vdash m$, define $H_\rho = H_{\rho_1}(x_1, x_2, \dots, x_n) \dots H_{\rho_k}(x_1, x_2, \dots, x_n)$ by

$$H_\rho = H_{\rho_1} H_{\rho_2} \dots H_{\rho_k}.$$

It follows from Theorem 8.35 that the coefficients $c_{\rho, \pi}$ in the expression

$$s_\pi = \sum_{\rho \vdash m} c_{\rho, \pi} H_\rho \quad (8.41)$$

can be evaluated by means of the determinant on the left-hand side of Equation (8.38). Among the interesting implications of this fact is a formula (Corollary 8.41) for χ_π in terms of π .

LEMMA 8.38 Let $x_1, x_2, \dots, x_n$ be independent indeterminates over $\mathbb{C}$. Suppose $\pi \vdash r$ and $\rho \vdash t$. Set $m = r + t$. Viewing S_r as the subgroup of S_m consisting of those permutations that fix $r + 1, r + 2, \dots, m$, and S_t as the subgroup consisting of those permutations that fix $1, 2, \dots, r$, then $G = S_r \times S_t$ is the Young subgroup $S_{[r, t]}$. Let ξ be the character of G defined by $\xi(\sigma, \tau) = \chi_\pi(\sigma) \chi_\rho(\tau)$.⁸ Then

$$\begin{aligned} s_\pi(x_1, x_2, \dots, x_n) s_\rho(x_1, x_2, \dots, x_n) \\ = \sum_{\nu \vdash m} (\xi, \chi_\nu)_{G} s_\nu(x_1, x_2, \dots, x_n). \end{aligned}$$

Proof From Equation (8.29),

$$\begin{aligned} s_\pi s_\rho &= \frac{1}{r!t!} \sum_{\sigma \in S_r} \sum_{\tau \in S_t} \chi_\pi(\sigma) \chi_\rho(\tau) \prod_{i=1}^r P_i^{c_i(\sigma)} \prod_{j=1}^t P_j^{c_j(\tau)} \\ &= \frac{1}{o(G)} \sum_{(\sigma, \tau) \in G} \xi(\sigma, \tau) \prod_{i=1}^m P_i^{c_i(\sigma, \tau)}. \end{aligned} \quad (8.42)$$

By the same token,

$$\begin{aligned} \sum_{\nu \vdash m} (\xi, \chi_\nu)_{G} s_\nu &= \frac{1}{o(G)} \sum_{(\sigma, \tau) \in G} \xi(\sigma, \tau) \sum_{\mu \in S_m} \left(\frac{1}{m!} \sum_{\nu \vdash m} \chi_\nu(\sigma, \tau) \chi_\nu(\mu) \right) \prod_{i=1}^m P_i^{c_i(\mu)}, \\ &= \frac{1}{o(G)} \sum_{(\sigma, \tau) \in G} \xi(\sigma, \tau) \prod_{i=1}^m P_i^{c_i(\sigma, \tau)}, \end{aligned} \quad (8.43)$$

by the Orthogonality Relations of the Second Kind. □

⁸See Exercise 29, Chapter 5.

THEOREM 8.39 If $\rho = [\rho_1, \rho_2, \dots, \rho_k] \vdash m$, then

$$H_\rho = \sum_{\nu \vdash m} K_{\nu, \rho} s_\nu,$$

where $K_{\nu, \rho}$ is a Kostka coefficient.

Proof From Example 8.23, $H_r = s_{[r]}$. Because $\chi_{[r]} = 1_r$, the principal character of S_r , it follows from Lemma 8.38 that

$$\begin{aligned} H_\rho &= \prod_{i=1}^k H_{\rho_i} \\ &= \prod_{i=1}^k s_{[\rho_i]} \\ &= \sum_{\nu \vdash m} (1_\rho, \chi_\nu) s_\rho s_\nu, \end{aligned}$$

where 1_ρ is the principal character of the Young subgroup S_ρ . The conclusion follows from the Frobenius Reciprocity Theorem and the definition of the Kostka coefficients. $\square$

COROLLARY 8.40 The matrix of coefficients $(c_{\rho, \pi})$ in Equation (8.41) is the inverse of the matrix of Kostka coefficients.

Proof Because (Exercise 23) $\{H_\rho(x_1, x_2, \dots, x_m) : \rho \vdash m\}$ is a basis of the space $SC_m[x_1, x_2, \dots, x_m]$ and (Theorem 8.28) $\{s_\pi : \pi \vdash m\}$ is another, the result is an immediate consequence of Theorem 8.39. $\square$

Corollary 8.40 leads to a family of explicit formulas for irreducible characters of symmetric groups. However, the description of these formulas requires a brief "suspension of disbelief". For the present application, let us interpret the formal "product"

$$[k_1][k_2] \dots [k_r]$$

to be the induced character $1_\rho^{S_m}$, where ρ is the partition whose parts are the positive integers $k_1, k_2, \dots, k_r$. Define $[0] = 1$, and interpret $[k]$ to be 0 whenever $k < 0$.

COROLLARY 8.41⁹ Suppose $\pi = [\pi_1, \pi_2, \dots, \pi_r] \vdash m$. Let A_π be the r -by- r matrix whose (i, j) -entry is the symbol $[\pi_i - i + j]$, then $\det(A_\pi) = \chi_\pi$, where χ_π is the irreducible character of S_m corresponding to π .

⁹Corollary 8.41 comes from [Frobenius (1900)].

Proof From the definition of the Kostka coefficients,

$$1_{\rho}^{S_m} = \sum_{\pi \vdash m} K_{\pi, \rho} \chi_{\pi}, \quad \rho \vdash m.$$

Therefore, from Corollary 8.40,

$$\chi_{\pi} = \sum_{\rho \vdash m} c_{\rho, \pi} 1_{\rho}^{S_m}, \quad \pi \vdash m. \quad (8.44)$$

The mechanical procedure described in the statement of the corollary emerges from the computations reflected by the identity

$$\sum_{\rho \vdash m} c_{\rho, \pi} H_{\rho} = \det (H_{\pi_i - i + j}).$$

□

EXAMPLE 8.42 Suppose $m = 3$. Let $\pi = [m - 1, 1] \vdash m$. Then, in the notation of Corollary 8.41,

$$\begin{aligned} A_{\pi} &= \begin{pmatrix} [\pi_1] & [\pi_1 + 1] \\ [\pi_2 - 1] & [\pi_2] \end{pmatrix} \\ &= \begin{pmatrix} [m - 1] & [m] \\ 1 & [1] \end{pmatrix}, \end{aligned}$$

and $\det(A_{\pi}) = [m - 1][1] - [m]$. Therefore,

$$\chi_{\pi} = (1_{[m-1,1]})^{S_m} - (1_{[m]})^{S_m}. \quad (8.45)$$

From the definition of induced characters,

$$1_{\rho}^{S_m}(\tau) = \frac{1}{o(S_{\rho})} \sum_{\sigma \in S_m} 1^{\#}(\sigma^{-1} \tau \sigma), \quad (8.46)$$

where $1^{\#}$ is the characteristic function of S_{ρ} in S_m , that is, $1^{\#}(\mu) = 1$, if $\mu \in S_{\rho}$, and 0 otherwise. When $\rho = [m - 1, 1]$, $\sigma^{-1} \tau \sigma \in S_{\rho} = S_{m-1} \times S_1$, if and only if $\sigma(m)$ is a fixed point of τ . Given a $\tau \in S_m$ having $c_1(\tau)$ fixed points, there are $c_1(\tau) \times (m - 1)!$ permutations $\sigma \in S_m$ that satisfy this criterion. Because $o(S_{[m-1,1]}) = (m - 1)!$, it follows from Equation (8.46) that

$$(1_{[m-1,1]})^{S_m}(\tau) = c_1(\tau), \quad \tau \in S_m. \quad (8.47)$$

Because $(1_{[m]})^{S_m} = 1_{[m]}$ is the principal character of S_m , Equations (8.45) and (8.47) confirm the formula

$$\chi_{[m-1,1]} = c_1 - 1, \quad m \geq 2. \quad (8.48)$$

□

EXAMPLE 8.43 Equation (8.48) may be viewed as the first of Frobenius's formulas for irreducible characters of symmetric groups. Other examples include

$$\chi_{[m-2,2]} = c_1(c_1 - 3)/2 + c_2, \quad m \geq 4; \quad (8.49)$$

$$\chi_{[m-2,1^2]} = (c_1 - 1)(c_1 - 2)/2 - c_2, \quad m \geq 3; \quad (8.50)$$

$$\chi_{[m-3,3]} = c_1(c_1 - 1)(c_1 - 5)/6 + (c_1 - 1)c_2 + c_3, \quad m \geq 6; \quad (8.51)$$

$$\chi_{[m-3,2,1]} = c_1(c_1 - 2)(c_1 - 4)/3 - c_3, \quad m \geq 5; \quad (8.52)$$

and

$$\chi_{[m-3,1^3]} = (c_1 - 1)(c_1 - 2)(c_1 - 3)/6 - (c_1 - 1)c_2 + c_3, \quad m \geq 4. \quad (8.53)$$

Let's confirm Equation (8.49). If $\pi = [m - 2, 2]$ then, in the notation of Corollary 8.41,

$$A_\pi = \begin{pmatrix} [m-2] & [m-1] \\ [1] & [2] \end{pmatrix}$$

and $\det(A_\pi) = [m-2][2] - [m-1][1]$. Therefore,

$$\begin{aligned} \chi_\pi &= (1_{[m-2,2]})^{S_m} - (1_{[m-1,1]})^{S_m} \\ &= (1_{[m-2,2]})^{S_m} - c_1, \end{aligned}$$

from Equation (8.47).

Given $\tau \in S_m$, there are two ways for $\sigma^{-1}\tau\sigma$ to be an element of $S_\rho = S_{m-2} \times S_2$. It could happen that $\sigma(m)$ and $\sigma(m-1)$ are fixed points of τ while $1, 2, \dots, m-2$ are permuted arbitrarily among them selves. There are

$$c_1(\tau)(c_1(\tau) - 1)(m-2)!$$

σ 's of this type. Alternatively, there are $2c_2(\tau)$ ways in which $\sigma(m-1)$ and $\sigma(m)$ can occupy the same 2-cycle of τ . For each of them, the remaining integers can be permuted in $(m-2)!$ ways. There are $2c_2(\tau)(m-2)!$ permutations $\sigma \in S_m$ fitting this description. Since $o(S_\rho) = (m-2)! \times 2$,

$$(1_{[m-2,2]})^{S_m}(\tau) = c_1(\tau)(c_1(\tau) - 1)/2 + c_2(\tau).$$

Putting it all together,

$$\begin{aligned} \chi_{[m-2,2]}(\tau) &= c_1(\tau)(c_1(\tau) - 1)/2 + c_2(\tau) - c_1(\tau) \\ &= c_1(\tau)(c_1(\tau) - 3)/2 + c_2(\tau). \end{aligned}$$

□

The techniques leading to Corollary 8.41 and the Frobenius formulas can also be used to shed new light on immanants.

THEOREM 8.44 Suppose $\pi \vdash m$. If $A \in \mathbf{C}_{m,m}$, then

$$d_\pi(A) = \sum_{\rho \vdash m} c_{\rho,\pi} \sum_{\rho}^{\rho} \text{per}(A[N_1]) \text{per}(A[N_2]) \dots \text{per}(A[N_k]),$$

where, for each $\rho = [\rho_1, \rho_2, \dots, \rho_k] \vdash m$, the sum $\sum_{\rho}^{\rho}$ is over all ordered sequences $(N_1, N_2, \dots, N_k)$ such that

$$\{1, 2, \dots, m\} = N_1 \cup N_2 \cup \dots \cup N_k,$$

and $o(N_i) = \rho_i$; $A[N_i]$ is the principal submatrix of A whose rows and columns are indexed by the elements of N_i ; and the $c_{\rho,\pi}$ are inverse Kostka coefficients.

Proof If $\rho = [\rho_1, \rho_2, \dots, \rho_k] \vdash m$ then

$$\begin{aligned} \sum_{\tau \in S_m} 1_{\rho}^{S_m}(\tau) \prod_{i=1}^m a_{i\tau(i)} &= \frac{1}{o(S_{\rho})} \sum_{\tau, \sigma \in S_m} 1^{\#}(\sigma^{-1}\tau\sigma) \prod_{i=1}^m a_{i\tau(i)} \\ &= \frac{1}{o(S_{\rho})} \sum_{\sigma \in S_m} \sum_{\tau \in S_{\rho}} \prod_{i=1}^m a_{\sigma(i), \sigma\tau(i)} \\ &= \frac{1}{o(S_{\rho})} \sum_{\sigma \in S_m} \sum_{\tau \in S_{\rho}} \prod_{i=1}^m (Q(\sigma^{-1})AQ(\sigma))_{i\tau(i)} \\ &= \frac{1}{o(S_{\rho})} \sum_{\sigma \in S_m} \prod_{i=1}^k \text{per}((Q(\sigma^{-1})AQ(\sigma))[\Gamma^* \rho_i^*]), \end{aligned} \quad (8.54)$$

where $Q(\sigma) = (\delta_{i, \sigma(j)})$ and $(Q(\sigma^{-1})AQ(\sigma))[\Gamma^* \rho_i^*]$ is the principal submatrix of $Q(\sigma^{-1})AQ(\sigma)$ lying in rows and columns $(\rho_1 + \rho_2 + \dots + \rho_{i-1}) + 1$ through $(\rho_1 + \rho_2 + \dots + \rho_{i-1}) + \rho_i$. As σ ranges over S_m , each product, $\prod \text{per}(A[N_i])$, occurs among the terms

$$\prod \text{per}((Q(\sigma^{-1})AQ(\sigma))[\Gamma^* \rho_i^*])$$

with multiplicity $o(S_{\rho}) = \prod \rho_i!$. Therefore, Equation (8.54) can be rewritten as

$$\sum_{\tau \in S_m} 1_{\rho}^{S_m}(\tau) \prod_{i=1}^m a_{i\tau(i)} = \sum_{\rho}^{\rho} \prod_{i=1}^k \text{per}(A[N_i]).$$

Multiply both sides of this identity by $c_{\rho,\pi}$ and sum on $\rho \vdash m$ to obtain

$$\sum_{\tau \in S_m} \left(\sum_{\rho \vdash m} c_{\rho,\pi} 1_{\rho}^{S_m}(\tau) \right) \prod_{i=1}^m a_{i\tau(i)} = \sum_{\rho \vdash m} c_{\rho,\pi} \sum_{\rho}^{\rho} \prod_{i=1}^k \text{per}(A[N_i]),$$

which, in view of Equation (8.44), is what was to have been proved. $\square$

EXAMPLE 8.45 Let $\pi = [r, s]$ where $r \geq s$ and $r + s = m$. Then

$$\begin{aligned} s_\pi &= \det \begin{pmatrix} H_r & H_{r+1} \\ H_{s-1} & H_s \end{pmatrix} \\ &= H_r H_s - H_{r+1} H_{s-1}. \end{aligned} \quad (8.55)$$

Because (Exercise 23) $\{H_\rho : \rho \vdash m\}$ is a basis of $SC_m[x_1, x_2, \dots, x_m]$, the inverse Kostka coefficients

$$c_{\rho, [r, s]} = \begin{cases} 1, & \text{if } \rho = [r, s] \\ -1, & \text{if } \rho = [r+1, s-1] \\ 0, & \text{otherwise} \end{cases}$$

are uniquely determined by Equations (8.41) and (8.55). Thus, from Theorem 8.44,

$$d_{[r, s]}(A) = \sum_{i=1}^{[r, s]} \text{per}(A[N_1]) \text{per}(A[N_2]) - \sum_{i=1}^{[r+1, s-1]} \text{per}(A[N_1]) \text{per}(A[N_2]),$$

for all $A \in \mathbb{C}_{m, m}$. In particular, setting $s = 1$,

$$d_{[m-1, 1]}(A) = \sum_{i=1}^m a_{ii} \text{per}(A(i|i)) - \text{per}(A). \quad (8.56)$$

□

As illustrated by Example 8.45, Theorem 8.44 is the foundation for a procedure to generate identities for immanants: Use Theorem 8.35 to express s_π as a sum of products of homogeneous symmetric functions. When (each) H_ρ is replaced with the sum of all possible products of permanents of complementary ρ_i -by- ρ_i principal submatrices of A , and s_π is replaced by $d_\pi(A)$, the result is an identity expressing $d_\pi(A)$ in terms of permanents. This procedure illustrates a general principle due to Littlewood.¹⁰

LITTLEWOOD'S THEOREM 8.46 *To any homogeneous polynomial relation of total degree m among Schur polynomials, there is an analogous relation for immanants obtained by replacing the Schur polynomials with the corresponding immanants of complementary principal submatrices and summing over all sets of complementary principal submatrices.*

¹⁰It appears to have been D.E. Littlewood who coined the term "immanant". Theorem 8.46 is stated without proof in [Littlewood (1958)]. A proof can be found, e.g., in [Goulden & Jackson (1992b)].

EXAMPLE 8.47¹¹ Suppose $\pi = [r, 1^{m-r}]$. Then, from Theorem 8.35,

$$s_{[r, 1^{m-r}]} = \det \begin{pmatrix} H_r & H_{r+1} & \dots & H_m \\ 1 & H_1 & \dots & H_{m-r} \\ 0 & 1 & \dots & H_{m-r-1} \\ & & \ddots & \\ 0 & 0 & \dots & H_1 \end{pmatrix}.$$

Expanding the determinant down the first column, we find that

$$\begin{aligned} s_{[r, 1^{m-r}]} &= H_r \times s_{[1^{m-r}]} - s_{[r+1, 1^{m-r-1}]} \\ &= s_{[r]} s_{[1^{m-r}]} - s_{[r+1, 1^{m-r-1}]}. \end{aligned}$$

Therefore, from Littlewood's Theorem, the single-hook immanant $d_r = d_{[r, 1^{m-r}]}$ satisfies the identity

$$d_r(A) = \sum^{[r, m-r]} d_{[r]}(A[N_1]) d_{[1^{m-r}]}(A[N_2]) - d_{r+1}(A),$$

or, equivalently,

$$d_{r+1}(A) = \sum^{[r, m-r]} \text{per}(A[N_1]) \det(A[N_2]) - d_r(A), \quad A \in \mathbf{C}_{m,m}. \quad (8.57)$$

Starting with $d_1(A) = \det(A)$, Equation (8.57) yields, inductively,

$$d_2(A) = \sum_{i=1}^m a_{ii} \det(A(i|i)) - \det(A);$$

$$d_3(A) = \sum_{i=1}^m \sum_{j>i}^m \text{per}(A[i, j|i, j]) \det(A(i, j|i, j)) - d_2(A);$$

$$d_4(A) = \sum_{i=1}^m \sum_{j>i}^m \sum_{k>j}^m \text{per}(A[i, j, k|i, j, k]) \det(A(i, j, k|i, j, k)) - d_3(A);$$

and so on.

These formulas seem to confirm the increasing difficulty of computing $d_r(A)$ as one goes from $d_1(A) = \det(A)$ to $d_m(A) = \text{per}(A)$.¹² $\square$

¹¹This example is taken from [Merris & Watkins (1985)].

¹²To be more quantitative, if the time to compute an m -by- m determinant is on the order of m^3 , then the time to compute d_r , using these formulas, is on the order of m^{r+2} . Interesting work on the computational complexity of immanants appears in [Barvinok], [Hartman (1985)], and [Valiant (1979)]. Related work can be found in [Goulden & Jackson (1992b)] and [Lloyd (1983)].

EXAMPLE 8.48 Suppose $\chi = \chi_\rho \in I(S_m)$. Then [Merris (1982)]

$$d_\chi(A) \det(A^{-1}) = d_\chi(A^{-1}) \det(A), \quad A \in GL(m, \mathbb{C}),$$

if and only if ρ is a partition of the form $[2^p, 1^q]$, where $2p + q = m$. If $\rho = [2^p, 1^q]$, then $\rho^* = [p + q, p]$ is among the partitions considered in Example 8.45. From Theorem 8.35,

$$\begin{aligned} s_\rho &= \det \begin{pmatrix} E_{p+q} & E_{p+q+1} \\ E_{p-1} & E_p \end{pmatrix} \\ &= E_{p+q} E_p - E_{p+q+1} E_{p-1}. \end{aligned}$$

Applying Littlewood's Theorem to this identity produces

$$\begin{aligned} d_\rho(A) &= \sum_{[p+q, p]} \det(A[N_1]) \det(A[N_2]) \\ &\quad - \sum_{[p+q+1, p-1]} \det(A[N_1]) \det(A[N_2]). \end{aligned}$$

□

THEOREM 8.49 (*Giambelli's Identity*¹³). Suppose $\pi \vdash m$. Let $r = f(\pi) = o(\{i : \pi_i \geq i\})$. Then the Schur polynomial, s_π , is the determinant of the r -by- r matrix whose (i, j) -entry is the Schur polynomial corresponding to the single-hook partition $[\pi_i - i + 1, 1^{\pi_j - j}]$.

EXAMPLE 8.50 If $\pi = [4, 3, 1]$, then $\pi^* = [3, 2^2, 1]$ and the length of the main diagonal of $F(\pi)$ is $r = f(\pi) = 2$. By Giambelli's Identity,

$$s_\pi = \det \begin{pmatrix} s_{[4, 1^2]} & s_{[4]} \\ s_{[2, 1^2]} & s_{[2]} \end{pmatrix},$$

that is, $s_{[4, 3, 1]} = s_{[4, 1^2]} s_{[2]} - s_{[4]} s_{[2, 1^2]}$. Combining this identity with Littlewood's Theorem yields

$$\begin{aligned} d_{[4, 3, 1]}(A) &= \sum_{\alpha \in Q_{4,1}} d_4(A[\alpha|\alpha]) \operatorname{per}(A(\alpha|\alpha)) \\ &\quad - \sum_{\beta \in Q_{4,1}} \operatorname{per}(A[\beta|\beta]) d_2(A(\beta|\beta)), \end{aligned}$$

for any 8-by-8 matrix A .

□

¹³Theorem 8.49 first appeared in [Giambelli (1903)]. A combinatorial proof can be found in [Egecioglu & Remmel (1988)].

Application to Graphs

Let $H = (V, E)$ be a graph with vertex set $V = \{1, 2, \dots, n\}$ and edge set $E = E(H)$. One way to describe H is by means of a coloring of the $m = C(n, 2)$ edges of the complete graph K_n . An edge of K_n is colored 1 if it is an edge of H and 2 if it is not. Thus, there is a one-to-one correspondence between the different graphs with vertex set V and the set $\Gamma_{m,2}$ of all functions $\alpha: \{1, 2, \dots, m\} \rightarrow \{1, 2\}$.

Suppose W is a vector space of dimension 2. Let $\{e_1, e_2\}$ be a basis of W . Then $\{e_\alpha^\otimes: \alpha \in \Gamma_{m,2}\}$ is a basis of $W^{\otimes m}$. Consequently, $W^{\otimes m}$ must be isomorphic to the free vector space generated by the different graphs on n vertices. Moreover, $\alpha \equiv \beta \pmod{S_n^{(2)}}$, if and only if there is a permutation $\sigma \in S_n^{(2)}$ such that $\beta = \alpha\sigma$, if and only if the graph corresponding to α is isomorphic to the graph corresponding to β . Evidently, there is a one-to-one correspondence between the nonisomorphic graphs on n vertices and the elements of Δ . Therefore, the free vector space spanned by the nonisomorphic graphs on n vertices must be isomorphic to $W_1(S_n^{(2)})$. At the very least, this means the number of nonisomorphic graphs on n vertices is equal to

$$\dim(W_1(S_n^{(2)})) = \sum_{\pi \vdash m} \chi_\pi(e)(1, \chi_\pi)_{S_n^{(2)}} f_\pi(2)/m!. \quad (8.58)$$

Because $f_\pi(2) = 0$ when $L(\pi) > 2$, we may restrict the sum in Equation (8.58) to those partitions having at most two parts. If $\pi = [p+q, p] \vdash m$, then $f_\pi(2) = (p+q+1)!p!$ and $\chi_\pi(e) = m!/\prod h_{ij}$, where the hook length product is $\prod h_{ij} = p!(p+q+1)!(q+1)$. Therefore, $\chi_\pi(e)f_\pi(2)/m! = q+1 = m-2p+1$. This proves the following.

THEOREM 8.51¹⁴ Suppose n is a positive integer. Let $m = C(n, 2)$ and denote by ξ_p the irreducible character of S_m corresponding to the partition $[p+q, p]$, $2p+q = m$. Then the number of nonisomorphic graphs on n vertices is given by the formula

$$\sum_{p=0}^{\lfloor m/2 \rfloor} (m-2p+1)(1, \xi_p)_{S_n^{(2)}}. \quad (8.59)$$

The missing ingredient in Theorem 8.51 is an analog of Young's Rule for computing the number of occurrences of the principal character in the restriction of $\xi_p = \chi_{[m-p, p]}$ to $S_n^{(2)}$.

EXAMPLE 8.52 Suppose $n = 4$. Then $m = C(4, 2) = 6$. From Examples 8.42 and 8.43, the characters $\xi_p = \chi_{[6-p, p]}$, $0 \leq p \leq 3$, are given by the formulas

¹⁴Theorem 8.51 first appeared in [Merris & Watkins (1983)].

$$\xi_0 = 1,$$

$$\xi_1 = c_1 - 1,$$

$$\xi_2 = c_1(c_1 - 3)/2 + c_2,$$

and

$$\xi_3 = c_1(c_1 - 1)(c_1 - 5)/6 + (c_1 - 1)c_2 + c_3.$$

The 24 permutations $\tilde{\sigma} \in S_4^{(2)}$ are tabulated in Figure 6.7. A brute-force computation based on this information yields $(1, \xi_p)_{S_4^{(2)}} = 1$, $p \neq 1$, and $(1, \xi_1)_{S_4^{(2)}} = 0$. From Equation (8.59), the number of non-isomorphic graphs on 4 vertices is

$$7 \times 1 + 5 \times 0 + 3 \times 1 + 1 \times 1 = 11.$$

□

Exercises

1. Prove Lemma 8.3.
2. Let χ be an irreducible character of $G = S_n$. Suppose $d_\chi(AB) = d_\chi(A)d_\chi(B)$, for all $A, B \in C_{n,n}$.
 - a. Prove that $\chi = \varepsilon$. (Hint: Theorem 8.2.)
 - b. What can be said if $G \neq S_n$?
3. Show that Equation (8.10) defines a function. (Hint: Theorems 5.15 and 5.16.)
4. Finish the proof of Theorem 8.9 by establishing
 - a. part (ii).
 - b. part (iii).
5. Use Equations (8.14), (8.15), and a direct computation to
 - a. show that $T_1(S_3, A)T_2(S_3, A) = 0$, thus confirming part of Theorem 8.9(i).
 - b. show that $T_1(S_3, A)^2 = T_1(S_3, A)$, thus confirming part of Theorem 8.9(i).
 - c. show that $T_1(S_3, A) + T_2(S_3, A) = T(S_3, \chi)$, thus confirming Theorem 8.9(ii).
 - d. show that $T_1(S_3, A)$ is not hermitian. Explain why this does not contradict Theorem 8.9(iii).
6. Let χ be an irreducible character of the subgroup G of S_m . Suppose $\sigma \rightarrow A(\sigma) = (a_{ij}(\sigma))$ is a representation of G affording χ . Let V be a vector space. Prove that $\dim(V_A^i(G)) = (\bar{\chi}, \eta)_G$, where η is the character of G afforded by the restriction of P to $V_\chi(G)$.
7. Prove that $A \rightarrow P_m(A)$ is an irreducible representation of $GL(n, \mathbb{C})$ for all $n \geq 1$. (Hint: Theorem 8.13.)
8. Prove that $A \rightarrow C_m(A)$ is an irreducible representation of $GL(n, \mathbb{C})$ if and only if $1 \leq m \leq n$. (Hint: Theorem 8.13.)
9. Let ξ be a linear character of the subgroup G of S_m . Suppose $m \leq n$. Prove that B_ξ^G is an irreducible representation of $GL(n, \mathbb{C})$ if and only if ξ^{S_m} , the character of S_m induced by ξ , is irreducible.
10. Let $G = D_4 \subset S_4$. Then the irreducible characters of G are given in Example 4.42. Suppose ξ is either χ_3 or χ_4 .
 - a. Prove that $T \rightarrow B_\xi^G(T)$, $T \in GL(V)$, is irreducible for any vector space V of dimension $n \geq 4$.
 - b. What about $n = 3$?
11. Suppose V is a vector space of dimension n . Let η be the character of the representation $\sigma \rightarrow P(\sigma)$, $\sigma \in S_m$, where $P(\sigma) \in L(V^{\otimes m}, V^{\otimes m})$ is the operator defined by

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

Show that

$$\eta = \sum_{\pi \vdash m} (\chi_{\pi}(e) f_{\pi}(n)/m!) \chi_{\pi}.$$

12. Suppose $\pi \vdash m$. Let $s_{\pi}(x_1, x_2, \dots, x_n)$ be the corresponding Schur polynomial.
 - a. Show that $s_{\pi}(1, 1, \dots, 1) = \chi_{\pi}(e) f_{\pi}(n)/m!$.
 - b. Use part (a) to prove that $s_{\pi}(x_1, x_2, \dots, x_n) = 0$ only if $n < L(\pi)$.
13. Prove Theorem 8.26.
14. Mimic Example 8.27 for the partition $\nu = [2, 2]$.
15. Express $H_3(x_1, x_2, \dots, x_n)$ as an explicit polynomial in the elementary symmetric functions. (Hint: Example 8.31.)
16. In the manner of Example 8.31, express $E_4(x_1, x_2, \dots, x_n)$ as an explicit polynomial in the homogeneous symmetric functions.
17. Prove that $E_m(x_1, x_2, \dots, x_n)$ can be expressed as a polynomial in the homogeneous symmetric functions. (Hint: Argue inductively from Equation (8.37).)
18. Prove that any polynomial, symmetric in the variables $x_1, x_2, \dots, x_n$, can be expressed as a polynomial in the homogeneous symmetric functions $H_m(x_1, x_2, \dots, x_n)$. (Hint: Exercise 17.)
19. Use the Murnaghan-Nakayama Rule to confirm Equation (8.49).
20. Confirm that Equations (8.49)–(8.53) are valid for $\tau = e$, that is, confirm that the character degrees afforded by the Frobenius formulas are consistent with the Frame-Robinson-Thrall hook length formula.
21. It follows from Theorem 8.28 that the monomial symmetric functions $M_{\pi}(x_1, x_2, \dots, x_n)$, $\pi \vdash m$, are linear combinations of Schur Polynomials. Express each of $M_{\pi}(x_1, x_2, \dots, x_n)$, $\pi \vdash 4$, as a linear combination of $s_{\rho}(x_1, x_2, \dots, x_n)$, $\rho \vdash 4$. (Hint: Example 4.57.)
22. Let $\pi = [3, 1] \vdash 4$. Show, by a direct computation, that
 - a. $\det \begin{pmatrix} H_3 & H_4 \\ H_0 & H_1 \end{pmatrix} = M_{[3,1]} + M_{[2^2]} + 2M_{[2,1^2]} + 3M_{[1^4]}.$
 - b. $\det \begin{pmatrix} E_2 & E_3 & E_4 \\ E_0 & E_1 & E_2 \\ 0 & E_0 & E_1 \end{pmatrix} = M_{[3,1]} + M_{[2^2]} + 2M_{[2,1^2]} + 3M_{[1^4]}.$
 - c. $s_{[3,1]}(a, b, c, d) = (a^3b + a^3c + \cdots + cd^3) + (a^2b^2 + a^2c^2 + \cdots + c^2d^2) + 2(a^2bc + a^2bd + \cdots + bcd^2) + 3abcd.$

23. Prove that $\{H_\pi(x_1, x_2, \dots, x_m) : \pi \vdash m\}$ is a basis of the vector space $SC_m[x_1, x_2, \dots, x_m]$.
24. Prove that $\bar{d}_{[m-1,1]}(A) \leq \text{per}(A)$, $A \in \mathcal{H}_m$. (Hint: Equations (7.12) and (8.56).)
25. It follows from Exercise 41, Chapter 6, that the substitution $y_i = P_i(x_1, x_2, \dots, x_n)$ in the cycle index polynomial $Z_{S_m}(y_1, y_2, \dots, y_m)$ produces the m -th homogeneous symmetric function $H_m(x_1, x_2, \dots, x_n)$. Use this fact to establish that
- $H_1 = P_1$.
 - $H_2 = (P_1^2 + P_2)/2$.
 - $H_3 = (P_1^3 + 3P_1P_2 + 2P_3)/6$.
 - $H_4 = (P_1^4 + 6P_1^2P_2 + 8P_1P_3 + 6P_4 + 3P_2^2)/24$.
26. Show that the Schur polynomial $s_{[3,1]} = (3P_1^4 + 6P_1^2P_2 - 6P_4 - 3P_2^2)/24$. (Hint: Example 8.45 and Exercise 25.)
27. For any $A \in C_{m,m}$, prove that

$$\begin{aligned} d_{[m-2,1^2]}(A) &= \sum_{i=1}^m \sum_{j \neq i} a_{ii} a_{jj} \text{per}(A(i, j|i, j)) \\ &\quad - \frac{1}{2} \sum_{i=1}^m \sum_{j \neq i} \text{per}(A[i, j|i, j]) \text{per}(A(i, j|i, j)) \\ &\quad - \sum_{i=1}^m a_{ii} \text{per}(A(i|i)) + \text{per}(A). \end{aligned}$$

28. For any $A \in C_{m,m}$, prove that the single-hook immanant

$$d_{m-2}(A) = \sum_{i=1}^m \sum_{j \neq i} \text{per}(A(i, j|i, j)) \det(A[i, j|i, j]) - d_{m-1}(A)$$

- directly from Exercise 27.
 - as in Example 8.47.
29. Prove that $(m - k - 1)\bar{d}_k(A) \leq m \text{per}(A)$, $A \in \mathcal{H}_m$.
30. Show that $s_{[6,4^2,2^2]} = s_{[6,1^4]}s_{[3,1^3]}s_{[2]} + s_{[6,1^3]}s_{[3]}s_{[2,1^4]} + s_{[6]}s_{[3,1^4]}s_{[2,1^3]} - s_{[6,1^3]}s_{[3,1^4]}s_{[2]} - s_{[6]}s_{[3,1^3]}s_{[2,1^4]} - s_{[6,1^4]}s_{[3]}s_{[2,1^3]}$.
31. If $\chi = \chi_{[2^p, 1^q]}$, where $2p + q = m$, then (Example 8.48)

$$d_X(A) \det(A^{-1}) = d_X(A^{-1}) \det(A),$$

for all invertible $A \in \mathbb{C}_{m,m}$. Confirm this equation for $m = 4$,

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & -1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}, \quad \text{and}$$

- a. $p = 1$.
 - b. $p = 2$.
32. Exhibit a 3-by-3 invertible matrix A such that
- $$\text{per}(A) \det(A^{-1}) \neq \text{per}(A^{-1}) \det(A).$$
33. Suppose $A \in \mathcal{H}_m$ does not have a zero row. Prove that the single-hook immanant $d_k(A) = 0$ if and only if $\text{rank}(A) \leq m - k$.
34. Let χ be an irreducible character of the subgroup G of S_m .
- a. If $d_X(A) \neq 0$ for all $A \in GL(m, \mathbb{C})$, prove that $d_X = \det$.
 - b. If $d_X(A) = 0$ for all singular $A \in \mathbb{C}_{m,m}$, prove that $d_X = \det$.
35. In the manner of Example 8.43, confirm
- a. Equation (8.50).
 - b. Equation (8.52).
36. Use the Murnaghan-Nakayama Rule to confirm
- a. Equation (8.50).
 - b. Equation (8.52).
37. Use Equation (8.49) to confirm the values given in Figure 4.5 for
- a. $\chi_{[3,2]}$.
 - b. $\chi_{[2^2,1]}$. (Hint: Theorem 4.47.)
38. Use Equation (8.50) to confirm the values given in Figure 4.5 for $\chi_{[3,1^2]}$.
39. Confirm the formula $s_{[2^2,1]} = M_{[2^2,1]} + 2M_{[2,1^3]} + 5M_{[1^5]}$ using
- a. Equation (8.29) and your answer to Exercise 37b.
 - b. the right-hand determinant in Equation (8.38). (Hint: Theorem 8.35.)
 - c. Giambelli's Identity.
40. Compute the Kostka coefficients $K_{[2^2,1],\rho}$, $\rho \vdash 5$. (Hint: Exercise 39.)
41. Show that the Kosta coefficient $K_{[3,1^2],[2,1^3]} = 3$.

Index of Notation

1_G	principal character of G	85
1_π	principal character of Young subgroup S_π	99
$(f, g)_G$	inner product of $f, g \in \mathbb{C}G$	86
$(\chi, \xi)_n$	$(\chi, \xi)_{S_n}$	99
A^r	r -th power of $A \geq 0$	44
$A^\star$	conjugate transpose of (matrix) A	28
A^t	transpose of (matrix) A	113
A_n	alternating group of degree n	57
$\text{adj}(A)$	classical adjoint or adjugate	226
$A(G)$	adjacency matrix of graph G	44
$A(G)$	group $\{A(\sigma) : \sigma \in G\}$	244
$A(i j)$	delete row i and column j	174
A/P	Schur complement of P in A	50
$A \geq 0$	A is positive semidefinite hermitian	36
$A \geq B$	$A - B \geq 0$	40
$A \cdot B$	Hadamard product of A and B	138
$A_1 \otimes A_2 \otimes \cdots \otimes A_m$	matrix Kronecker product	137
$A^{\otimes m}$	$A \otimes A \otimes \cdots \otimes A$ (m -times)	144
$A^{[m]}$	variation on $P_m(A)$	242
$(a) \succ (b)$	(a) majorizes (b)	3
$A(\alpha \beta)$	delete rows α and columns β	175
$A[\alpha \beta]$	keep rows α and columns β	175
α^c	sequence complementary to $\alpha \in Q_{r,m}$	174
$\mathcal{B}^\#$	induced basis of $V_X(G)$	234
$B^\star$	induced bases of $V_X(G)$	238
$C_m(A)$	m -th compound of (matrix) A	236
$C(n, r)$	binomial coefficient n -choose- r	8, 70
$C_G(\sigma)$	conjugacy class of σ in G	89

C_π	conjugacy class in S_n of cycle type π	93
$C_\sigma(x)$	cycle of σ containing x	53
$c_t(\sigma)$	number of cycles of length t in σ	57
$c(\sigma)$	total number of cycles in σ	57
$\mathbb{C}$	complex numbers	12
$\mathbb{C}G$	set of functions $f: G \rightarrow \mathbb{C}$	86, 117
$\mathbb{C}_{m,n}$	complex m -by- n matrices	28
$\mathbb{C}^n$	n -tuples of complex numbers	28
$\mathbb{C}[x_1, x_2, \dots, x_n]$	polynomials in $x_1, x_2, \dots, x_n$	12
$\mathbb{C}_m[x_1, x_2, \dots, x_n]$	homogeneous polynomials of degree m	180
$\mathcal{C}_m$	$\{c: S_m \rightarrow \mathbb{C} d_c(A) \geq 0, A \in \mathcal{H}_m\}$	215
$\mathcal{C}_m^+$	$\{c: S_m \rightarrow \mathbb{C} c \text{ is positive semidefinite}\}$	216
$\bar{\chi}$	$\bar{\chi}(\sigma) = \chi(\sigma), \sigma \in G$	234
χ^G	character of G induced by χ	96
χ_π	irreducible character of S_n	99, 291
χ_r	single-hook character $\chi_{[r, 1^{n-r}]}$	106
D_4	symmetry group of the square	65, 93
$d(G)$	degree sequence of graph G	15
$d^*(G)$	conjugate degree sequence	15
$d_G(v) = d(v)$	degree of vertex v in G	14
d_χ	generalized matrix function	213
$\bar{d}_\chi$	$d_\chi / \chi(e)$	214
d_π	d_{χ_π}	213
$\underline{\Delta}$	system of distinct reps modulo G	163
$\bar{\Delta}$	distinguished subset of Δ	164
$\hat{\Delta}$	distinguished subset of Ω	167, 238
$\delta_{i,j}$	Kronecker delta	27
e	identity permutation	53
e_n	identity permutation of S_n	53
e_n	n -th basis vector	141
e_α^*	$T(G, \chi)(e_\alpha^\otimes)$	161
$E_n(x_1, x_2, \dots, x_k)$	n -th elementary symmetric function	9
ε	alternating character	57
ε_π	restriction of ε to S_π	99
$F(A)$	numerical range of matrix A	37
$F(\pi)$	Ferrers diagram of the partition π	2
$f(\pi)$	trace of partition π	17
$f_\pi(x)$	polynomial associated with $F(\pi)$	278
$F(\sigma)$	number of fixed points of σ	59
$g(n, m)$	number of n -vertex, m -edge graphs	193
G^c	complement of graph G	20

$G_{n,k}$	nondecreasing functions in $\Gamma_{n,k}$	12
G_x	stabilizer subgroup	59
G_α	stabilizer subgroup	162
$GL(n, \mathbb{C})$	invertible matrices in $\mathbb{C}_{n,n}$	75
$GL(V)$	invertible linear operators on V	78
$\Gamma_{n,k}$	functions $\alpha: \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, k\}$	9, 142
$h(A)$	Hadamard's matrix function	42
$H_n(x_1, x_2, \dots, x_n)$	n -th homogeneous symmetric function	11
H_m	$H_m(x_1, x_2, \dots, x_n)$	11, 207
H_ρ	$H_{\rho_1} H_{\rho_2} \dots H_{\rho_r}$	290, 291
$\mathcal{H}_m$	$\{A \in \mathbb{C}_{m,m}: A \geq 0\}$	215
I_n	n -by- n identity matrix	36
I_V	identity operator on V	34
$I(G)$	irreducible characters of G	85
$K(A)$	induced matrix	235
$K(T)$	induced transformation	185
K_n	complete graph on n vertices	19
$K_{\pi,\rho}$	Kostka coefficient	100
$L(G)$	Laplacian matrix of graph G	44
$L(\pi)$	length of the partition π	1
$L(V, W)$	linear transformations from V to W	27
$M(V_1, V_2, \dots, V_m)$	m -linear functionals	127
$M_\pi(x_1, x_2, \dots, x_k)$	monomial symmetric function	6
$MP(\xi)$	multilinearity partition of $\xi \in I(G)$	170
$m_t(\alpha)$	multiplicity of t in α	180
$\mu(\alpha)$	multiplicity partition of α	169
$\binom{n}{\pi}$	a multinomial coefficient	8
$o(S)$	cardinality of the set S	2
$o(\sigma)$	order of the element $\sigma \in S_n$	56
O_x	orbit of G containing x	58
Ω	distinguished subset of $\Gamma_{m,n}$	163
$\text{p adj } (A)$	permanental adjoint of (matrix) A	237
$\text{per } (A)$	permanent of (matrix) A	214
$P_m(A)$	m -th induced power of (matrix) A	236
$P_n(x_1, x_2, \dots, x_k)$	$x_1^n + x_2^n + \dots + x_k^n$	8
$p(n)$	number of partitions of n	21
$P(\sigma)$	permutation operator on $V^{\otimes m}$	151
$\pi \vdash n$	π is a partition of n	1

π^*	conjugate of the partition π	2
$\prod f_i$	product of linear functionals	122
$\prod(A)$	Schur power matrix	218
$\mathcal{Q}_{n,k}$	increasing functions in $\Gamma_{n,k}$	9
$r(\alpha)$	$\alpha(1) + \alpha(2) + \cdots + \alpha(p)$, $\alpha \in \Gamma_{p,n}$	175
$s(A)$	eigenvalues of hermitian A	40
$s(G)$	largest $n - 1$ eigenvalues of $L(G)$	46
S_n	symmetric group of degree n	53
$S_n^{(2)}$	pair group	194
s_α	$\dim(\langle e_{\alpha\sigma}^* : \sigma \in G \rangle)$	165
$s_\pi(x_1, x_2, \dots, x_n)$	Schur polynomial	281
s_π	$s_\pi(x_1, x_2, \dots, x_n)$	283
S_π	Young subgroup	98
$SC_n[x_1, x_2, \dots, x_k]$	symmetric homogeneous polynomials	12, 284
$[T]$	matrix representation of T	33
T^*	adjoint of T	33
$\text{tr}(A)$	trace of A	28
$T(V)$	image of $T \in L(V, W)$	34
$T \geq 0$	T is positive semidefinite hermitian	35
$T_1 \otimes T_2 \otimes \cdots \otimes T_m$	induced transformation	133
$T^{\otimes m}$	$T \otimes T \otimes \cdots \otimes T$ (m -times)	144
$T(G, \chi)$	symmetrizer	153
$V' = L(V, \mathbb{C})$	dual space of V	27
$V^{(2)}$	two element subsets of V	13
$\ v\ $	norm of the vector v	29
$\langle v_1, v_2, \dots, v_k \rangle$	subspace spanned by the indicated vectors	30
$V_1 \times V_2 \times \cdots \times V_m$	cartesian product space	121
$V_1 \otimes V_2 \otimes \cdots \otimes V_m$	tensor product space	126
$V^{\otimes m}$	m -th tensor power of V	141
$\wedge^m V$	m -th exterior power of V	172
$V_\chi(G)$	symmetry class of tensors	154
$v_1 \otimes v_2 \otimes \cdots \otimes v_m$	decomposable tensor	127, 130
$v_\alpha^\otimes$	$v_{\alpha(1)} \otimes v_{\alpha(2)} \otimes \cdots \otimes v_{\alpha(m)}$	142
$v_1 * v_2 * \cdots * v_m$	decomposable symmetrized tensor	155
v_α^*	$v_{\alpha(1)} * v_{\alpha(2)} * \cdots * v_{\alpha(m)}$	161
$v_1 \wedge v_2 \wedge \cdots \wedge v_m$	decomposable element of $\wedge^m V$	172
$v_\alpha^\wedge$	$v_{\alpha(1)} \wedge v_{\alpha(2)} \wedge \cdots \wedge v_{\alpha(m)}$	173
$v_1 \bullet v_2 \bullet \cdots \bullet v_m$	decomposable element of $V_1(S_m)$	179
$v_\alpha^\bullet$	$v_{\alpha(1)} \bullet v_{\alpha(2)} \bullet \cdots \bullet v_{\alpha(m)}$	180
$W^\perp$	orthogonal complement of W	30

$W_1 + W_2$	sum of subspaces	34
$W_1 \oplus W_2$	direct sum of subspaces	34
$W_1 \perp W_2$	orthogonal direct sum of subspaces	34
$w(\alpha)$	weight of $\alpha \in \Gamma_{m,n}$	188
$W_G(x_1, x_2, \dots, x_n)$	pattern inventory	189
$W_G^\chi(x_1, x_2, \dots, x_n)$	character weighted pattern inventory	189
$x \equiv y \pmod{G}$	equivalence modulo G	58
$Z(CG)$	conjugacy class functions in CG	89, 118
$Z_G(y_1, y_2, \dots, y_m)$	cycle index polynomial	189
$Z_G^\chi(y_1, y_2, \dots, y_m)$	character weighted cycle index polynomial	189

References

- A.R. Amir-Moez, *An Introduction to Elements of Multilinear Algebra*, seminar notes, Texas Tech Press, Lubbock, TX.
- T. Ando, Concavity of certain maps on positive definite matrices and applications to Hadamard products, *Linear Algebra Appl.*, **26** (1979), 203–241.
- T. Ando, Inequalities for permanents, *Hokkaido Math. J.*, **10** (1981), 18–36.
- S.R. Arikati and U.N. Peled, Degree sequences and majorization, *Linear Algebra Appl.*, **199** (1994), 179–211.
- Y.-H. Au-Yeung, Some inequalities for the rational power of a nonnegative definite matrix, *Linear Algebra Appl.*, **7** (1973), 347–350.
- K. Balasubramanian, A method for nuclear spin statistics in molecular spectroscopy, *J. Chem. Phys.*, **74** (1981), 6824–6829.
- K. Balasubramanian, Generating functions for the nuclear spin statistics of nonrigid molecules, *J. Chem. Phys.*, **75** (1981), 4572–4585.
- K. Balasubramanian, An algorithm for the generation of nuclear spin species and nuclear spin statistical weights, *J. Comput. Chem.*, **3** (1982), 69–74.
- K. Balasubramanian, Computer generation of nuclear spin species and nuclear spin statistical weights, *J. Comput. Chem.*, **3** (1982), 75–88.
- K. Balasubramanian, Applications of combinatorics and graph theory to spectroscopy and quantum chemistry, *Chem. Rev.*, **85** (1985), 599–618.
- K. Balasubramanian, Imminant polynomials of graphs, *Theoretica Chimica Acta*, **85** (1993), 379–390.
- R.B. Bapat, A bound for the permanent of the Laplacian matrix, *Linear Algebra Appl.*, **74** (1986), 219–223.
- R.B. Bapat, A note on the permanental Hadamard inequality, *Linear & Multilinear Algebra*, **30** (1991), 205–208.
- R.B. Bapat, Interpolating the determinantal and the permanental Hadamard inequality, *Linear & Multilinear Algebra*, **32** (1992), 335–337.

- R.B. Bapat and A.K. Lal, Inequalities for the q -permanent, *Linear Algebra Appl.*, **197&198** (1994), 397–409.
- R.B. Bapat and V.S. Sunder, An extremal property of the permanent and the determinant, *Linear Algebra Appl.*, **76** (1986), 153–163.
- M. Barnabei, A. Brini, and G.-C. Rota, On the exterior calculus of invariant theory, *J. Algebra*, **96** (1985), 120–160.
- A.I. Barvinok, Computational complexity of immanants and representations of the general linear group, *Functional Anal. Appl.*, **24** (1990), 144–145.
- A.I. Barvinok, Statistical sums in combinatorial optimization and related questions, *Functional Anal. Appl.*, to appear.
- D.K. Baxter, The existence of matrices with prescribed characteristic and permanent polynomials, *Linear Algebra Appl.*, **22** (1978), 175–189.
- L.B. Beasley, Maximal groups on which the permanent is multiplicative, *Canad. J. Math.*, **21** (1969), 495–497; Corrigendum, *ibid.* **22** (1970), 192.
- L.B. Beasley and J.L. Brenner, Bounds for permanents, determinants, and Schur functions, *J. Algebra*, **10** (1968), 134–148.
- L.B. Beasley and L.J. Cummings, Permanent groups, *Proc. Amer. Math. Soc.*, **34** (1972), 351–355.
- L.B. Beasley and L.J. Cummings, Permanent groups II, *Proc. Amer. Math. Soc.*, **40** (1973), 358–364.
- L.B. Beasley and L.J. Cummings, Permanent semigroups, *Linear & Multilinear Algebra*, **5** (1978), 297–302.
- L.B. Beasley and L.J. Cummings, Multiplicative properties of generalized matrix functions, *Linear & Multilinear Algebra*, **11** (1982), 23–31.
- L.B. Beasley and L.J. Cummings, On the uniqueness of generalized matrix functions, *Proc. Amer. Math. Soc.*, **87** (1983), 229–232.
- L.B. Beasley and L.J. Cummings, Schur extensions of monomial matrix groups, *Linear Algebra Appl.*, **161** (1992), 11–25.
- R. Bellman, Notes on Matrix Theory II, *Amer. Math. Monthly*, **60** (1953), 173–175.
- G. Birkhoff, Tres observaciones sobre el algebra lineal, *Univ. Nac. Tucumán Rev. Ser. A* **5** (1946), 147–150.
- H. Boerner, *Representations of Groups*, American Elsevier, New York, 1970.
- M. Borowiecki and T. Jozwiak, Computing the permanent polynomial of a multigraph, *Discrete Math.*, **5** (1982), 9–16.
- P. Botti and R. Merris, Almost all trees share a complete set of immanantal polynomials, *J. Graph Theory*, **17** (1993), 467–476.
- N. Bourbaki, Algèbre, Chapter III, *Algèbre Multilinéaire*, Hermann, Paris, 1948.
- M. Bozejko and R. Speicher, An example of a generalized Brownian motion, *Comm. Math. Phys.*, **137** (1991), 519–531.
- J.L. Brenner and R.A. Brualdi, Eigenschaften der permanentenfunktion, *Archiv der Math.*, **18** (1967), 585–586.

- R.A. Brualdi and J.L. Goldwasser, Permanent of the Laplacian matrix of trees and bipartite graphs, *Discrete Math.*, **48** (1984), 1–21.
- R.A. Brualdi and H.J. Ryser, *Combinatorial Matrix Theory*, Cambridge University Press, Cambridge, 1991.
- J.W. Bunce, The Binet–Cauchy theorem and inequalities for the Schur power matrix, *Linear & Multilinear Algebra*, **30** (1991), 5–11.
- W. Burnside, *Theory of Groups of Finite Order*, Dover, New York, 1955.
- C.F. Chan, Positive semi-definiteness in the complex group algebra, *Linear & Multilinear Algebra*, **27** (1990), 189–198.
- C.F. Chan, Letter to the editor, *Linear & Multilinear Algebra*, **35** (1993), 355–356.
- G.-H. Chan, A note on symmetrizers of rank one, *Nanta Math.*, **11** (1978a), 130–133.
- G.-H. Chan, On the triviality of a symmetry class of tensors, *Linear & Multilinear Algebra*, **6** (1978b), 73–82.
- G.-H. Chan, k -characters and the triviality of symmetry classes, *Linear Algebra Appl.*, **25** (1979a), 139–149.
- G.-H. Chan, (k) -characters and the triviality of symmetry classes II, *Nanta Math.*, **12** (1979b), 7–15.
- G.-H. Chan and M.H. Lim, Nonzero symmetry classes of smallest dimension, *Canad. J. Math.*, **32** (1980), 957–968.
- G.-H. Chan and M.H. Lim, A conjecture involving generalized matrix functions, *Linear Algebra Appl.*, **36** (1981), 157–163.
- O. Chan and T.K. Lam, *Hook immanantal inequalities for Laplacians*, unpublished manuscript, 1996.
- O. Chan, T.K. Lam, and H.-K. Tang, *Ordering trees via immanants*, unpublished manuscript, 1996.
- S.-C. Chang, On the vanishing of a (G, σ) space, *Chinese J. Math.*, **4** (1976), 1–7.
- J. Chollet, Is there a permanental analogue to Oppenheim's inequality?, *Amer. Math. Monthly*, **89** (1982), 57–58.
- V. Chvátal and P.L. Hammer, Aggregation of inequalities in integer programming, *Ann. Discrete Math.*, **1** (1977), 145–162.
- J.E. Cohen, Supermultiplicative inequalities for the permanent of nonnegative matrices, *Math. Mag.*, **65** (1992), 41–44.
- R.W. Cottle, Manifestations of the Schur complement, *Linear Algebra Appl.*, **8** (1974), 189–211.
- R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Interscience, New York, 1953.
- L.J. Cummings, Cyclic symmetry classes, *J. Algebra*, **40** (1976), 401–405.
- L.J. Cummings and R. W. Robinson, Linear symmetry classes, *Canad. J. Math.*, **28** (1976), 1311–1319.
- C. Davis, The Toeplitz–Hausdorff Theorem Explained, *Canad. Math. Bull.*, **14** (1971), 245–246.

- J.A. Dias da Silva, On the Schur inequality, *Linear & Multilinear Algebra*, **7** (1979), 343–357.
- J.A. Dias da Silva, Irreducible representations of the full linear group, *Linear Algebra Appl.*, **40** (1981), 161–182.
- J.A. Dias da Silva, On the μ -colorings of a matroid, *Linear & Multilinear Algebra*, **27** (1990), 25–32.
- J.A. Dias da Silva, New conditions for equality of decomposable symmetrized tensors, *Linear Algebra Appl.*, **245** (1996), 353–372.
- J.A. Dias da Silva and M. da Purificação Coelho, (λ, G) -Critical matrices, *Linear Algebra Appl.*, **140** (1990), 1–11.
- J.A. Dias da Silva and A. Fonseca, On the multilinearity partition of an irreducible character, *Linear & Multilinear Algebra*, **20** (1987), 203–218.
- J.A. Dias da Silva and A. Fonseca, Nonzero star products, *Linear & Multilinear Algebra*, **27** (1990), 49–55.
- J.A. Dias da Silva and A. Fonseca, Symmetry classes of tensors: dual indices, *Linear & Multilinear Algebra*, **39** (1995), 245–253.
- D.Ž. Djoković and J. Malzan, Monomial irreducible characters of the symmetric group, *J. Algebra*, **35** (1975), 153–158; Erratum, *Ibid.* **38** (1976), 251.
- P. Doubilet, J. Fox, and G.-C. Rota, The elementary theory of the symmetric group, *Combinatorics, Representation Theory and Statistical Methods in Groups* (T.V. Narayana, et al., eds.), *Pure and Appl. Math.*, v. **57**, Dekker, New York, 1980.
- P. Doubilet, G.-C. Rota, and J. Stein, On the foundations of combinatorial theory: ix combinatorial methods in invariant theory, *Studies in Appl. Math.* **53** (1974), 185–216.
- M.A. Duffner, A note on singular matrices satisfying certain polynomial identities, unpublished manuscript, 1995.
- Ö.N. Eğecioğlu and J.B. Remmel, A combinatorial proof of the Giambelli identity for Schur functions, *Advances in Math.*, **70** (1988), 59–86.
- Ö. N. Eğecioğlu and J.B. Remmel, A combinatorial interpretation of the inverse Kostka matrix, *Linear & Multilinear Algebra*, **26** (1990), 59–84.
- G.M. Engel, Regular equimodular sets of matrices for generalized matrix functions, *Linear Algebra Appl.*, **7** (1973), 243–274.
- I. Faria, Permanent roots and the star degree of a graph, *Linear Algebra Appl.*, **64** (1985), 255–265.
- I. Faria, Multiplicity of integer roots of polynomials of graphs, *Linear Algebra Appl.*, **225** (1995), 15–35.
- W. Feit, *Characters of Finite Groups*, Benjamin, New York, 1967.
- M. Fiedler, Geometry of the numerical range of matrices, *Linear Algebra Appl.*, **37** (1981), 81–96.
- M. Fiedler, *Special Matrices and Their Applications in Numerical Mathematics*, Martinus Nijhoff, Boston, 1986.

- E. Fischer, Über quadratische Formen mit reellen Koeffizienten, *Monat. Math. Physik*, **16** (1905), 234–249.
- E. Fischer, Über den Hadamardschen Determinantensatz, *Archiv der Math. und Phys.*, **13**(3) (1907), 32–40.
- A. Fonseca, On the multilinearity partition of an irreducible character II, *Linear & Multilinear Algebra*, **24** (1989), 191–198.
- A. Fonseca, A survey of the multilinearity partition, *Linear Algebra Appl.*, **170** (1992), 188–193.
- J.S. Frame, G. de B. Robinson, and R.M. Thrall, The hook graphs of the symmetric group, *Canad. J. Math.*, **6** (1954), 316–325.
- R. Freese, Inequalities for generalized matrix functions based on arbitrary characters, *Linear Algebra Appl.*, **7** (1973), 337–345.
- S. Friedland, Matrices with prescribed off-diagonal elements, *Israel J. Math.*, **11** (1972), 184–189.
- S. Friedland, On inverse multiplicative eigenvalue problems for matrices, *Linear Algebra Appl.*, **12** (1975), 127–137.
- S. Friedland, Maximality of the monomial group, *Linear & Multilinear Algebra*, **18** (1985), 1–7.
- G. Frobenius, Über die Charaktere der symmetrischen Gruppe, *Sitzungsberichte Akad. Wiss.*, Berlin (1900), 516–534.
- D. Gale, A theorem on flows in networks, *Pacific J. Math.*, **7** (1957), 1073–1082.
- C. Gamas, Conditions for a symmetrized decomposable tensor to be zero, *Linear Algebra Appl.*, **108** (1988), 83–119.
- R. Gardner, Book review, Invariant Theory by T. A. Springer, *Bull. Amer. Math. Soc. (N.S.)*, **2** (1980), 246–256.
- A.M. Garsia and J. Remmel, Symmetric functions and raising operators, *Linear & Multilinear Algebra*, **10** (1981), 15–43.
- G.Z. Giambelli, Alcune proprietà delle funzioni simmetriche caratteristiche, *Atti Accad. Sci. Torino Cl. Sci. Fis. Mat. Natur.*, **38** (1903), 323–344.
- P.M. Gibson, Eigenvalues of complex tridiagonal matrices, *Proc. Edinburgh Math. Soc.*, **17** (1971), 317–319.
- P.M. Gibson, Localization of the zeros of the permanent of a characteristic matrix, *Proc. Amer. Math. Soc.*, **31** (1972), 18–20.
- P.M. Gibson, Real permanental roots of doubly stochastic matrices, *Linear Algebra Appl.*, **21** (1978), 289–291.
- J.L. Goldwasser, Permanent of the Laplacian matrix of trees with a given matching, *Discrete Math.*, **61** (1986), 197–212.
- M.C. Golumbic, Threshold graphs and synchronizing parallel processes, *Combinatorics* (A. Hajnal and V.T. Sós, eds.), North-Holland, Amsterdam, 1978.

- I.P. Goulden and D.M. Jackson, The enumeration of directed closed Euler trails and directed Hamiltonian circuits by Lagrangian methods, *Europ. J. Combinatorics*, **2** (1981), 131–135.
- I.P. Goulden and D.M. Jackson, Immanants, Schur functions, and the MacMahon master theorem, Research Report CORR 89–33, Faculty of Mathematics, University of Waterloo, 1989.
- I.P. Goulden and D.M. Jackson, Immanants of combinatorial matrices, *J. Algebra*, **148** (1992), 305–324.
- J.A. Green, Polynomial Representations of GL_n , *Lecture Notes in Math*, **830**, Springer-Verlag, Berlin, 1980.
- C. Greene, Proof of a conjecture on immanants of the Jacobi-Trudi matrix, *Linear Algebra Appl.*, **171** (1992), 65–79.
- W.H. Greub, *Multilinear Algebra*, Springer-Verlag, Berlin, 1967.
- R. Grone, A note on the dimension of an orbital subspace, *Linear Algebra Appl.*, **17** (1977a), 283–286.
- R. Grone, Decomposable tensors as a quadratic variety, *Proc. Amer. Math. Soc.*, **64** (1977b), 227–230.
- R. Grone, Eigenvalues and the degree sequences of graphs, *Linear & Multilinear Algebra*, **39** (1995), 133–136.
- R. Grone, C.R. Johnson, E.M. Sa, and H. Wolkowicz, A note on maximizing the permanent of a positive definite hermitian matrix, given the eigenvalues, *Linear & Multilinear Algebra*, **19** (1986), 389–393.
- R. Grone, C.R. Johnson, E.M. Sa, and H. Wolkowicz, Normal matrices, *Linear Algebra Appl.*, **87** (1987), 213–225.
- R. Grone and R. Merris, A bound for the complexity of a simple graph, *Discrete Math.*, **69** (1988), 97–99.
- R. Grone, R. Merris, and W. Watkins, A Hadamard dominance theorem for a class of immanants, *Linear & Multilinear Algebra*, **19** (1986), 167–171.
- R. Grone, R. Merris, and W. Watkins, Cones in the group algebra related to Schur's determinantal inequality, *Rocky Mtn. J. Math.*, **18** (1988), 137–146.
- J. Hadamard, Resolution d'une question relative aux determinants, *Bull. Sci. Math.*, **2** (1893), 240–248.
- J. Hadamard, *The Psychology of Invention in the Mathematical Field*, Princeton University Press, 1945, reprinted by Dover, New York, 1954.
- M. Haiman, Hecke algebra characters and immanant conjectures, *J. Amer. Math. Soc.*, **6** (1993), 569–595.
- M. Hamermesh, *Group Theory*, Addison-Wesley, Reading, MA, 1962.
- P.L. Hammer, T. Ibaraki and B. Simeone, Threshold sequences, *SIAM J. Alg. Disc. Meth.*, **2** (1981), 39–49.
- G.H. Hardy, J.E. Littlewood, and G. Pólya, *Inequalities*, Cambridge University Press, London, 1967.

- W. Hartmann, On the complexity of immanants, *Linear & Multilinear Algebra*, **18** (1985), 127–140.
- R.E. Hartwig and S.B. Morris, The universal flip matrix and the generalized Faro-Shuffle, *Pacific J. Math.*, **58** (1975), 445–455.
- F. Hausdorff, Der Wertworrat einer Bilinearform, *Math. Z.*, **3** (1919), 314–316.
- E. Haynsworth, Determination of the inertia of a partitioned hermitian matrix, *Linear Algebra Appl.*, **1** (1968), 73–81.
- P.B. Henderson and Y. Zalcstein, A graph-theoretic characterization of the PV_{chunk} class of synchronizing primitives, *SIAM J. Comput.*, **6** (1977), 88–108.
- P. Heyfron, Immanant dominance orderings for hook partitions, *Linear & Multilinear Algebra*, **24** (1988), 65–78.
- P. Heyfron, Some inequalities concerning immanants, *Math. Proc. Cambridge Phil. Soc.*, **109** (1991), 15–30.
- P. Heyfron, A generalization of Hadamard's inequality, *Linear & Multilinear Algebra*, **32** (1992), 75–84.
- J. Hollas, *Modern Spectroscopy*, Wiley, New York, 1987.
- R.R. Holmes, Orthogonal bases of symmetrized tensor spaces, *Linear & Multilinear Algebra*, **39** (1995), 241–243.
- R.R. Holmes and T.-Y. Tam, Symmetry classes of tensors associated with certain groups, *Linear & Multilinear Algebra*, **32** (1992), 21–31.
- R. Horaud and H. Sossa, Polyhedral object recognition by indexing, *Pattern Recognition*, **28** (1995), 1855–1970.
- A. Horn, Doubly stochastic matrices and the diagonal of a rotation matrix, *American J. Math.*, **76** (1954), 620–630.
- R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, Cambridge, 1985.
- D.B. Hunter, The construction of irreducible normalized representations of the general linear group, *Linear & Multilinear Algebra*, **13** (1983), 357–366.
- I.M. Isaacs, *Character Theory of Finite Groups*, Academic Press, New York, 1976.
- C.G.J. Jacobi, De functionibus alternantibus earumque divisione per productum e differentiis elementorum conflatum, *J. für die Reine und Angewandte Math.*, **22** (1841), 360–371; *Gesammelte Werke*, **3**, 441–452.
- G. James and A. Kerber, *The Representation Theory of the Symmetric Group*, Addison-Wesley, Reading, MA, 1981.
- G. James and M. Liebeck, Permanents and immanants of hermitian matrices, *Proc. London Math. Soc.*, **55** (1987), 243–265.
- G. James and M. Liebeck, *Representations and Characters of Groups*, Cambridge Univ. Press, Cambridge, 1993.
- C.R. Johnson, Normality and the numerical range, *Linear Algebra Appl.*, **15** (1976), 89–94.
- C.R. Johnson, R. Merris and S. Pierce, Inequalities involving immanants and diagonal products for H -matrices and positive definite matrices, *Portugaliae Math.*, **43** (1985–86), 43–54.

- W.M. Kantor, Homogeneous designs and geometric lattices, *J. Combinatorial Theory Ser. A*, **38** (1985), 66–74.
- O. Kim, J. Chollet, R. Brown, and D. Rauschenberg, Orthonormal bases of symmetry classes with computer-generated examples, *Linear & Multilinear Algebra*, **21** (1987), 91–106.
- R.C. King, The dimensions of irreducible representations of linear groups, *Canad. J. Math.*, **22** (1970), 436–448.
- C. Kostka, Über den Zusammenhang zwischen einigen Formen von symmetrischen Funktionen, *J. für die Reine und Angewandte Math.* **93** (1882), 89–123.
- A. Kräuter, On the distance between two permanental roots of a matrix, *Linear Algebra Appl.*, **93** (1987), 39–55.
- J.P.S. Kung and G.-C. Rota, The invariant theory of binary forms, *Bull. Amer. Math. Soc. (NS)*, **10** (1984), 27–85.
- M.K. Kwong, Conditions under which the order of two non-negative hermitian matrices is preserved when taking integral powers, *Linear Algebra Appl.*, **18** (1977), 223–228.
- A.K. Lal, *Coxeter graphs and positive matrices*, PhD Dissertation, Indian Statistical Institute, New Delhi, 1992.
- T.Y. Lam, Young diagrams, Schur functions, the Gale-Ryser theorem and a conjecture of Snapper, *J. Pure & Applied Algebra*, **10** (1977), 81–94.
- D.W. Lewis, Trace forms, Kronecker sums, and the shuffle matrix, *Linear & Multilinear Algebra*, **40** (1996), 221–227.
- E.H. Lieb, Proofs of some conjectures on permanents, *J. Math. & Mech.*, **16** (1966), 127–134.
- E.H. Lieb, Convex trace functions and the Wigner-Yanase-Dyson conjecture, *Advances in Math.*, **11** (1973), 267–288.
- R.A. Liebler and M.R. Vitale, Ordering the partition characters of the symmetric group, *J. Algebra*, **25**, (1973), 487–489.
- M.H. Lim, Rank k vectors in symmetry classes of tensors, *Canad. Math. Bull.*, **19** (1976), 67–76.
- D.E. Littlewood, *The Theory of Group Characters*, Oxford Univ. Press, London, 1958.
- E.K. Lloyd, Some little-known work of MacMahon on permanents, *Ars Combinatoria*, **16** (1983), 319–323.
- J.S. Lomont, *Applications of Finite Groups*, Dover, New York, 1993.
- I.G. Macdonald, *Symmetric Functions and Hall Polynomials*, 2nd. ed., Oxford Univ. Press, 1995.
- M. Marcus, The permanent analogue of the Hadamard determinant theorem, *Bull. Amer. Math. Soc.*, **69** (1963), 494–496.
- M. Marcus, *Finite Dimensional Multilinear Algebra*, Part I, Marcel Dekker, New York, 1973.
- M. Marcus, *Finite Dimensional Multilinear Algebra*, Part II, Marcel Dekker, New York, 1975.

- M. Marcus and J. Chollet, Linear groups defined by decomposable tensor equalities, *Linear & Multilinear Algebra*, **8** (1980), 207–212.
- M. Marcus and J. Chollet, The index of a symmetry class of tensors, *Linear & Multilinear Algebra*, **11** (1982), 277–281.
- M. Marcus and J. Chollet, On the equality of decomposable symmetrized tensors, *Linear & Multilinear Algebra*, **13** (1983), 253–266.
- M. Marcus and J. Chollet, Construction of orthonormal bases in higher symmetry classes of tensors, *Linear and Multilinear Algebra*, **19** (1986), 133–140.
- M. Marcus and W.R. Gordon, The structure of bases in tensor spaces, *Amer. J. Math.*, **92** (1970), 623–640.
- M. Marcus and R. Merris, A relation between the permanent and determinantal adjoints, *J. Australian Math. Soc.*, **15** (1973), 270–271.
- M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Allyn & Bacon, Boston, 1964. Reprinted by Prindle, Weber & Schmidt, Boston.
- M. Marcus and H. Minc, Generalized matrix functions, *Trans. Amer. Math. Soc.*, **116** (1965a), 316–329.
- M. Marcus and H. Minc, Permanents, *Amer. Math. Monthly*, **72** (1965b), 577–591.
- M. Marcus and M. Newman, Inequalities for the permanent function, *Annals of Math.*, **75** (1962), 47–62.
- M. Marcus and S. Pierce, Extensions of the Minkowski inequality, *Linear Algebra Appl.*, **1** (1968), 13–27.
- M. Marcus and G. Soules, Some inequalities for combinatorial matrix functions, *J. Combinatorial Theory*, **2** (1967), 145–163.
- A.W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York, 1979.
- B.D. McKay, On the spectral characteristics of trees, *Ars Combinat.*, **3** (1977), 219–232.
- R. Merris, Equality of decomposable symmetrized tensors, *Canad. J. Math.*, **27** (1975a), 1022–1024.
- R. Merris, Two problems involving Schur functions, *Linear Algebra Appl.*, **10** (1975b), 155–162.
- R. Merris, Relations among generalized matrix functions, *Pacific J. Math.*, **62** (1976), 153–161.
- R. Merris, Nonzero decomposable symmetrized tensors, *Linear Algebra Appl.*, **17** (1977), 287–292.
- R. Merris, The structure of higher degree symmetry classes of tensors II, *Linear & Multilinear Algebra*, **6** (1978), 171–178.
- R. Merris, Extensions of the Minkowski determinant theorem, *Portugaliae Math.*, **38** (1979), 149–153.
- R. Merris, Pattern inventories associated with symmetry classes of tensors, *Linear Algebra Appl.*, **29** (1980), 225–230.

- R. Merris, Extensions of the Hadamard determinant theorem, *Israel J. Math.*, **46** (1983), 301–304.
- R. Merris, The second immanantal polynomial and the centroid of a graph, *SIAM J. Alg. Disc. Meth.*, **7** (1986), 484–503.
- R. Merris, Degree maximal graphs are Laplacian integral, *Linear Algebra Appl.*, (1994a), 381–389.
- R. Merris, Laplacian matrices of graphs: A survey, *Linear Algebra Appl.*, **197/198** (1994b), 143–176.
- R. Merris, A survey of graph Laplacians, *Linear & Multilinear Algebra*, **39** (1995), 19–31.
- R. Merris, *Combinatorics*, PWS, Boston, 1996.
- R. Merris and S. Pierce, The Bell numbers and r -fold transitivity, *J. Combinatorial Theory*, **12** (1972), 155–157.
- R. Merris and S. Pierce, Elementary divisors of higher degree associated transformations, *Linear & Multilinear Algebra*, **1** (1973), 241–250.
- R. Merris and W. Watkins, Tensors and graphs, *SIAM J. Alg. Disc. Meth.*, **4** (1983), 534–547.
- R. Merris and W. Watkins, An invariant theory approach to graph enumeration, *Linear Algebra Appl.*, **61** (1984), 277–285.
- R. Merris and W. Watkins, Inequalities and identities for generalized matrix functions, *Linear Algebra Appl.*, **64** (1985), 223–242.
- H. Minc, *Permanents*, *Encyclopedia of Math.*, v. 6, Addison–Wesley, Reading, 1978.
- H. Minc, Theory of permanents 1978–1981, *Linear & Multilinear Algebra*, **12** (1983), 227–263.
- H. Minkowski, Diskontinuitätsbereich für arithmetische Äquivalenz, *J. für die Reine und Angewandte Math.*, **129** (1905), 220–274; *Gesammelte Abhandlungen*, Ch. 21.
- F.D. Murnaghan, The characters of the symmetric group, *Amer. J. Math.*, **59** (1937), 739–753.
- G. Murtaza and M.A. Rashid, Duality of a Young diagram describing a representation and dimensionality formulas, *J. Math. Phys.*, **14** (1973), 1196–1198.
- T. Nakayama, On some modular properties of irreducible representations of a symmetric group I and II, *Japanese J. Math.*, **17** (1940), 165–184 and 411–423.
- M. Newman, *Matrix Representations of Groups*, U.S. National Bureau of Standards, Applied Math. Series, **60**, U.S. Govt. Printing Office, Washington, D.C. 1967.
- G.N. de Oliveira, A conjecture and some problems on permanents, *Pacific J. Math.*, **32** (1970), 495–499.
- G.N. de Oliveira, Note on the function $\text{per}(\lambda I - A)$, *Rev. Fac. Ciencias Lisboa (Ser. 2)*, **13** (1971), 199–201.
- G.N. de Oliveira, On the multiplicative inverse eigenvalue problem, *Canad. Math. Bull.*, **15** (1972), 189–193.
- G.N. de Oliveira, *Generalized Matrix Functions*, Estudos, Instituto Gulbenkian de Ciência, Oeiras, 1973.

- G.N. de Oliveira, A.P. Santana, and J.A. Dias da Silva, Note on the equality of star products, *Linear & Multilinear Algebra*, **14** (1983), 157–163.
- G.N. de Oliveira and J.A. Dias da Silva, Equality of decomposable symmetrized tensors and $\star$ -matrix groups, *Linear Algebra Appl.*, **49** (1983), 191–219.
- G.N. de Oliveira and J.A. Dias da Silva, On matrix groups defined by certain polynomial identities, *Portugaliae Math.*, **43** (1985–86), 77–92.
- A. Oppenheim, Inequalities connected with definite hermitian forms, *J. London Math. Soc.*, **5** (1930), 114–119.
- J. Orlin, The minimal integral separator of a threshold graph, *Ann. Disc. Math.*, **1** (1977), 415–419.
- D.S. Passman, *Permutation Groups*, Benjamin, New York, 1968.
- T.H. Pate, Inequalities relating groups of diagonal products in a Gram matrix, *Linear & Multilinear Algebra*, **11** (1982), 1–17.
- T.H. Pate, Immanants and decomposable tensors that symmetrize to zero, *Linear & Multilinear Algebra*, **28** (1990), 175–184.
- T.H. Pate, Immanant inequalities and partition node diagrams, *J. London Math. Soc.*, **46** (1992), 65–80.
- T.H. Pate, Inequalities involving immanants, *Linear Algebra Appl.*, **212/213** (1994a), 31–44.
- T.H. Pate, Immanant inequalities, induced characters, and rank two partitions, *J. London Math. Soc.*, **49** (1994b), 40–60.
- T.H. Pate, Row appending maps, Ψ -functions, and immanant inequalities for hermitian positive semidefinite matrices, manuscript, 1996.
- U.N. Peled, Threshold graph enumeration and series-product identities, *Congressus Numerantium*, **29** (1980), 735–738.
- R.D. Poshusta, Algebrants in many electron quantum mechanics. II. New computational algorithms, *International J. Quantum Chem.*, **25** (1991), 225–234.
- R.D. Poshusta and D.B. Kinghorn, Algebrants in many-electron quantum mechanics: Applications of generalized determinants or matrix functions, *International J. Quantum Chem.*, **41** (1992), 15–42.
- R.C. Read, The use of S -functions in combinatorial analysis, *Canad J. Math.*, **20** (1968), 808–841.
- G.-C. Rota and J.A. Stein, Symbolic method in invariant theory, *Proc. Natl. Acad. Sci. USA*, **83** (1986), 844–847.
- E. Ruch and I. Gutman, The branching extent of graphs, *J. Combinatorics, Information & System Sci.*, **4** (1979), 285–295.
- H.J. Ryser, Combinatorial properties of matrices of 0's and 1's, *Canad J. Math.*, **9** (1957), 371–377.
- E.M. de Sá, Multiple roots of diagonal multiples of a square matrix, *Discrete Math.*, **36** (1981), 57–67.
- B.E. Sagan, The ubiquitous Young tableau, *Invariant Theory and Tableaux, IMA Volumes in Math. Appl.* **19** (D. Stanton, ed.), Springer-Verlag, New York, 1990, pp. 262–298.

- B.E. Sagan, *The Symmetric Group*, Wadsworth & Brooks/Cole, Belmont, CA, 1991.
- I.V. Schensted, *A Course on the Application of Group Theory to Quantum Mechanics*, NEO Press, Peaks Island, ME, 1976.
- I. Schur, Potenzreihen im Innern des Einheitskreises, *J. für die Reine und Angewandte Math.*, **147** (1917), 205–232.
- I. Schur, Über endliche Gruppen und Hermitesche Formen, *Math. Z.*, **1** (1918), 184–207.
- I. Schur, Über eine Klasse von Mittelbildungen mit Anwendungen die Determinantentheorie, *Sitzungsberichte Math. Gesellschaft Berlin*, **22** (1923), 9–20.
- A.J. Schwenk, Almost all trees are cospectral, *New Directions in the Theory of Graphs*, Academic Press, New York, 1973, 275–307.
- G. Sierksma and H. Hoogeveen, Seven criteria for integral sequences being graphic, *J. Graph Theory*, **15** (1991), 223–231.
- N.J.A. Sloane, Error-correcting codes and invariant theory: New applications of a nineteenth-century technique, *Amer. Math. Monthly*, **84** (1977), 82–107.
- E. Snapper, Group characters and nonnegative integral matrices, *J. Algebra*, **19** (1971), 520–535.
- G. Soules, *Matrix functions and the Laplace expansion theorem*, PhD Dissertation, Univ. Calif. Santa Barbara, July 1966.
- G. Soules, Constructing symmetric nonnegative matrices, *Linear & Multilinear Algebra*, **13** (1983), 241–251.
- G. Soules, An approach to the permanental dominance conjecture, *Linear Algebra Appl.*, **201** (1994), 211–229.
- R.P. Stanley, Theory and application of plane partitions, Part 1, *Studies in Appl. Math.*, **50** (1971), 167–188.
- R.P. Stanley, Invariants of finite groups and their applications to combinatorics, *Bull. Amer. Math. Soc. (N.S.)*, **1** (1979), 475–511.
- R.P. Stanley and J.R. Stembridge, On immanants of Jacobi-Trudi matrices and permutations with restricted position, *J. Combinatorial Theory Ser. A*, **62** (1993), 261–279.
- W. Steeb, *Kronecker Products of Matrices and Applications*, Wissenschaftsverlag, Mannheim, 1991.
- J.R. Stembridge, Some conjectures for immanants, *Canad J. Math.*, **44** (1992), 1079–1099.
- V. Stok, Лерсаненты: Теорѣя ц прѣлоисенѣя. — Межѣузоѣскѣѣ сѣборнѣк. омѣ. пер. Г. П. ЕгорѣчѣѣКрЛІИ Красмѣярек. 180ѣ. (1990), 73–88.
- G.P. Thomas, A combinatorial interpretation of the wreath product of Schur functions, *Canad J. Math.*, **28** (1976a), 879–884.
- G.P. Thomas, *Further Results on Baxter Sequences and Generalized Schur Functions*, Lecture Notes in Math. no. 579, Springer Verlag, Berlin, 1976b.
- O. Toeplitz, Das Algebraische Analogen zu einem Satze von Fejér, *Math. Z.*, **2** (1918), 187–197.
- N. Trudi, Intorno un determinante più generale, *Giornale Mat.*, **2** (1864), 152–158, 180–186.

- J. Turner, Generalized matrix functions and the graph isomorphism problem, *SIAM J. Appl. Math.*, **16** (1968), 520–526.
- L.G. Valiant, The complexity of computing the permanent, *Theoret. Comp. Sci.*, **8** (1979), 189–201.
- A. Vrba, The permanent of the Laplacian matrix of a bipartite graph, *Czech. J. Math.*, **36** (1986a), 7–17.
- A. Vrba, Principal subpermanents of the Laplacian matrix, *Linear & Multilinear Algebra*, **19** (1986b), 335–346.
- B.-Y. Wang and M.-P. Gong, A high symmetry class of tensors with an orthogonal basis of decomposable symmetrized tensors, *Linear & Multilinear Algebra*, **30** (1991a), 61–64.
- B.-Y. Wang and M.-P. Gong, The subspaces and orthonormal bases of symmetry classes of tensors, *Linear & Multilinear Algebra*, **30** (1991b), 195–204.
- E.T.H. Wang, A new class of finite permanent groups, *J. Combinatorial Theory*, **17** (1974), 261–264.
- W. Watkins, The cone of positive generalized matrix functions, *Linear Algebra Appl.*, **181** (1993), 1–28.
- R. Westwick, A note on symmetry classes of tensors, *J. Algebra*, **15** (1970), 309–311.
- D.E. White, Monotonicity and unimodality of the pattern inventory, *Advances in Math.*, **38** (1980a), 101–108.
- D.E. White, A Pólya interpretation of the Schur function, *J. Combinatorial Theory, Ser. A*, **28** (1980b), 272–281.
- S.G. Williamson, On a class of combinatorial inequalities, *J. Combinatorial Theory*, **6** (1969), 359–369.
- S.G. Williamson, Symmetry operators of Krantz products in enumerative combinatorial theory, *J. Combinatorial Theory*, **11** (1971), 122–138.
- B.G. Wybourne, *Symmetry Principles and Atomic Spectroscopy*, Interscience, New York, 1970.

Index

- abelian group 91, 116
- adjacency matrix of a graph 44
- adjacent
 - edges of a graph 13
 - vertices of a graph 13
- adjoint
 - classical 237, 261
 - of a linear transformation 33, 141
 - permanental 237
- adjugate 237, 261
- algebrant 220
- alternating
 - character 57, 116
 - group 57
- Amir-Moez, A.R. viii
- Ando, T. 144, 229, 237
- Arikati, S.R. 46
- associated transformation 185
- automorphisms of a graph 194, 252
- Au-Yeung, Y.-H. 50

- Balasubramanian, K. 198, 255
- Bapat, R.B. 221, 224, 255, 256, 262
- Barnabei, M. 312
- Barvinok, A.I. 296
- Baxter, D.K. 254
- Beasley, L.B. 230, 254, 258
- Bell numbers 70, 71
- Bellman, R. 262

- Bessel's inequality 32
- bilinear function 122
- binary code 207
- Binet-Cauchy theorem
 - (see Cauchy-Binet theorem)
- Birkhoff, G. 5
- bisymmetric operators 184, 275
- Boerner, H. 106, 107, 276, 278
- border of a Ferrers diagram 107
- border strip 107
- Born-Oppenheimer approximation 198
- Borowiecki, M. 255
- Botti, P. 254
- boundary of a Ferrers diagram 107
- boundary part 107
- Bourbaki, N. viii, 132
- Bozejko, M. 261
- branching theorem 109
- Brenner, J.L. 254
- Brini, A. 243
- Brualdi, R. 99, 254, 255
- Bunce, J.W. 257
- Burnside's lemma 59, 168
- Burnside's theorem 84
- Butler, P.H. 278

- Cartan, E. vii, 121
- cartesian product space 121
- Cauchy-Binet theorem 229, 231

- Cauchy-Schwarz inequality 31, 228
- Cauchy's identity 211
- Cayley's theorem 62, 77
- center of a group 115
- Chan, C.F. 219, 224
- Chan, G.-H. 229, 280
- Chan, O. 255, 263
- Chang, S.C. 280
- character 85
 - degree 89
 - table 92, 95, 116, 209
- character weighted
 - cycle index polynomial 189
 - pattern inventory 189
- characteristic function 90, 119
- Chollet, J. 167, 233, 256, 280
- Chvátal, V. 46
- class function 89
- class number 89
- classical adjoint 237, 261
- Coelho, M. da P. 233
- Cohen, J.E. 237
- Cohen-Macaulay algebras 243
- coimmanantal trees 254ff
- color patterns 186, 208
- coloring 186ff, 298
- commutative diagram 124
- commutator 116
- commutator subgroup 116
- complement
 - of a graph 20, 52
 - orthogonal 30
- complementary sequence 174
- complete
 - graph 19, 194, 298
 - set of inequivalent irreducible representations 81
- completely symmetric tensors 179
- compound matrix 236, 261, 279ff, 300
- computational complexity 296
- cone 256
- conform 171
- conjugacy class 89, 93, 119
- conjugate
 - degree sequence 15, 46
 - groups 110
 - partition 2, 172, 226
 - permutations 58, 110
 - transpose 28, 36
- contragredient representation 114, 234, 274
- convex
 - combination 5, 37
 - hull 37, 50, 222
- coordinate representation 48
- correlation matrix 50, 257
- correspond 63
- Cottle, R. 51
- Courant, R. 38
- Courant-Fisher theorem 39
- cross product 205
- Cummings, L.J. 230, 258, 280
- cycle
 - in a graph 20
 - index polynomial 189ff, 207ff
 - of a permutation 53
 - structure 54
 - type 54, 55, 62
- cyclic group 59
- D_4 65, 92, 97, 192, 197ff
- Davis, C. 38
- decomposable
 - symmetrized tensor 155ff, 169ff, 179, 226, 233
 - tensor 126, 130ff, 145
- degree
 - of a character 89
 - of a permutation 53
 - of a permutation group 58
 - of a representation 75, 270
 - of a vertex 14
 - sequence 15, 46
- Dias da Silva, J.A. viii, 167, 170–172, 219, 233, 276
- direct sum
 - of matrices 36, 39, 76
 - of vector spaces 34, 155, 165
- Dirichlet, P.G.L. 42
- disjoint cycle factorization 54
- disjoint cycles of a permutation 53
- Djoković, D.Ž. 277
- dominate 40
- dot product 29

- Doubilet, P. 99, 243, 289
- double coset 118
- double coset number 99
- doubly stochastic matrix 5
- doubly transitive permutation group 60, 62, 70, 88, 118
- dual
 - basis 28
 - space 27, 128
- Duffner, M.A. 233
- edge of a graph 13
- edge coloring of the complete graph 194, 298
- Eğecioğlu, Ö.N. 289, 297
- elementary symmetric function(s) 9, 22ff, 240, 281, 285ff
- energy 198
- Engel, G.M. 254
- equivalent
 - colorings 186
 - cycles of a permutation 53
 - modulo a permutation group 58, 161
 - representations 78, 86
- even permutation 57
- exterior power 172
- faithful representation 75
- Faria, I. 255, 263
- Feit, W. 110
- Ferrers, N.M. 2
- Ferrers diagram 2, 99ff, 225, 277
- Fiedler, M. 27, 38
- field of values 37
- Fischer, E. 38, 213
- Fischer's inequality 213ff, 224, 256
- fixed point
 - character 88
 - of a permutation 53, 59
- Fonseca, A. viii, 170, 171, 205
- Fox, J. 99, 289
- Frame, J.S. 106
- Frame-Robinson-Thrall hook length
 - formula 106, 277, 278
- Fredholm, E. 42
- free vector space 125, 252
- Freese, R. 165, 224
- Freese's theorem 165, 203
- Friedland, S. 253, 254
- Frobenius, G. 84, 291
- Frobenius
 - notation 24
 - reciprocity theorem 98, 203
- Frobenius's formulas for symmetric group
 - characters 293
- full linear group 265
- fully reducible 78, 270
- fundamental theorem of symmetric
 - functions 9, 284
- Gale, D. 24
- Gale-Ryser theorem 25, 99
- Gamas, C. 171
- Gamas's theorem 171, 204
- Garsia, A. 289
- Gardner, R. 243
- general linear group 265
- generalized matrix function 213, 226
- Giambelli, G.Z. 297
- Giambelli's identity 297
- Gibson, P.M. 254
- Goldwasser, J.L. 255
- Golumbic, M.C. 46
- Gong, M.-P. 167
- Gordon, W.R. 181
- Goulden, I. 255, 289, 295, 296
- Gram matrix 49
- Gram-Schmidt process 29
- graph 13
 - automorphism(s) 194, 252
 - coimmanantal 254
 - cycle 20
 - invariant 14, 46, 251ff
 - isomorphism 45
- graphic sequence 16
- Grassmann space 172
- Green, J.A. 270
- Greene, C. 289
- Greub, W.H. viii, 132
- Grone, R. viii, 38, 46, 167, 181, 215, 225, 237, 255, 257
- Gutman, I. 17, 46
- Hadamard, J. 42, 213
- Hadamard

- matrix 150
- product 138
- theorem for permanents 224ff
- Hadarnard's
 - function 42, 225, 262
 - theorem 42, 213, 256, 260, 262
- Haiman, M. 289
- Hamermesh, M. 276
- Hamiltonian
 - cycle 264
 - (differential) operator 198
 - path 264
- Hammer, P.L. 46
- Hardy, G.H. 5
- Hartman, W. 296
- Hartwig, R.E. 147
- Hasse diagram 4
- Hasselbarth, W. 17
- Hausdorff, F. 38, 222
- Haynsworth, E. 51
- height of a border strip 107
- Henderson, P.B. 46
- hermitian
 - matrix 34ff
 - operator 34, 48
- Heyfron, P. 225
- Heyfron's
 - inequalities 255
 - theorem 225
- Hodge star mapping 205
- Hollas, J. 111, 198
- Holmes, R.R. 167
- homogeneous
 - polynomial(s) 9, 180ff, 246ff, 268
 - symmetric function(s) 11, 207, 240, 281, 285ff, 291ff, 301
- Hoogetveen, H. 17
- hook 106
- hook length 106
- Horaud, R. 254
- Horn, A. 42
- Horn, R. 27
- Hunter, D.B. 276
- Ibaraki, T. 46
- identity
 - matrix 36
 - operator 34
- image
 - of a linear transformation 34, 124
 - of a multilinear function 124
- immanant(s) 213, 283, 294, 295
- immanantal polynomial 254ff
- incident 13
- induced
 - action 67, 187, 194, 251
 - basis 136ff, 142, 234
 - character 96ff, 116, 203, 292
 - inner product 141, 148, 162, 226
 - matrix 235, 261
 - permutation 161
 - power matrix 236, 242ff, 248, 280ff, 300
 - representation 96, 115
 - transformation 133ff, 148, 185, 227, 239, 271ff
- inequivalent irreducible representations 81
- inner product 28, 35
- inner product space 28, 86
- interlacing inequalities 43
- invariant
 - of a graph 14, 46, 251ff
 - of a group 58, 245
 - similarity 254
 - subspace 48, 81
- inversion 261
- irreducible
 - character 85ff
 - representation 78
- irreducible set of operators 81
- isolated vertex 13, 15
- isomorphic
 - graphs 14, 45
 - universal pairs 126
 - vector spaces 28
- isomorphism
 - of graphs 14
 - of vector spaces 28
- Issacs, I.M. 110
- Jackson, D. 255, 289, 295, 296
- Jacobi, C.G.J. 289
- Jacobi's identity 237, 288
- Jacobi-Trudi matrix 289
- James, G. 99, 101, 103, 105–107, 111, 224, 276, 278

- Johnson, C.R. 27, 38, 237, 254, 257
 Jozwiak, T. 255
- Kantor, W.M. 62
 Kerber, A. 99, 101, 103, 105–107, 276, 278
 Kim, O. 167
 King, R.C. 278
 Kinghorn, D.B. 220
 Kostka, C. 100
 Kostka
 coefficient 100ff, 119, 240, 280ff, 291ff
 matrix 103ff
 Kräuter, A. 254
 Kronecker
 delta 27
 power 144, 149, 270ff
 product 137ff, 147
 Kung, J.P.S. 243
 Kwong, M.S. 50
- Lal, A.K. 237, 262
 Lam, T.K. 255, 263
 Lam, T.Y. 103
 Laplace expansion theorem 175
 Laplacian
 matrix of a graph 44, 46, 51, 254
 spectrum 46
 length
 of a binary code 207
 of a cycle of a permutation 53
 of a hook 106
 of a partition 1
 of a reduction 80
 Lewis, D.W. 147
 lexicographic order 10, 103, 136ff, 142, 234
 Lieb, E.H. 144, 224
 Lieb's inequality 224, 226
 Liebeck, M. 111, 224
 Liebler, R.A. 103
 Lim, M.H. 159, 229
 linear
 character 85, 116, 234ff
 functional 27, 28, 31
 operator 27
 representation 91
- Littlewood, D.E. 5, 276, 278, 289, 295
 Littlewood's theorem 295ff
 Lloyd, E.K. 296
 Lomont, J.S. 111
- Macdonald, I.G. 270, 289
 majorization 3, 4, 40, 46, 103, 120, 169ff, 226, 258
 Malzan, J. 277
 Marcus, M. viii, 27, 132, 167, 175, 177, 178, 181, 205, 224, 229, 233, 237, 257, 270, 276, 280
 Marcus's inequality 224
 Marshall, A.W. 5, 42
 Maschke's theorem 79, 113, 270
 matrix group 245
 matrix representation
 of a group 75ff
 of a linear transformation 33, 36, 48
 of an induced linear transformation 138
 maximal graphic partition 19
 McKay, B.D. 254
 Merris, R. 6, 46, 70, 103, 159, 167, 190, 215, 224, 225, 229, 237, 253–255, 257, 259, 296–298
 Minc, H. 27, 175, 224, 257
 Minkowski, H. 229
 Minkowski determinant theorem 229
 m -linear function 122
 model 127, 131ff, 156
 Molien series 249
 Molien's theorem 248
 monomial
 matrix 253
 similarity 254
 symmetric function 6, 187, 208ff, 281
 Morris, S.B. 147
 multilinear
 extension 123ff
 function 122, 145
 multilinearity partition 170ff, 204
 multinomial
 coefficient 6
 theorem 6, 8
 multiple transitivity 70
 multiplicity partition 169ff, 180, 209, 239, 282

- Murnaghan, F.D. 107
 Murnaghan-Nakayama rule 107
 Murtaza, G. 278
- Nakayama, T. 107
 Newman, M. 84, 177
 Newton, I. 22
 Newton's identities 23, 283, 286
 NMR
 spectroscopy 197ff
 wave equation 199
 norm of a vector 29
 normal
 matrix 37, 49
 operator 37
 normalized
 generalized matrix function 214
 immanant 225ff
 nuclear
 magnetic resonance (NMR) 198
 spin character 199
 spin quantum number 199
 spin species 199
 nullity partition 172
 numerical range 37, 221
- odd permutation 57
 Oliveira, G.N. de viii, 233, 254
 Olkin, I. 5, 42
 Oppenheim, A. 256
 Oppenheim's inequality 256
 orbit 58
 orbital subspace 165
 order of a permutation 56
 orientation of a graph 51
 oriented vertex-edge incidence matrix 52
 Orlin, J. 46
 orthogonal
 complement 30
 direct sum 34, 155, 165
 projection 34, 48
 vectors 29
 orthogonality relations
 of the first kind 85, 90, 118
 of the second kind 91, 118
 orthonormal
 basis 29, 30, 37, 90, 234ff
 set of vectors 29, 86
- pair group 194, 251, 298
 Parseval's identity 32, 230
 part of a partition 1
 partition 1, 21, 94, 100, 109
 partitioned matrix 213
 Passman, D. 170
 Pate, T. viii, 171, 221, 225, 226, 258
 Pate's theorem 225
 path 263
 pattern inventory 187ff
 character weighted 189
 Pauli exclusion principle 220
 Peled, U.N. 46
 permanent 214, 224, 262, 294ff
 permanental
 adjoint 237
 compound 237
 dominance conjecture 224ff, 255
 polynomial 263
 roots 263
 permutation 53
 group 58
 matrix 75, 77, 95
 operator 147, 151, 184, 300
 representation 88, 95
 similar 45, 254
 Petersen graph 14
 Pierce, S. viii, 70, 167, 181, 229, 254, 258
 plane symmetry 65
 Pólya, G. 5
 Pólya's theorem 190, 193
 polynomial
 invariant 245ff, 263
 representation 266ff
 Poshusta, R.D. 220
 positive definite hermitian
 matrix 36, 42, 44, 49
 operator 35
 positive semidefinite function 216
 positive semidefinite hermitian
 matrix 36, 44
 operator 35
 power sum symmetric function 8, 22ff, 190, 284
 principal
 character 99, 116
 representation 76
 projection operator 34

- q-permanent 262
- quantum mechanical
 - energy 198
 - spin functions 220
 - state 198
 - wave function 198
- quaternion group 115
- rank partition 172
- Rashid, M.A. 278
- rational representation 265ff
- reach of a multilinear function 124
- Read, R.C. 289
- reducible representation 78
- reduction of a representation 80
- reflections 67
- regular
 - boundary part 107
 - representation 77, 88, 99, 216
- Remmel, J. 289, 297
- representation 75
- representation theorem for linear functionals 31
- r -fold transitive 70
- Robinson, R.W. 280
- Rota, G.-C. 99, 243, 289
- rotational symmetries of the cube 66, 68, 72, 112, 208ff
- Ruch, E. 17, 46
- Ryser, H.J. 24, 99
- Sá, E.M. de 38, 237, 254, 257
- Sagan, B. 101, 106, 107, 276
- Santana, A.P. 233
- Schensted, I.V. 111, 198, 220
- Schrödinger equation 198
- Schur, I. 42, 51, 84, 213, 214, 221
- Schur
 - complement 50, 51, 257, 259, 260
 - concave/convex 24, 42
 - polynomial(s) 189, 240, 281ff, 295ff, 301
 - power matrix 218ff, 257
 - product 138
 - relations 82ff, 274
 - triangularization theorem 36, 239
- Schur's
 - inequality 214ff, 227ff, 259ff
 - lemma 81
- Schwarz inequality (see Cauchy-Schwarz inequality)
- Schwenk, A.J. 254
- self adjoint operator 34
- self conjugate partition 2, 100
- semiregular permutation group 71
- Sierksma, G. 17
- Silva, J.A.D. da (see Dias da Silva, J.A.)
- Simeone, B. 46
- single-hook
 - character 106, 225
 - immanant(s) 225, 255, 264, 296
- skew Schur polynomial 289
- skew symmetric
 - matrices 156, 202
 - tensors 172
- Sloane, N.J.A. 243
- Snapper, E. 103
- Sossa, H. 254
- Soules, G. 175, 224
- Soules's conjecture 224
- spanning set of vectors 30
- spectral theorem 37, 49
- spectrum
 - Laplacian 46
 - NMR 198ff
 - of a matrix 39, 40
- Speicher, R. 261
- stabilizer subgroup 59, 162, 202
- standard inner product 28, 29, 36
- Stanley, R. 243, 289
- star 263
- stationary state(s) 198
- Stein, J. 243
- Stembridge, J.R. 289
- Stirling Number of the
 - First Kind 69
 - Second Kind 149
- Strok, V. 254
- submatrix 174
- sum of vector spaces 34
- Sunder, V.S. 221, 256
- superposition of states 198
- Sylvester-Franke theorem 241
- symmetric
 - functions, fundamental theorem 9, 284
 - group 53, 93

- homogeneous polynomials 12, 284
- polynomial 6
- with respect to G and χ 157, 202
- symmetries of the cube 66ff
- symmetrizer 153
- symmetry 64ff
 - class of tensors 154, 272
 - group 65ff
- system of distinct representatives for
 - equivalence classes mod G 163
 - inequivalent color patterns 187
- Tam, T.Y. 167
- Tang, H.-K. 263
- tensor
 - algebras 132
 - power 141
 - product 126ff
- Thomas, G.P. 289
- Thrall, R.M. 106
- threshold graph 19, 20, 25, 46
- Toeplitz, O. 38, 222
- Toeplitz-Hausdorff theorem 38, 222
- trace
 - function 28, 29, 47
 - of a partition 17
- transitive homomorphic image 63, 71
- transitive permutation group 58, 60, 62, 70, 95
- transpose 28, 157
- transposition 56, 120
- tree 20, 25, 254
- trivial extension 215
- Trudi, N. 289
- Turner, J. 254
- unit vector 29
- unitary
 - matrix 36, 48
 - operator 36, 49
 - representation 110
- universal factorization property 124, 158
 - for symmetric multilinear functions 158
- universal pair 124, 145
- Valiant, L.G. 296
- vertex
 - degree 14
 - isolated 13
 - of a graph 13
- Vitale, M.R. 103
- volume element 205
- Vrba, A. 255
- Wang, B.-Y. 167
- Wang, E.T.H. 230
- Watkins, W. viii, 215, 219, 224, 225, 253, 259, 296, 298
- Watkins's theorem 215, 227ff
- weight of a color pattern 187ff
- Westwick, R. 280
- Weyl's [irrelevance] principle 265ff
- White, D. 103, 190
- Williamson, S.G. 190, 259
- Wolkowicz, H. 38, 237, 257
- Wybourne, B.G. 278
- Young, A. 98
- Young
 - subgroup 98, 226, 236ff
 - tableau 16, 105, 106
- Young's rule 100
- Zalcstein, Y. 46