

Symmetry Studies

Marlos A. G. Viana

$$\begin{array}{c}
 \left[\begin{array}{cccccccc}
 aaa & aac & aag & aat & caa & cac & cag & cat \\
 aca & acc & acg & act & cca & ccc & ccg & cct \\
 aga & agc & agg & agt & cga & cgc & cgg & cgt \\
 ata & atc & atg & att & cta & ctc & ctg & ctt \\
 gaa & gac & gag & gat & taa & tac & tag & tat \\
 gca & gcc & gcg & gct & tca & tcc & tcg & tct \\
 gga & ggc & ggg & ggt & tga & tgc & tgg & tgt \\
 gta & gtc & gtg & gtt & tta & ttc & ttg & ttt
 \end{array} \right]
 \end{array}
 \xrightarrow{x}
 \begin{array}{c}
 \left[\begin{array}{cccccccc}
 3 & 2 & 2 & 2 & 2 & 1 & 1 & 1 \\
 2 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
 2 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
 2 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
 2 & 1 & 1 & 1 & 2 & 1 & 1 & 1 \\
 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0
 \end{array} \right]
 \end{array}$$

$$\begin{array}{c}
 \left[\begin{array}{c|cccccc}
 * & a & b & c & d & e & f \\
 a & a & b & c & d & e & f \\
 b & b & a & e & f & c & d \\
 c & c & f & a & e & d & b \\
 d & d & e & f & a & b & c \\
 e & e & d & b & c & f & a \\
 f & f & c & d & b & a & e
 \end{array} \right]
 \end{array}
 \xrightarrow{x}
 \begin{array}{c}
 \left[\begin{array}{cccccc}
 28 & 24 & 29 & 43 & 37 & 67 \\
 24 & 28 & 37 & 67 & 29 & 43 \\
 29 & 67 & 28 & 37 & 43 & 24 \\
 43 & 37 & 67 & 28 & 24 & 29 \\
 37 & 43 & 24 & 29 & 67 & 28 \\
 67 & 29 & 43 & 24 & 28 & 37
 \end{array} \right]
 \end{array}$$

Preface

These lecture notes cover the content of the short course on Symmetry Studies jointly sponsored by EURANDOM and the Euler Institute for Discrete Mathematics and its Applications (EIDMA), held on the campus of the Eindhoven University of Technology, The Netherlands, in the Spring of 2005.

Chapter 1 is dedicated to identifying the language and the basic components of structured data and symmetries studies. It introduces examples defining and connecting the notions of symmetry, classification and experimentation in the natural sciences.

Chapter 2 is an introduction to the theory of representation of finite groups written within the context of developing the tools and techniques for the analysis of structured data.

Appendix A includes one of the workshops developed for the short course, with selected comments and solutions. A number of computer routines for symbolic logic programs utilized in the text appear in Appendix B. The content of these Notes was abstracted from selected chapters of *Structured Data - An Introduction to the Study of Symmetry in Applications*, by M. Viana, to appear in print elsewhere.

I am indebted to Professor Arjeh M. Cohen (Eindhoven University of Technology) who taught, with great insight and enthusiasm, the algebraic aspects of the course. This event would not have happened without the organizational effort of Mrs. Henny Houben (EIDMA) and the support and coordination of Professors Alessandro Di Bucchianico and Henry Wynn (EURANDOM). Peter van de Ven (EURANDOM) assisted with the current revision of these Notes. Many thanks!

Last, but not least, I acknowledge the friendly and warm atmosphere induced by all participants throughout that week (including those who did not believe that a Friday afternoon session would really happen!). It was great fun!

Marlos A. G. Viana
The University of Illinois at Chicago
USA

February 02, 2006.

Contents

| | |
|---|-----|
| Preface | iii |
| Chapter 1. Symmetry, Classification and the Analysis of Structured Data | 1 |
| 1.1. Introduction | 1 |
| 1.2. Symmetry and classification | 1 |
| 1.3. Symmetry and measurement | 3 |
| 1.4. A data structure induced by a molecular framework | 4 |
| 1.5. A data structure induced by short DNA sequences | 6 |
| 1.6. Symmetries acting on a structure | 8 |
| 1.7. Symmetry and probability laws | 10 |
| 1.8. Symmetry and classification in classical mechanics | 13 |
| 1.9. Canonical decompositions | 17 |
| Appendix -Maxwell-Boltzmann canonical distribution | 20 |
| Further reading | 20 |
| Exercises | 22 |
| Chapter 2. Algebraic Methods for Data Analysis | 29 |
| 2.1. Introduction | 29 |
| 2.2. Permutations, groups and homomorphisms | 32 |
| 2.3. Group actions and orbits | 40 |
| 2.4. Linear representations | 45 |
| 2.5. Reducibility | 52 |
| 2.6. Characters of a representation | 59 |
| 2.7. Schur's Lemma and applications | 60 |
| 2.8. The canonical decomposition | 67 |
| 2.9. The standard decomposition | 70 |
| 2.10. Inference | 72 |
| 2.11. Summary | 84 |
| Further reading | 85 |
| Exercises | 86 |
| Appendix A: Workshop | 89 |
| 2.12. Comments and solutions | 90 |
| Appendix B: Computing Algorithms | 93 |
| Bibliography | 97 |
| Index | 101 |

CHAPTER 1

Symmetry, Classification and the Analysis of Structured Data

Lecture Notes at EURANDOM - EIDMA, The Netherlands
March 14, 2005

Abstracted from M. Viana *Structured Data - An Introduction to the Study of Symmetry in Applications* ©2005 by M. A. G. Viana

1.1. Introduction

George Pólya, in his introduction to mathematics and plausible reasoning, observes that

A great part of the naturalist's work is aimed at describing and classifying the objects that he observes. A good classification is important because it reduces the observable variety to relative few clearly characterized and well ordered types.

Pólya's narrative introduces us directly to the practical aspect of partitioning a large number of objects by exploring certain rules of equivalence among them. This is how symmetry will be understood in the present text: as a set of rules with which we may describe certain commonalities, invariants or regularities among objects or concepts. The classification of crystals, for example, is based on symmetries in their molecular framework.

Included in the naturalist's methods of description is the delicate notion of measuring something on these objects and recording their data, so that the classification of the objects may be related to the classification or partitioning of their corresponding data. Pólya's picture also includes the notion of interpreting, or characterizing, the resulting types of varieties. That is, the naturalist has a better result when he can explain why certain varieties fall into the same type or category.

This chapter is an introduction to the interplay among symmetry, classification and experimental data, which is the driving motive underlying any symmetry study and is often present in the basic sciences. Our purpose here is suggesting that the principles derived from such interplay can lead to novel ways of looking at data, of planning experiments, and, potentially, of facilitating contextual explanation. We will observe the intertwined presence of symmetry, classification and experimental data in a number of classical examples from Chemistry, Biology and Physics. The partition of disciplines, however, is only eventual. In fact, we strongly believe that the reader will benefit from a synthetic reading of the material. Many principles and techniques will repeat across different disciplines, and it is exactly that cross-section of knowledge that constitutes the higher motivation and foundation of the proposed symmetry studies.

1.2. Symmetry and classification

In grade school we were amused (for a little while at least!) by drawings and games with colorful patterns or motifs, such as



repeated periodically along an horizontal straight line. These bands can be classified according to the distinct generating rules for capturing invariance, or regularity, such as horizontal translations, vertical and horizontal line reflections, point reflections and rotations. Wallpaper, textile and

tapestry designs explore these rules of invariance, or symmetry in two dimensions, adding the technical difficulty of artistically and graphically designing these repeating motifs within the finite boundaries of the work.

These planar transformations can be used to classify the fonts of the 26 capital letters A, B, . . . , Z used in this typesetting into the five classes

$$\{A, M, T, U, V, W, Y\}, \{B, C, D, E\}, \{N, S, Z\}, \{H, I, O, X\}, \{F, G, J, K, L, P, Q, R\}$$

according to the set of symmetries best describing them. The fonts in the set $\{H, I, O, X\}$, for example, are characterized by having horizontal (h), vertical (v) and point (o) symmetries. The letter H and its transformed image $v(H)$ under the vertical reflection coincide, that is $v(H) = H$. We then say that the object H *has the symmetry of* the transformation v. Similarly, H has the symmetry of the horizontal reflection, that is $h(H) = H$. Therefore, H must remain invariant under the iterated action vh of these two symmetry transformations, that is $(vh)(H) = v(h(H)) = v(H) = H$. Because the resulting composition of v and h is a point reflection ($vh = o$), we learn that H has the symmetries of v, h and o. Clearly, in addition, H and any other letter remain invariant when the identity symmetry transformation (indicated by 1), is applied to it. It turns out that the set

$$S = \{1, v, h, o\}$$

of symmetry transformations, together with the operation of composition of transformations satisfy the multiplication table

$$(1.1) \quad S: \begin{array}{c|cccc} * & 1 & v & h & o \\ \hline 1 & 1 & v & h & o \\ v & v & 1 & o & h \\ h & h & o & 1 & v \\ o & o & h & v & 1 \end{array},$$

and confer to the pair $(S, *)$ the algebraic properties of a *finite group*, briefly: $t * t'$ is in S for all t, t' in S, the existence of the identity element 1 satisfying $t * 1 = 1 * t = t$, the inverse $t^{-1} \in S$ of $t \in S$ satisfying $t * t^{-1} = t^{-1} * t = 1$, and the associativity of $*$.

We then say that the set $\{H, I, O, X\}$ has the symmetry of $S = \{1, v, h, o\}$. On the other hand, the set $\{F, G, J, K, L, P, Q, R\}$ has the symmetry of the identity transformation alone, whereas $\{A, M, T, U, V, W, Y\}$ has the symmetry of $\{1, v\}$.

Table (1.1) gives us the opportunity of introducing from the start the convention that the (algebraic) product $\sigma * \tau$ of the row element σ with the column element τ is to be understood as the (analytic) composition $\tau\sigma$ of the two transformations, that is $\sigma * \tau = \tau\sigma$. This convention adopted from now on is meant to accommodate both interpretations as much as possible.

The linear transformations

$$\{e_1, e_2\} \xrightarrow{r(t)} \{t(e_1), t(e_2)\}, \quad t \in S,$$

defined when the symmetries t in S are applied to the canonical basis $\{e_1, e_2\}$ of $\mathcal{V} = \mathbb{R}^2$, namely,

$$r(1) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad r(v) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad r(h) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad r(o) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix},$$

provide an example of a linear representation of $(S, *)$ in \mathcal{V} . Note, for example, that

$$r(v * o) = r(h) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = r(v)r(o),$$

and that in general,

$$r(t * t') = r(t)r(t') \quad \text{for all } t, t' \in S.$$

This is the homomorphic property between multiplication in S and multiplication of linear transformations in \mathcal{V} , which is characteristic of such representations, to be studied in detail in Chapter

2. Similarly, the transformations

$$\{e_1, e_v, e_h, e_o\} \xrightarrow{\rho(t)} \{e_{t*1}, e_{t*v}, e_{t*h}, e_{t*o}\}, \quad t \in S,$$

define a linear representation

$$\rho(1) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \rho(v) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \rho(h) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad \rho(o) = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

of $(S, *)$ in $\mathcal{V} = \mathbb{R}^4$.

The important point here for the data analyst is the fact that these linear representations transfer the symmetries described by S to the vector space \mathcal{V} of potential measurements.

1.3. Symmetry and measurement

Our text is limited to classical, structural measurements which are independent of the past history of the system. Even at the particle level, measurements such as the atomic binding energy, spin, magnetic moment or the difference in energy between the various stationary states are structural properties. The position or the momentum of a particle, or a component of its angular momentum are not structural. We are restricted to stationary measurements, which are distinct from measurement of a probability. As a consequence of this assumption, we do not need to include in our analysis of the data the interface between the observer and the observed, and the recorded data coincide with the measurement.

Here is an example of assigning data to symmetries: The left hand matrix in (1.2) shows a Sloan Chart developed for use in the Early Treatment Diabetic Retinopathy Study. Adjacent to the chart is a table with the symmetry transformations of the individual fonts and their estimated probability of being incorrectly identified. The Sloan Charts and the study of the individual Sloan fonts appear in Ferris 3rd, Freidlin, Kassoff, Green and Milton (1993, Table 5).

| (1.2) | C O H Z V | <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="padding: 5px;">Letter</th> <th style="padding: 5px;">Symmetries</th> <th style="padding: 5px;">Difficulty</th> </tr> </thead> <tbody> <tr><td style="padding: 5px;">Z</td><td style="padding: 5px;">1, o</td><td style="padding: 5px;">0.844</td></tr> <tr><td style="padding: 5px;">N</td><td style="padding: 5px;">1, o</td><td style="padding: 5px;">0.774</td></tr> <tr><td style="padding: 5px;">H</td><td style="padding: 5px;">1, o, v, h</td><td style="padding: 5px;">0.688</td></tr> <tr><td style="padding: 5px;">V</td><td style="padding: 5px;">1, v</td><td style="padding: 5px;">0.636</td></tr> <tr><td style="padding: 5px;">R</td><td style="padding: 5px;">1</td><td style="padding: 5px;">0.622</td></tr> <tr><td style="padding: 5px;">K</td><td style="padding: 5px;">1</td><td style="padding: 5px;">0.609</td></tr> <tr><td style="padding: 5px;">D</td><td style="padding: 5px;">1, h</td><td style="padding: 5px;">0.556</td></tr> <tr><td style="padding: 5px;">S</td><td style="padding: 5px;">1, o</td><td style="padding: 5px;">0.516</td></tr> <tr><td style="padding: 5px;">O</td><td style="padding: 5px;">1, o, v, h</td><td style="padding: 5px;">0.470</td></tr> <tr><td style="padding: 5px;">C</td><td style="padding: 5px;">1, h</td><td style="padding: 5px;">0.393</td></tr> </tbody> </table> | Letter | Symmetries | Difficulty | Z | 1, o | 0.844 | N | 1, o | 0.774 | H | 1, o, v, h | 0.688 | V | 1, v | 0.636 | R | 1 | 0.622 | K | 1 | 0.609 | D | 1, h | 0.556 | S | 1, o | 0.516 | O | 1, o, v, h | 0.470 | C | 1, h | 0.393 |
|-----------|-----------|--|------------|------------|------------|---|------|-------|---|------|-------|---|------------|-------|---|------|-------|---|---|-------|---|---|-------|---|------|-------|---|------|-------|---|------------|-------|---|------|-------|
| | Letter | | Symmetries | Difficulty | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | Z | | 1, o | 0.844 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | N | | 1, o | 0.774 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | H | | 1, o, v, h | 0.688 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | V | | 1, v | 0.636 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | R | | 1 | 0.622 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | K | | 1 | 0.609 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | D | | 1, h | 0.556 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | S | | 1, o | 0.516 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | O | | 1, o, v, h | 0.470 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | C | | 1, h | 0.393 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | S Z N D C | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | V K C N R | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | K C R H N | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | Z K D V C | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H V O R K | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| R H S O N | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| K S V R H | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H N K C D | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| N D V K O | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| D H O S Z | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| V R N D O | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| C Z H K S | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| O R Z S K | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Fix the first line (COHZV) in the chart and to each symmetry transformation t in $S = \{1, v, h, o\}$ assign the number $x(t)$ of letters in the selected line with the symmetry of t . Thus, $x(1) = 5$ and $x(v) = x(h) = x(o) = 3$, so that

$$x' = (5, 3, 3, 3)$$

is an example of data indexed by the S , and a point in the vector space $\mathcal{V} = \mathbb{R}^4$. The representation ρ of $(S, *)$ in \mathcal{V} introduced above then connects the symmetries in S with the data vector $x \in \mathcal{V}$.

It is precisely this connection that will facilitate, and often determine, the analysis of the data x . Clearly, should the Sloan lines be selected at random, then x is a random vector subject to statistical inference. The vector x is an example of data indexed by a particular structure- or, simply, an example of structured data.

We will learn from the theory developed in the next chapter that the appropriate vector space for displaying the summaries of any outcome associated with each line in the chart, such as its mean line difficulty or contrast sensitivity, is determined by the invariants

$$x(1) + x(o) + x(v) + x(h), \quad x(1) + x(o) - x(v) - x(h),$$

$$x(1) + x(v) - x(o) - x(h), \quad x(1) + x(h) - x(o) - x(v).$$

1.4. A data structure induced by a molecular framework

In Chemistry, symmetry is often the link between the determination of certain structural properties of molecules and their specific measurements such as the infra-red spectra, ultra-violet spectra, dipole moments and optical activity. That is, symmetry is relevant to characterizing regularities in experimental data. The fact that a symmetry operation, when applied to the framework of a molecule moves it in a way that its final position is physically indistinguishable from its initial position, implicitly says that the physical data extracted from the structure remain constant after each attempt to alter the initial structure [See Bishop (1973, p. 10)]. This *immunity to change*, characterized by certain invariant physical properties, can be used by the chemist to classify elementary molecules. No matter how we turn the hydrogen molecule about a line of axial symmetry joining the two nuclei, its electron density data read the same and reflect the fact that the bond in the hydrogen atom has cylindrical symmetry. This illustrates, literally, Rosen's (1995) view of symmetry as the object's immunity to change. Other transformations may, however, alter the position of the molecular structure in space, so that the original compound and the transformed one have the same molecular formula and different structure position. Such isomeric (consisting of the same parts) compounds are called *stereoisomers*.

A symmetry argument is sufficient to classify simple molecules according to their measurable capacity of rotating the plane of polarization of a ray of light (optical activity). The fact that two molecules are mirror-image symmetric and not superimposable characterizes their optical activity data. This property, called *chirality*, or handedness, is observable in many other aspects of nature. Typically, these molecules have many similar measurable physical properties, such as boiling point, melting point or index of refraction. The chemist tells them apart by their response to polarized light. One molecule may rotate the plane of polarization to the right whereas the fellow isomer may rotate it to the left. Chiral is the Greek word for hands- which do not have superimposable mirror images- hands are chiral objects. Again, symmetry establishes practical links between the object and the experimental data and, as we will see in the sequence, these relationship can be useful to the data analyst.

The symmetry transformations that leave the stable configuration of a molecule physically indistinguishable, are generally known as *point groups*. The name indicates that at least one point in the molecular framework remains fixed. For example, the following transformations in \mathbb{R}^3 are represented here with the standard notation used by chemists:

- (1) E: the identity operator;
- (2) C_2 : a rotation by 180 deg around the z-axis;
- (3) i: an *inversion* or point reflection through the origin (0, 0, 0);
- (4) σ_H : a reflection on the xy-plane.

These transformations, $\mathcal{C}_{2h} = \{E, C_2, i, \sigma_h\}$, together with the operation of composition of transformations, define the point group \mathcal{C}_{2h} . Its multiplication table is given by

$$(1.3) \quad \mathcal{C}_{2h} : \begin{array}{c|cccc} * & E & C_2 & i & \sigma_h \\ \hline E & E & C_2 & i & \sigma_h \\ C_2 & C_2 & E & \sigma_h & i \\ i & i & \sigma_h & E & C_2 \\ \hline \sigma_h & \sigma_h & i & C_2 & E \end{array}.$$

For example, the 180 deg rotation (C_2) around the z-axis followed by an inversion (i) through the origin is equivalent to a reflection on the xy-plane, that is, $iC_2 = \sigma_h$. The planar structure of a dichloroethene $C_2H_2Cl_2$ -*trans* molecule is among the molecules characterized by the symmetries of the point group \mathcal{C}_{2h} .

The molecular framework of the dichloroethene molecule can be used as a data structure. To see this, consider a rectangular parallelepiped with vertices $\{(\pm 2, \pm 1, \pm 1)\}$ expressed as the set of labels

$$V = \{ \text{abb}, \text{abB}, \text{aBb}, \text{aBB}, \text{Abb}, \text{AbB}, \text{ABb}, \text{ABB} \}.$$

For example, AbB is the label for the point $(-2, 1, -1)$. Note that the labels in V transform according to permutations in V under the action of each symmetry transformations of \mathcal{C}_{2h} , thus determining a linear representation

$$V \xrightarrow{\rho(t)} \{t(s); s \in V\}, \quad t \in \mathcal{C}_{2h}$$

of \mathcal{C}_{2h} in $\mathcal{V} = \mathbb{R}^8$. Because the set V is interpreted as set of labels for experimental conditions, we say that V is a *structure*, and that the vector

$$x = (x(s))_{s \in V},$$

of experimental responses and a point in the data space \mathcal{V} is the corresponding *structured data*.

The algebraic tools to be introduced later on in the sequence will identify a reduction of the data space \mathcal{V} into 4 two-dimensional subspaces $\mathcal{V}_1, \dots, \mathcal{V}_4$ of \mathcal{V} , with corresponding invariants

$$x(\alpha\text{bb}) + x(\alpha\text{bB}) + x(\alpha\text{Bb}) + x(\alpha\text{BB}), \quad x(\alpha\text{bb}) - x(\alpha\text{bB}) - x(\alpha\text{Bb}) + x(\alpha\text{BB}),$$

$$x(\alpha\text{bb}) - x(\alpha\text{bB}) + x(\alpha\text{Bb}) - x(\alpha\text{BB}), \quad x(\alpha\text{bb}) + x(\alpha\text{bB}) - x(\alpha\text{Bb}) - x(\alpha\text{BB}),$$

where $\alpha = a, A$. This is only one of infinitely many such reductions, all isomorphic to each other. Data analysts may recognize these invariants as two copies of a 2-factor factorial experiment, indexed by the two regular faces, $(2, \pm 1, \pm 1)$ and $(-2, \pm 1, \pm 1)$, of the parallelepiped, and distinguished by the transposition (aA) of a and A. What makes the analogy work is the fact that the molecular framework of the molecule is invariant under its point group. The connection between point groups in molecular chemistry and factorial experiments thus suggests a strategy for discovering additional forms of experimentation and plausible descriptors of the data based on the derived invariants.

We conclude this section with the observation that

$$\{e_1, e_2, e_3\} \xrightarrow{R(t)} \{t(e_1), t(e_2), t(e_3)\}, \quad t \in \mathcal{C}_{2h}$$

defines a linear representation

$$R(E) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad R(C_2) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad R(i) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \quad R(\sigma_h) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

of \mathcal{C}_{2h} in $\mathcal{V} = \mathbb{R}^3$. Moreover, the reader may want to compare the two multiplication tables 1.1 and 1.3 and conclude that they are particular realizations of an *abstract* multiplication table

$$(1.4) \quad G : \begin{array}{c|cccc} * & 1 & a & b & c \\ \hline 1 & 1 & a & b & c \\ a & a & 1 & c & b \\ b & b & c & 1 & a \\ c & c & b & a & 1 \end{array},$$

defined among the symbols $G = \{1, a, b, c, d\}$. That is, they correspond to each other *isomorphically*. Consequently, the representations r and ρ of $(S, *)$ and ϱ and R of $(\mathcal{C}_{2h}, *)$ are representations of the group G , of dimensions 2, 4, 8 and 3 respectively.

1.5. A data structure induced by short DNA sequences

A biological sequence is a finite string of symbols from a finite alphabet (\mathcal{A}) of residues, such as the linear string

[ctccttgggatattgatgatctgtagtgtctacagaaaaattgtgggtcacagtctattat](#),

in which the symbols are letters in the alphabet $\mathcal{A} = \{a, g, t, c\}$. Here the symbols represent adenine (a), guanine (g), thymine (t) and cytosine (c) molecules in DNA (deoxyribonucleic acid) sequences. The adjacency of two symbols in the linear string means that the two molecules are chemically binded to each other. There are many more common alphabets, representing

- the nucleotides adenine (a), guanine (g), cytosine (c), uracil (u) in RNA (ribonucleic acid) sequences: $\mathcal{A} = \{a, g, t, u\}$;
- the classes $u = \{a, g\}$ of purine and $y = \{c, t\}$ of pyrimidine residues: $\mathcal{A} = \{u, y\}$,

or the larger class of amino acids in protein sequences. The length of *global* or complete sequences, in base pairs, ranges from 10^3 (single-stranded virus) to 10^9 (mammals). The standard code then translates DNA triplets into specific amino acids. For example, the set V of DNA triplets constitutes an example of a simple set of labels or indices for experimental and analytical studies in molecular biology.

Similarly to the data structure induced by the molecular framework fixed by the point group \mathcal{C}_{2h} , the structure

$$V = \{ttt, ttc, tta, \dots, gga, ggg\}$$

defined by these 64 simple sequences (s) in length of three written with a four-letter alphabet $\mathcal{A} = \{a, g, c, t\}$, is a structure indexing potential molecular constructs or measurements, $x(s)$, such as the triplet's molecular weight or its frequency of occurrence in a larger, reference sequence. Consequently, then,

$$x = (x(s))_{s \in V} = (x(ttt), x(ttc), x(tta), \dots, x(gga), x(ggg)),$$

is a data vector indexed by the structure V , or, simply, a structured data. The structure may be amalgamated, for example, by rewriting each word with the shorter alphabet $\mathcal{A} = \{u, y\}$ of purine-pyrimidine residues. The new structure

$$V = \{yyy, yyu, \dots, uuu\}$$

of triplets of purine-pyrimidines has $2^4 = 16$ points or labels, and

$$x = (x(yyy), x(yyu), \dots, x(uuu)),$$

are the corresponding structured data.

Here is another simple structured data in which is the structure is the set

$$V = \{a, g, c, t\} \times \{a, g, c, t\}$$

of ordered pairs of DNA nucleotides. It has $4 \times 4 = 16$ points in it.

Given two local DNA sequences

$$I = \text{ttttcgctatggaacctgggatcttttagtttgaatgggagagccattccgcctggaaaaattagataaggtaag},$$

$$J = \text{tttccgcatggaacctggaatagttgctcaaaagtgaggcaaccgcttaggtttgaaaaattagataaggcg},$$

we measure, in each point (i, j) of V , the frequency $x(i, j)$ with which the residue i in the sequence I aligns with the residue j in sequence J along the two sequences. Here are the resulting structured data:

| | | | | |
|-----------------|----|---|----|----|
| $i \setminus j$ | a | c | g | t |
| a | 17 | 2 | 4 | 0 |
| c | 3 | 5 | 1 | 3 |
| g | 3 | 1 | 15 | 1 |
| t | 1 | 4 | 2 | 15 |

We observe, in addition, that a sequence in length of ℓ is a function or mapping

$$s : L \rightarrow \mathcal{A},$$

where $L = \{1, 2, \dots, \ell\}$ is the set for the ordered positions in which the residues in the alphabet \mathcal{A} are located. The set of all such mappings is indicated by \mathcal{A}^L . For example,

$$(1.5) \quad V = \left[\begin{array}{cccccccc} \text{aaa} & \text{ggg} & \text{ccc} & \text{ttt} & \text{aag} & \text{aac} & \text{aat} & \text{gga} \\ \text{ggc} & \text{ggt} & \text{cca} & \text{ccg} & \text{cct} & \text{tta} & \text{ttg} & \text{ttc} \\ \text{aga} & \text{aca} & \text{ata} & \text{gag} & \text{gcg} & \text{gtg} & \text{cac} & \text{cgc} \\ \text{ctc} & \text{tat} & \text{tgt} & \text{tct} & \text{gaa} & \text{caa} & \text{taa} & \text{agg} \\ \text{cgg} & \text{tgg} & \text{acc} & \text{gcc} & \text{tcc} & \text{att} & \text{gtt} & \text{ctt} \\ \text{agc} & \text{gac} & \text{cga} & \text{acg} & \text{gca} & \text{cag} & \text{agt} & \text{atg} \\ \text{tga} & \text{gat} & \text{gta} & \text{tag} & \text{act} & \text{atc} & \text{tca} & \text{cat} \\ \text{cta} & \text{tac} & \text{get} & \text{gtc} & \text{tcg} & \text{cgt} & \text{ctg} & \text{tgc} \end{array} \right]$$

is the set of all mappings $s : \{1, 2, 3\} \rightarrow \{a, g, t, c\}$ and a structure with which potential measurements can be indexed.

Indicating by $|\mathcal{A}|$ the number of elements in the set \mathcal{A} , we refer to the $|\mathcal{A}|^\ell$ sequences in V as $|\mathcal{A}|$ -sequences in length of ℓ . Every binary sequence in length of four, with $\mathcal{A} = \{u, y\}$, is a mapping

$$s : \{1, 2, 3, 4\} \rightarrow \{u, y\}.$$

The structure V has 16 points, namely,

$$(1.6) \quad V = \left[\begin{array}{c|cccccccccccccccc} s & 1 & 16 & 15 & 14 & 12 & 8 & 13 & 11 & 7 & 10 & 6 & 4 & 9 & 5 & 3 & 2 \\ \hline s(1) & y & u & y & u & u & u & y & y & y & u & u & u & y & y & y & u \\ s(2) & y & u & u & y & u & u & y & u & u & y & y & u & y & y & u & y \\ s(3) & y & u & u & u & y & u & u & y & u & y & u & y & y & u & y & y \\ s(4) & y & u & u & u & u & y & u & u & y & u & y & y & u & y & y & y \end{array} \right].$$

The numbers in the first row are reference labels for each point in V , based on the fact that the c^ℓ elements of $V = C^L$ can be indexed by the (base c) representation

$$(1.7) \quad I(s) = 1 + \sum_{j=1}^{\ell} (s(j) - 1)c^{j-1}, \quad s \in C^L,$$

with $C = \{1, \dots, c\}$ and $L = \{1, \dots, \ell\}$.

1.6. Symmetries acting on a structure

In the previous sections we have illustrated the classification of objects using symmetry transformations and the representation of these symmetries as linear transformations in the vector space for the data observed on those objects. In this section we go one step further and reduce the structure into sets of symmetrically equivalent points.

Indicate by S_ℓ the set of all permutations (one-to-one mappings) τ over a set with $\ell < \infty$ elements. For example, the set S_3 of all permutations of 3 symbols includes the identity transformation, indicated here by 1; three transpositions,

$$(12) = \begin{bmatrix} 1 \rightarrow 2 \\ 2 \rightarrow 1 \\ 3 \rightarrow 3 \end{bmatrix}, \quad (13) = \begin{bmatrix} 1 \rightarrow 3 \\ 2 \rightarrow 2 \\ 3 \rightarrow 1 \end{bmatrix}, \quad (23) = \begin{bmatrix} 1 \rightarrow 1 \\ 2 \rightarrow 3 \\ 3 \rightarrow 2 \end{bmatrix},$$

and two cyclic permutations,

$$(123) = \begin{bmatrix} 1 \rightarrow 2 \\ 2 \rightarrow 3 \\ 3 \rightarrow 1 \end{bmatrix}, \quad (132) = \begin{bmatrix} 1 \rightarrow 3 \\ 2 \rightarrow 1 \\ 3 \rightarrow 2 \end{bmatrix}.$$

In the notation introduced above, each permutation is written as a sequence of cycles, such as (132) indicating $1 \rightarrow 3 \rightarrow 2 \rightarrow 1$. This is a cycle of length (or order) 3. In the notation, cycles of length one are omitted. A transposition is a cycle of length 2. In summary, we write

$$S_3 = \{1, (12), (13), (23), (123), (132)\}.$$

A subset of S_3 is the set C_3 of all cyclic permutations in length of 3, that is, $C_3 = \{1, (123), (132)\}$. The reader may want to evaluate the multiplication table of $(S_3, *)$, in analogy to Tables (1.1), (1.3) or (1.4) and verify that $(S_3, *)$ is finite group. The order, or number of elements, of S_3 is 6. The order of C_3 is 3. The reader may also identify the transformations in S_3 with the symmetry transformations of a regular triangle with vertices indexed by 1, 2, 3.

Given a sequence $s \in C^L$, a permutation τ in S_ℓ and a permutation $\sigma \in S_c$, then the composites

$$s\tau^{-1} : L \xrightarrow{\tau^{-1}} L \xrightarrow{s} C, \quad \text{and} \quad s\sigma : L \xrightarrow{s} C \xrightarrow{\sigma} C$$

are also a sequences in length of ℓ in V (using the inverse permutation will be justified later). The composition $s\tau^{-1}$ is called a *composition on the left*, whereas $s\sigma$ is a *composition on the right*. If, say,

$$\tau = \begin{bmatrix} 1 \rightarrow 2 \\ 2 \rightarrow 3 \\ 3 \rightarrow 4 \\ 4 \rightarrow 1 \end{bmatrix} \quad \text{and} \quad s = \begin{bmatrix} 1 \rightarrow a \\ 2 \rightarrow a \\ 3 \rightarrow g \\ 4 \rightarrow c \end{bmatrix}, \quad \text{then,} \quad s\tau^{-1} = \begin{bmatrix} 1 \rightarrow c \\ 2 \rightarrow a \\ 3 \rightarrow a \\ 4 \rightarrow g \end{bmatrix}.$$

The fact that the composition $s\tau^{-1}$ of a mapping $s \in V$ and a permutation $\tau \in S_\ell$ results in another mapping in V leads to the construction of the sets of all mappings sharing the symmetries defined by a given problem. These sets are called *symmetry orbits*, or simply orbits. Similarly, two mappings are classified as equivalent (\sim) when one is obtained from the other by composing it with a symmetry permutation, that is, $s \sim f \iff f = s\tau^{-1}$. Consequently, equivalent mappings define the same orbit. We write

$$\mathcal{O}_s = \{s\tau^{-1}; \tau \in S_\ell\}$$

to indicate the permutation orbit of a mapping s resulting from composing it with S_ℓ on the left. Similarly, $\mathcal{O}_s = \{s\sigma; \sigma \in S_c\}$ is a permutation orbit, resulting from composing s with $\sigma \in S_c$ on the right.

When the symmetries of interest are the cyclic permutations, we obtain the corresponding cyclic orbits. For example, starting with the sequence cgg in length of three and composing on the left with all three cyclic permutations in C_3 we obtain the orbit $\mathcal{O}_{cgg} = \{cgg, gcg, ggc\}$.

Similarly, starting with the sequence uuyuuy in length of six and composing on the left with all six cyclic permutations in C_6 we obtain the orbit $\mathcal{O}_{uuyuuy} = \{uuyuuy, yuuyuu, uyuyuy\}$.

1.6.1. Permutation orbits for binary sequences in length of four. The mapping space V of all binary sequences in length of four has $2^4 = 16$ points, each representing one sequence, as shown in Matrix (1.8). The numbers in the first row are identifiers for each sequence generated from (1.7).

$$(1.8) \quad \left[\begin{array}{c|cccccccccccccccc} s & 1 & 16 & 15 & 14 & 12 & 8 & 13 & 11 & 7 & 10 & 6 & 4 & 9 & 5 & 3 & 2 \\ \hline s(1) & y & u & y & u & u & u & y & y & y & u & u & u & y & y & y & u \\ s(2) & y & u & u & y & u & u & y & u & u & y & y & u & y & y & u & y \\ s(3) & y & u & u & u & y & u & u & y & u & y & u & y & y & u & y & y \\ s(4) & y & u & u & u & u & y & u & u & y & u & y & y & u & y & y & y \end{array} \right].$$

Consider the left composition $(s\tau^{-1})$ of sequences in V with the symmetries in S_4 . The group S_4 has 6 transpositions, 3 elements of order 2, 8 elements of order 3 and 6 elements of order 4. These permutations are indicated in the first column of Matrix (1.9).

$$(1.9) \quad \begin{array}{c|cccccccccccccccc} S_4 \backslash s & 1 & 16 & 15 & 14 & 12 & 8 & 13 & 11 & 7 & 10 & 6 & 4 & 9 & 5 & 3 & 2 \\ \hline 1 & 1 & 16 & 15 & 14 & 12 & 8 & 13 & 11 & 7 & 10 & 6 & 4 & 9 & 5 & 3 & 2 \\ (34) & 1 & 16 & 15 & 14 & 8 & 12 & 13 & 7 & 11 & 6 & 10 & 4 & 5 & 9 & 3 & 2 \\ (23) & 1 & 16 & 15 & 12 & 14 & 8 & 11 & 13 & 7 & 10 & 4 & 6 & 9 & 3 & 5 & 2 \\ (24) & 1 & 16 & 15 & 8 & 12 & 14 & 7 & 11 & 13 & 4 & 6 & 10 & 3 & 5 & 9 & 2 \\ (12) & 1 & 16 & 14 & 15 & 12 & 8 & 13 & 10 & 6 & 11 & 7 & 4 & 9 & 5 & 2 & 3 \\ (13) & 1 & 16 & 12 & 14 & 15 & 8 & 10 & 11 & 4 & 13 & 6 & 7 & 9 & 2 & 3 & 5 \\ (14) & 1 & 16 & 8 & 14 & 12 & 15 & 6 & 4 & 7 & 10 & 13 & 11 & 2 & 5 & 3 & 9 \\ \hline (234) & 1 & 16 & 15 & 12 & 8 & 14 & 11 & 7 & 13 & 4 & 10 & 6 & 3 & 9 & 5 & 2 \\ (243) & 1 & 16 & 15 & 8 & 14 & 12 & 7 & 13 & 11 & 6 & 4 & 10 & 5 & 3 & 9 & 2 \\ (123) & 1 & 16 & 14 & 12 & 15 & 8 & 10 & 13 & 6 & 11 & 4 & 7 & 9 & 2 & 5 & 3 \\ (124) & 1 & 16 & 14 & 8 & 12 & 15 & 6 & 10 & 13 & 4 & 7 & 11 & 2 & 5 & 9 & 3 \\ (132) & 1 & 16 & 12 & 15 & 14 & 8 & 11 & 10 & 4 & 13 & 7 & 6 & 9 & 3 & 2 & 5 \\ (134) & 1 & 16 & 12 & 14 & 8 & 15 & 10 & 4 & 11 & 6 & 13 & 7 & 2 & 9 & 3 & 5 \\ (142) & 1 & 16 & 8 & 15 & 12 & 14 & 7 & 4 & 6 & 11 & 13 & 10 & 3 & 5 & 2 & 9 \\ (143) & 1 & 16 & 8 & 14 & 15 & 12 & 6 & 7 & 4 & 13 & 10 & 11 & 5 & 2 & 3 & 9 \\ \hline (12)(34) & 1 & 16 & 14 & 15 & 8 & 12 & 13 & 6 & 10 & 7 & 11 & 4 & 5 & 9 & 2 & 3 \\ (13)(24) & 1 & 16 & 12 & 8 & 15 & 14 & 4 & 11 & 10 & 7 & 6 & 13 & 3 & 2 & 9 & 5 \\ (14)(23) & 1 & 16 & 8 & 12 & 14 & 15 & 4 & 6 & 7 & 10 & 11 & 13 & 2 & 3 & 5 & 9 \\ \hline (1234) & 1 & 16 & 14 & 12 & 8 & 15 & 10 & 6 & 13 & 4 & 11 & 7 & 2 & 9 & 5 & 3 \\ (1243) & 1 & 16 & 14 & 8 & 15 & 12 & 6 & 13 & 10 & 7 & 4 & 11 & 5 & 2 & 9 & 3 \\ (1324) & 1 & 16 & 12 & 8 & 14 & 15 & 4 & 10 & 11 & 6 & 7 & 13 & 2 & 3 & 9 & 5 \\ (1342) & 1 & 16 & 12 & 15 & 8 & 14 & 11 & 4 & 10 & 7 & 13 & 6 & 3 & 9 & 2 & 5 \\ (1432) & 1 & 16 & 8 & 15 & 14 & 12 & 7 & 6 & 4 & 13 & 11 & 10 & 5 & 3 & 2 & 9 \\ (1423) & 1 & 16 & 8 & 12 & 15 & 14 & 4 & 7 & 6 & 11 & 10 & 13 & 3 & 2 & 5 & 9 \end{array}$$

The sequences $s \in V$ are identified in the first row by their labels shown in Matrix (1.8). The resulting compositions $s\tau^{-1}$ are shown in the adjacent columns. For example, if

$$\tau = (1234) = \begin{bmatrix} 1 \rightarrow 2 \\ 2 \rightarrow 3 \\ 3 \rightarrow 4 \\ 4 \rightarrow 1 \end{bmatrix}, \quad \text{and } s = \begin{bmatrix} 1 \rightarrow u \\ 2 \rightarrow u \\ 3 \rightarrow y \\ 4 \rightarrow u \end{bmatrix}, \quad \text{then } s\tau^{-1} = \begin{bmatrix} 1 \rightarrow u \\ 2 \rightarrow u \\ 3 \rightarrow u \\ 4 \rightarrow y \end{bmatrix},$$

so that the composition with $\tau^{-1} = (1432)$ takes the sequence uuyu (label 12) into the sequence uuyu (label 8). In particular, these two sequences are in the same orbit.

The resulting orbits (indicating the sequences by their labels) may be expressed as

$$\begin{aligned}\mathcal{O}_0 &= \{1\}, \\ \mathcal{O}_1 &= \{9, 5, 3, 2\}, \\ \mathcal{O}_2 &= \{13, 11, 7, 10, 6, 4\}, \\ \mathcal{O}_3 &= \{15, 14, 12, 8\}, \\ \mathcal{O}_4 &= \{16\},\end{aligned}$$

so that

$$(1.10) \quad V = \mathcal{O}_0 \cup \mathcal{O}_1 \cup \mathcal{O}_2 \cup \mathcal{O}_3 \cup \mathcal{O}_4$$

forms a disjoint partition of V . We observe that the orbit \mathcal{O}_k has exactly

$$|\mathcal{O}_k| = \binom{\ell}{k}$$

elements and is characterized by the number of purines (u) in the sequences, that is,

$$\mathcal{O}_k = \{s \in V; |s^{-1}(u)| = k\}, \quad k = 0, \dots, 4.$$

The orbit volume $|\mathcal{O}|$ is a *symmetry invariant* in the resulting reduction. It stays constant regardless of how the purine-pyrimidine positions in the sequences are shuffled.

The reader may identify, in this example, all the steps described in Pólya's reasoning, introduced earlier on in the chapter, namely: description, classification and interpretation of the objects of interest.

A good classification is important because it reduces the observable variety to relative few clearly characterized and well ordered types.

The effect of composing V with S_ℓ on the left is that of removing the order of the positions—equivalently, any two sequences are then equivalent, similar or indistinguishable, when they differ only by reordering the position of the letters or residues. As a result, we obtain the space called *quotient space*, in which the elements are the resulting 5 permutation orbits $\mathcal{O}_0, \mathcal{O}_1, \dots, \mathcal{O}_4$. These orbits are characterized by the number of, say, purines. That is, orbit \mathcal{O}_i is composed of those sequences with exactly i purines in it.

The remarkable aspect of this simple example, and its consequences for the planning and analysis of experimental data, is the varied structural classifications that can be obtained from the same initial set of labels by introducing different groups of symmetry and different actions. Permutation groups can act on the set of positions, $L = \{1, 2, 3, 4\}$, and on the alphabet $\mathcal{A} = \{u, y\}$ of residues. To each subgroup of S_4 acting on the set $L = \{1, 2, 3, 4\}$ of positions, a new partition of V can be obtained, with a corresponding new partition in the data space. These techniques will be fully explored in the sequence.

1.7. Symmetry and probability laws

Consider again the structure V of binary sequences in length of ℓ and let P indicate a probability model in the space V , where a group G of symmetries is identified. We say that P has the symmetry of the group G if P is constant (uniform) over each one of the orbits of V . For example, if

$$(1.11) \quad P(s) = P(s\tau^{-1})$$

for all sequences s in V and permutations τ in $G = S_\ell$, then the probability law P should be constant in the position-symmetry orbits. Because s is now a random variable, the purine-pyrimidine levels

$$(\text{number of purines, number of pyrimidines}) = (i, \ell - i)$$

are also random, and consequently, the probability laws

$$(1.12) \quad \mathcal{L}_i = \left(\frac{i}{\ell}, \frac{\ell - i}{\ell}\right), \quad i = 0, 1, \dots, \ell,$$

associated with the orbits described in (1.10) are also random. Here are the possible probability laws for purine-pyrimidine levels from binary sequences in length of four:

$$\mathcal{L}_0 = (0, 1), \quad \mathcal{L}_1 = \left(\frac{1}{4}, \frac{3}{4}\right), \quad \mathcal{L}_2 = \left(\frac{2}{4}, \frac{2}{4}\right), \quad \mathcal{L}_3 = \left(\frac{3}{4}, \frac{1}{4}\right), \quad \mathcal{L}_4 = (1, 0).$$

The likelihood of each law is therefore determined by the probability of seeing a sequence which is associated with the law- because all sequences in the orbit \mathcal{O}_i lead to the law \mathcal{L}_i and conversely, we see that \mathcal{L}_i occurs with probability $P(\mathcal{O}_i)$; shortly,

$$\text{Probability of law } \mathcal{L}_i = P(\mathcal{O}_i).$$

Clearly, if the law P is such that all sequences are equally likely (P is said to be uniform), then condition (1.11) is satisfied and

$$(1.13) \quad \text{Probability of law } \mathcal{L}_i = P(\mathcal{O}_i) = \frac{|\mathcal{O}_i|}{|V|} = \frac{\binom{\ell}{i}}{|V|}.$$

We have, for binary sequences in length of four,

$$P(\mathcal{O}_0) = \frac{1}{16}, \quad P(\mathcal{O}_1) = \frac{4}{16}, \quad P(\mathcal{O}_2) = \frac{6}{16}, \quad P(\mathcal{O}_3) = \frac{4}{16}, \quad P(\mathcal{O}_4) = \frac{1}{16},$$

so that the most likely distribution of purine-pyrimidine levels, under uniformly distributed sequences in V , is

$$\mathcal{L}_2 = \left(\frac{1}{2}, \frac{1}{2}\right).$$

Example 1.7.1 (Four-sequences in length of three). Let $\mathcal{A} = \{a, c, g, t\}$. The space V of all four-sequences in length of three has $|V| = 4^3 = 64$ sequences. The random variables generated by equivalence of the positions are the frequencies of

$$(\text{adenines, cytosines, guanines, thymines}) = (f_a, f_c, f_g, f_t),$$

with $f_a + f_c + f_g + f_t = 3$. Consequently, the corresponding probability laws

$$\mathcal{L}_\lambda = \left(\frac{f_a}{3}, \frac{f_c}{3}, \frac{f_g}{3}, \frac{f_t}{3}\right)$$

are also random. The index λ in \mathcal{L}_λ indicates the corresponding orbit type, in analogy with expression (1.12), in which \mathcal{O}_0 and \mathcal{O}_4 belong to class \mathcal{O}_{40} , \mathcal{O}_1 and \mathcal{O}_3 belong to the class \mathcal{O}_{31} and \mathcal{O}_2 coincides with \mathcal{O}_{22} . We obtain these indices as the possible *integer partitions* of 3 in length of 4, namely the nonnegative integers $\{n_1, \dots, n_4\}$ with $n_1 \geq n_2 \geq n_3 \geq n_4 \geq 0$ satisfying $n_1 + \dots + n_4 = 3$. Consequently, there are 3 *types* of orbits, namely \mathcal{O}_{3000} , \mathcal{O}_{2100} and \mathcal{O}_{1110} , and corresponding laws:

$$\begin{aligned} \lambda = 3000 &\rightarrow \mathcal{L}_{3000} = (1, 0, 0, 0), \\ \lambda = 2100 &\rightarrow \mathcal{L}_{2100} = \left(\frac{2}{3}, \frac{1}{3}, 0, 0\right), \\ \lambda = 1110 &\rightarrow \mathcal{L}_{1110} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0\right). \end{aligned}$$

Similarly to expression (1.13) we now obtain

$$\begin{aligned} \text{Probability of a law type } \mathcal{L}_{3000} &= P(\mathcal{O}_{3000}) = \frac{\binom{3}{3,0,0,0}}{|V|} = \frac{3!}{3!0!0!0!} \frac{1}{64} = \frac{1}{64}, \\ \text{Probability of a law type } \mathcal{L}_{2100} &= P(\mathcal{O}_{2100}) = \frac{\binom{3}{2,1,0,0}}{|V|} = \frac{3!}{2!1!0!0!} \frac{1}{64} = \frac{3}{64}, \\ \text{Probability of a law type } \mathcal{L}_{1110} &= P(\mathcal{O}_{1110}) = \frac{\binom{3}{1,1,1,0}}{|V|} = \frac{3!}{1!1!1!0!} \frac{1}{64} = \frac{6}{64}, \end{aligned}$$

so that, under the assumption that all 4-sequences in length of 3 are equally likely (uniform probability), the most probable distribution by levels of nucleotides comes from the *class* of distribution

given by \mathcal{L}_{11110} , each of which has the highest probability, $6/64$. Simple combinatorics show that there are

$$\frac{4!}{3!1!} = 4$$

orbits of type $\lambda = 1110$, namely

$$(1.14) \quad \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0\right), \left(\frac{1}{3}, \frac{1}{3}, 0, \frac{1}{3}\right), \left(\frac{1}{3}, 0, \frac{1}{3}, \frac{1}{3}\right), \left(0, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right).$$

These are the most probable probability laws describing the nucleotide levels. See also Exercise 1.15. \square

1.7.1. Exchangeability. Consider an urn with 5 distinct marbles numbered 1, 2, 3, 4, 5, each one of color, say, yellow (y) or green (g). The possible urn configurations may be considered as non-observable events whereas the color or the number of a marble drawn from the urn are observable events. The urn compositions are represented by the structure $V = C^L$ of all mappings s defined in $L = \{1, 2, 3, 4, 5\}$ with values in $C = \{y, g\}$. Here it is natural to classify the possible configurations by the number of, say, yellow marbles. That is, making the marbles distinguishable by color only. This classification, as we now know, follows from letting the permutations in S_5 act on the structure V according to the rule

$$\varphi(\tau, s) = s\tau^{-1}$$

and counting two configurations s and f as equivalent when there is a permutation τ connecting (via $f = s\tau^{-1}$) the two mappings. The resulting classes of equivalent mappings are exactly the sets of urn compositions with 0, 1, \dots , 5 yellow marbles, and the exchangeable probability laws in V are convex combinations of those laws assigning equal or uniform probability to equivalent members.

A probability law w in C^L is exchangeable if $w(s) = w(s\tau^{-1})$ for all $\tau \in S_\ell$. Consequently, a probability law is exchangeable if it is a constant function in each of the left permutation orbits. A probability law w in C^∞ is exchangeable if w is exchangeable for all finite ℓ .

In its simplest form, De Finetti's Theorem states that to every exchangeable probability law w there corresponds a distribution F , concentrated in $[0, 1]$, such that, for all $0 < \ell$, for all $s \in C^L$, the representation,

$$w(s) = \int_0^1 \theta^k (1 - \theta)^{\ell-k} F(d\theta)$$

holds, where k is the number of ones in s .

All finite binary sequences have the same structure, and finite forms of exchangeability can be defined, with the resulting finite-type De Finetti theorems.

1.7.2. Partial exchangeability. Consider again the urn with 5 distinct marbles with the added information that marbles 1, 2 and 3 are larger in volume than marbles 4 and 5. The size-related or partial exchangeability erases the number-labels within each one of the two groups separately. The equivalent mappings representative of partially exchangeable sequences arise from the structure product of $V_1 = L_1^C$ and $V_2 = L_2^C$, where $L_1 = \{1, 2, 3\}$ and $L_2 = \{4, 5\}$. Two mappings (s, f) and (u, v) in $V_1 \times V_2$ are equivalent when there are permutations τ in S_3 and σ in S_2 such that $(s, f) = (u\tau^{-1}, v\sigma^{-1})$. The resulting classes of similar mappings are exactly the sets of configurations with k_1 yellow smaller marbles and k_2 yellow larger marbles, $k_1 = 0, 1, 2, 3$, $k_2 = 0, 1, 2$. If we indicate by $|\mathcal{O}_i|$ the size of the equivalence class (under the corresponding partial symmetry) for marbles of the same size, then the partially exchangeable probability laws are convex combinations of laws assigning equal or uniform probability $1/(|\mathcal{O}_1||\mathcal{O}_2|)$ to equivalent members.

1.7.3. Bilateral exchangeability. Indicate the left eye by OS and the right eye by OD, and let $L = \{OS, OD\}$. Also, let $C = \{1, 0\}$, where 1 stands for the condition that the eye is examined with lense refraction, and 0 stands for the condition that the eye is examined without lense refraction. At each point s in the mapping space $V = C^L$ we annotate a numerical expression of the resulting visual acuity $x(s) \in \mathcal{V}$, or some frequency data related to the acuity response from

a group of subjects. The hypothesis that the visual acuity response of the visual system due to lense refracting is indifferent to left-right indexing is described by making the permutations in S_2 act on V by $(\tau, s) \mapsto s\tau^{-1}$. Similarly to the previous examples, this action simplifies or factors the original structure V (with 4 labels) into 3 equivalency classes, or orbits $\mathcal{O}_1, \mathcal{O}_2, \mathcal{O}_3$, namely:

- (1) \mathcal{O}_1 bilateral refraction, with 1 label,
- (2) \mathcal{O}_2 monocular refraction, with 2 labels,
- (3) \mathcal{O}_3 without refraction, with 1 label.

These orbits in V , in turn, define the corresponding summaries and analysis in the data space \mathcal{V} . This structure is applicable to any bilateral biological system.

1.8. Symmetry and classification in classical mechanics

In physics as in chemistry, we find that certain physical properties of a system remain unchanged under certain transformations of such system. Riley, Hobson and Bence (2002) observe that

If a physical system is such that after application of a particular symmetry transformation the final system is indistinguishable from the original system then its behavior, and hence the functions that describe its behavior, must have the corresponding property of invariance when subject to the same transformations.

The study of these transformations is a study of the symmetries of the system. More generally, as Bacry (1963) shows, the study of the symmetries of a physical system often suggests the study of the symmetries of certain physical laws and theories, and not infrequently, leads to symmetry-related principles, such as Kepler's Law of planetary motion (*a planet covers equal elliptic areas in equal times in its trajectory, relative to a focal point*), the principle of time-reversal invariance or the Relativity Principle.

The following quote is from von Mises (1957, p.200), with the notation partially adapted. The theory studies the distributions of a certain number ℓ of molecules over f positions in the velocity space under the assumption that all possible f^ℓ distributions have the same probability. Given two molecules A and B , and three different positions a, b, c then the number of different distributions is 9, since each of the three positions of A , namely Aa, Ab, Ac can be combined with each of B . According to the classical theory, all these distributions, as random events, have the same probability, $1/9$.

A new theory, first suggested by the Indian physicist Bose¹, and developed by Einstein, chooses another assumption regarding the equal probabilities. Instead of considering single molecules and assuming that each molecule can occupy all positions in the velocity space with equal probability, the new theory starts with the concept of *repartition*. This is given by the number of molecules at each place of the velocity space, without paying attention to the individual molecules. From this point of view, only six 'partitions' are possible for two molecules on three places, namely, both molecules may be together at a , at b , or at c , or they may be separated, one at a and one at b , one at a and one at c , or one at b and one at c . According to the Bose-Einstein theory, each of these six cases has the same probability, $1/6$. In the classical theory, each of these three possibilities would have the probability of $1/9$, each of the other three, however, $2/9$, because, in assuming individual molecules, each of the last three possibilities can be realized in two different ways: A can be in a , and B in b , or vice versa, B can be in a , and A in b .

¹Satyendranath Bose, Born: 1 Jan 1894 in Calcutta, India Died: 4 Feb 1974 in Calcutta, India

The Italian physicist Fermi² advanced still another hypothesis. He postulated that only such distributions are possible- and possess equal probabilities- in which all molecules occupy different places. In our example of two molecules and three positions, there would only be three possibilities, each having the probability $1/3$; i.e., one molecule in a and one in b ; one in a and one in c ; one in b and one in c .

The arguments in von Mises' narrative can be expressed with the language of a symmetry study as follows: Let $L = \{A, B\}$, $C = \{a, b, c\}$ and V the set of all mappings $s : L \rightarrow C$, that is,

$$V = \left[\begin{array}{c|ccccccccc} s & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ \hline s(A) & a & b & c & a & b & a & c & b & c \\ s(B) & a & b & c & b & a & c & a & c & b \end{array} \right].$$

Under the Maxwell-Boltzmann (MB) model, it is assumed that all points or configurations in the space V are equally likely, or *uniformly distributed*, that is:

$$P(s) = \frac{1}{|V|} = \frac{1}{9}, \quad \text{for all } s \in V.$$

The volume $|V| = c^\ell$ of V is called the Maxwell-Boltzmann statistic.

Under the Bose-Einstein (BE) model, it is assumed that all points in the quotient space V/S_2 of V by the action $s\tau^{-1}$ of shuffling the molecules' labels (in $L = \{A, B\}$) are uniformly distributed. Thus, in the BE model, the uniform probability applies to the set of *orbits* of V obtained by label symmetry. The following matrix summarizes the action $s\tau^{-1}$ of S_2 on V :

$$(1.15) \quad \left[\begin{array}{c|ccccccccc} \sigma \backslash s & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ \hline 1 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ (12) & 1 & 2 & 3 & 5 & 4 & 7 & 6 & 9 & 8 \end{array} \right],$$

so that the six orbits in the quotient space V/S_2 are

$$\mathcal{O}_{11} = \{1\}, \quad \mathcal{O}_{12} = \{2\}, \quad \mathcal{O}_{13} = \{3\}, \quad \mathcal{O}_{21} = \{4, 5\}, \quad \mathcal{O}_{22} = \{6, 7\}, \quad \mathcal{O}_{23} = \{8, 9\},$$

each one of these having probability of $1/6$. A probability law in V/S_2 such as

$$P(s) = \begin{cases} 1/6 & \text{when } s \in \{\mathcal{O}_{11}, \mathcal{O}_{12}, \mathcal{O}_{13}\}, \\ 1/12 & \text{when } s \in \{\mathcal{O}_{21}, \mathcal{O}_{22}, \mathcal{O}_{23}\}, \end{cases}$$

would be consistent with the assumptions of the BE model. The Bose-Einstein statistic is the number

$$\binom{c + \ell - 1}{\ell}$$

of distinct orbits in the quotient space. In the example, there are $\binom{4}{2} = 6$ distinct orbits.

The Fermi-Dirac (FD) model assumes that only the injective mappings

$$V_I \equiv \left[\begin{array}{c|cccccc} s & 4 & 5 & 6 & 7 & 8 & 9 \\ \hline s(A) & a & b & a & c & b & c \\ s(B) & b & a & c & a & c & b \end{array} \right] \subset V$$

are admissible representations of the physical system, and that a uniform probability law is assigned to the resulting orbits in the quotient space of V_I by the action $s\tau^{-1}$ of shuffling the molecules'

²Enrico Fermi was born in Rome on 29th September, 1901. The Nobel Prize for Physics was awarded to Fermi for his work on the artificial radioactivity produced by neutrons, and for nuclear reactions brought about by slow neutrons. He died in Chicago on 29th November, 1954.

labels. Therefore, starting with

$$\left[\begin{array}{c|cccccc} \sigma \backslash s & 4 & 5 & 6 & 7 & 8 & 9 \\ \hline 1 & 4 & 5 & 6 & 7 & 8 & 9 \\ (12) & 5 & 4 & 7 & 6 & 9 & 8 \end{array} \right],$$

we obtain the three orbits $\mathcal{O}_1 = \{4, 5\}$, $\mathcal{O}_2 = \{6, 7\}$, $\mathcal{O}_3 = \{8, 9\}$ in the quotient space V_I/S_2 . To each of these, a probability of $1/3$ is assigned. In the present example, a probability law in V_I/S_2 given by

$$P(s) = \begin{cases} 1/6 & \text{when } s \in \{\mathcal{O}_1, \mathcal{O}_2, \mathcal{O}_3\}, \\ 0 & \text{otherwise,} \end{cases}$$

would be consistent with the assumptions of the FD model. The Fermi-Dirac statistic is the number $\binom{c}{\ell}$ of distinct orbits in the quotient space. In the example, we observed $\binom{3}{2} = 3$ distinct orbits.

Table 1.1 summarizes the domain of the uniform law in each of the models discussed above.

TABLE 1.1. Maxwell-Boltzmann, Bose-Einstein and Fermi-Dirac Probabilities.

| Model | Domain of the Uniform Law | Probability |
|-------------------|---------------------------|----------------------------|
| Maxwell-Boltzmann | V | $1/c^\ell$ |
| Bose-Einstein | V/G | $1/\binom{c+\ell-1}{\ell}$ |
| Fermi-Dirac | V_I/G | $1/\binom{c}{\ell}$ |

1.8.1. Macrostates and microstates in thermodynamics. Consider six numbered molecules indexed by the set $L = \{1, 2, 3, 4, 5, 6\}$ and four energy levels, indicated by the set $\mathcal{E} = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4\}$. The energy configurations are mappings

$$s : L \rightarrow \mathcal{E},$$

so that there is a total of $|\mathcal{E}|^{|L|} = 4^6 = 4096$ accessible microstates. We pass from microstates to measurable macrostates by dividing the space by similarities that result among the molecules when their identifying labels are erased. This is in analogy to erasing the position of the nucleotides in a four-sequence in length of six, as discussed in Example 1.7.1. Algebraically, this is obtained by letting the permutations in S_6 act on (by shuffling) the molecule labels in the set L . The composition rule is $s\tau^{-1}$. The resulting classes \mathcal{O}_λ of orbits are then the energy macrostates realized by the system. Here, λ indicates the possible integer partitions (n_1, n_2, n_3, n_4) of 6, that is, n_1, n_2, n_3 and n_4 are non negative integers such that $6 = n_1 + n_2 + n_3 + n_4$ with $n_1 \geq n_2 \geq n_3 \geq n_4$. The resulting classes, their volume $|\mathcal{O}_\lambda|$, usually indicated by Ω_λ in the thermodynamics context, and their number Q_λ of quantal states are:

| λ | Ω_λ | Q_λ | $\Omega_\lambda \times Q_\lambda$ |
|-----------|------------------|-------------|-----------------------------------|
| 6000 | 1 | 4 | 4 |
| 5100 | 6 | 12 | 72 |
| 4200 | 15 | 12 | 180 |
| 4110 | 30 | 12 | 360 |
| 3300 | 20 | 6 | 120 |
| 3210 | 60 | 24 | 1440 |
| 3111 | 120 | 4 | 480 |
| 2220 | 90 | 4 | 360 |
| 2211 | 180 | 6 | 1080 |
| total | 522 | 84 | 4096 |

There are $Q = 6$ quantal states associated with the most probable ($\Omega = 180$) orbit type, $\lambda = 2211$. Also note that

$$\sum_{\lambda} Q_{\lambda} = \binom{|\mathcal{E}| + |\mathbf{L}| - 1}{|\mathbf{L}|} = \binom{9}{6} = 84$$

is the Bose-Einstein statistic.

1.8.2. Boltzmann's Entropy Theorem. In Boltzmann model all particles are considered to be distinguishable, so that a uniform probability can be assigned to the ensemble. However, the passage from the accessible microstates to macrostates is equivalent to obtaining a partition of the ensemble V of accessible microstates into orbits of symmetry realized by the symmetric group acting on V according to the composition rule $s\tau^{-1}$. It is an important observation that the mean energy level

$$(1.16) \quad \bar{\mathcal{E}} = \frac{1}{\ell} \sum_i \mathcal{E}_i f_i,$$

where $f_i = |s^{-1}(\mathcal{E}_i)|$ indicates the number of molecules at the energy level \mathcal{E}_i , of any configuration in V , is an invariant under the composition rule $s\tau^{-1}$ and, therefore, depends only on the orbit (macrostate) realized by the configuration. Boltzmann reasoned that the molecule-energy configurations in V evolved from least probable configurations to most probable configurations, so that the quest for describing the equilibrium energy distribution in the ensemble requires the determination of the most likely configurations in V . This, in turn, requires the determination of the macrostate (orbit) with the largest volume Ω , conditioned on the fact that mean energy of the isolated ensemble must remain constant. Given a configuration s with f_1 particles at the energy level \mathcal{E}_1 , f_2 particles at the level \mathcal{E}_2 , f_3 particles at the level \mathcal{E}_3 , etc, its orbit \mathcal{O}_s has volume

$$(1.17) \quad |\mathcal{O}_s| = \frac{\ell!}{f_1! f_2! f_3! \dots}$$

We have then a well-defined mathematical problem: find the macrostate identified by f_1, f_2, \dots which maximizes (1.17) for a given mean energy level $\bar{\mathcal{E}}$. The solution is $f_i = \ell P(\mathcal{E}_i)$, where

$$(1.18) \quad P(\mathcal{E}_i) = \frac{e^{-\beta \mathcal{E}_i}}{\sum_j e^{-\beta \mathcal{E}_j}}$$

is the Maxwell-Boltzmann canonical distribution. It describes the most likely energy distribution of the ensemble. For reference, its classical derivation is outlined in the Appendix. Similar calculations can be obtained for the models of Fermi-Dirac and Bose-Einstein.

We conclude this example noting that the constrained minimization of $\sum f_i \ln f_i$ is equivalent to the constrained maximization of

$$H = - \sum_i \frac{f_i}{\ell} \ln \left(\frac{f_i}{\ell} \right)$$

which is the *entropy* of the probability law associated with the orbit of f_1, f_2, f_3, \dots . The entropy, usually indicated by S in thermodynamics, is a physical characteristic (such as temperature, mass) of the gas and at the same time, a measure of uniformity in its thermodynamical probability law. The canonical distribution corresponds to an ensemble configured to its maximum entropy. Boltzmann's statistical expression

$$S = k \ln \Omega$$

for the equilibrium entropy relates the equilibrium or limit number of accessible microstates, Ω , and k , the (known now as) Boltzmann constant 1.3807×10^{-23} K J/molecule. A volume of gas, left to itself, will almost always be found in the state of the most probable distribution.

1.8.3. Maxwell-Boltzmann Law for velocities in a perfect gas. In this example we outline the classical derivation of Maxwell-Boltzmann Law. In the context of the orbit method, Maxwell's assumptions e.g., Ruhla (1989, Ch.4) led to the searching of a probability law, indicated here by F , for the random velocity vector (\mathbf{v}) satisfying the following conditions: First, the component-velocities are statistically independent and identically distributed, so that the law F should have the form

$$F(\mathbf{v}) = f(v_x)f(v_y)f(v_z),$$

where f indicated the common probability law for the component-velocities. The *isotropic condition* states that F should be invariant under all central rotations, indicated here by U , in the three-dimensional Euclidian space \mathbb{R}^3 . Denoting by $SO(3, \mathbb{R})$ the collection of all such rotations, we write the isotropic condition as,

$$(1.19) \quad F(U\mathbf{v}) = F(\mathbf{v}), \text{ for all } U \in SO(3, \mathbb{R}).$$

Note the analogy between the isotropic condition and the invariance condition described by Expression (1.11). These two conditions lead to the probability law which has the form

$$(1.20) \quad F(\mathbf{v}) = A^3 e^{-\mu \|\mathbf{v}\|^2},$$

where $v = \|\mathbf{v}\| = \sqrt{v_x^2 + v_y^2 + v_z^2}$ is the *speed* in the velocity vector \mathbf{v} . The constants are determined from additional physical considerations. The orbits \mathcal{O}_v in the quotient space are exactly those velocity vectors \mathbf{v} in \mathbb{R}^3 with common speed v .

1.9. Canonical decompositions

In the previous sections we have illustrated the classification of objects using symmetry transformations, the representation of these symmetries as linear transformations in the vector space for the data observed on those objects, and the reduction (Section 1.6) of the structure into sets of symmetrically equivalent points. The formal connection between those basic steps and the context of statistical inference depends of the notion of canonical projections or decompositions, studied in detail in the following chapter. Canonical projections appear in many fundamental and practical terms in Chemistry and Physics.

To illustrate, recall that the symmetry among the molecules, imposed by the Bose-Einstein argument described in Section 1.8, led to the classification of the points in V into six orbits

$$\mathcal{O}_{11} = \{1\}, \quad \mathcal{O}_{12} = \{2\}, \quad \mathcal{O}_{13} = \{3\}, \quad \mathcal{O}_{21} = \{4, 5\}, \quad \mathcal{O}_{22} = \{6, 7\}, \quad \mathcal{O}_{23} = \{8, 9\},$$

derived from the action

$$\left[\begin{array}{c|cccccccc} \sigma \backslash s & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ \hline 1 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ (12) & 1 & 2 & 3 & 5 & 4 & 7 & 6 & 9 & 8 \end{array} \right]$$

$s\tau^{-1}$ of S_2 on V . We observe that the action

$$\left[\begin{array}{c|cccccccc} \sigma \backslash s & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ \hline 1 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{array} \right]$$

of the identity permutation on V can be represented by the identity matrix I in \mathbb{R}^9 . We write $\rho(1) = I$. Similarly, the action

$$\left[\begin{array}{c|cccccccc} \sigma \backslash s & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ \hline (12) & 1 & 2 & 3 & 5 & 4 & 7 & 6 & 9 & 8 \end{array} \right]$$

of the permutation $t = (12)$ on V can be linearly represented by

$$\rho(t) = \text{Diag} (1, 1, 1, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}).$$

This is the same principle described earlier on in the chapter to represent the action of the point group C_{2h} on the framework of the dichloroethene molecule. The evaluation of

$$\mathcal{P}_1 = [\rho(1) + \rho(t)]/2 = \frac{1}{2}\text{Diag} (2, 2, 2, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}),$$

and

$$\mathcal{P}_2 = [\rho(1) - \rho(t)]/2 = \frac{1}{2}\text{Diag} (0, 0, 0, \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix})$$

shows that

$$(1.21) \quad \mathcal{P}_1^2 = \mathcal{P}_1, \quad \mathcal{P}_2^2 = \mathcal{P}_2, \quad \text{and } \mathcal{P}_1\mathcal{P}_2 = 0,$$

thus showing that \mathcal{P}_1 and \mathcal{P}_2 are (commuting) projections onto orthogonal subspaces \mathcal{V}_1 and \mathcal{V}_2 of $\mathcal{V} = \mathbb{R}^9$ with

$$(1.22) \quad \boxed{\mathbf{I} = \mathcal{P}_1 + \mathcal{P}_2}.$$

Equivalently, \mathcal{V} reduces into the direct sum $\mathcal{V}_1 \oplus \mathcal{V}_2$ of subspaces of dimensions $d_1 = \text{tr } \mathcal{P}_1 = 6$ and $d_2 = \text{tr } \mathcal{P}_2 = 3$, respectively.

When properties (1.21) are verified, the decomposition (1.22) is called a *canonical decomposition* and \mathcal{P}_1 and \mathcal{P}_2 are the *canonical projections* associated with the action of G on V . These decompositions play a fundamental part in all symmetry studies. The study of their derivations and properties is the primary aim of Chapter 2.

An important application of all canonical reductions is the computation and interpretation of vectors that *transform* according to certain basic properties (called irreducible representations) of the group of symmetries. In the present example, that means that for any vector $x \in \mathcal{V}$, the new vector $y_1 = \mathcal{P}_1x$ transforms according to the one-dimensional identity (or symmetric) representation of S_2 , whereas $y_2 = \mathcal{P}_2x$ transforms according to another one-dimensional representation of S_2 , namely one that assigns 1 to the identity and -1 to the transposition (this representation is called the *sign* or anti-symmetric representation, and will be introduced in Chapter 2).

In quantum chemistry these projections play a central role in determining whether a chemical bonding can take place in a molecule. The bonding, points Riley et al. (2002, p.948), is strongly dependent upon whether the wavefunction of the two atoms forming a bond transform according to the same (irreducible) representation. Typically, these reductions take place in a infinite dimension Hilbert space \mathcal{H} , in such way that the invariant subspaces in the reduction of \mathcal{H} define the *properties* of the quantum system. Consequently, properties are identified with the corresponding (Hermitian) projections. For example, the projections \mathbf{I} and 0 correspond to the sure property and the impossible property, whereas the projection $\mathbf{I} - \mathcal{P}$ corresponds to the negation of the property associated with \mathcal{P} . The properties associated with two commuting projections \mathcal{P} and \mathcal{Q} are said to be compatible, in which case the projection $\mathcal{P}\mathcal{Q} = \mathcal{Q}\mathcal{P}$ represents the conjunction of \mathcal{P} and \mathcal{Q} , whereas the projection $\mathcal{P}\mathcal{Q} + \mathcal{P}(\mathbf{I} - \mathcal{Q}) + (\mathbf{I} - \mathcal{P})\mathcal{Q} = \mathcal{P} + \mathcal{Q} - \mathcal{P}\mathcal{Q}$ is associated with the disjunction of the two properties. If, in addition, $\mathcal{P}\mathcal{Q} = 0$, the properties are compatible, mutually exclusive, and the disjunction is given by the sum $\mathcal{P} + \mathcal{Q}$. See, for example Omnès (1994) or Faris (1996) for a review of Omnès' work.

Unit vectors y in the Hilbert space are associated with the *states* of the system, and determine a mathematical specification of probabilities for all properties. These probabilities are obtained from the fact that associated to each set of mutually exclusive properties $\mathcal{P}_1, \dots, \mathcal{P}_h$ whose disjunction is sure, that is,

$$\boxed{\mathbf{I} = \mathcal{P}_1 + \dots + \mathcal{P}_h},$$

there is a decomposition

$$1 = \|y\|^2 = y'y = y'\mathcal{P}_1y + \dots + y'\mathcal{P}_hy$$

which is interpreted as a probability distribution among the corresponding properties $\mathcal{P}_1, \dots, \mathcal{P}_h$. Each state $y \in \mathcal{H}$ then provides a probabilistic description,

$$y \rightarrow (y'\mathcal{P}_1y, \dots, y'\mathcal{P}_hy) = (\|\mathcal{P}_1y\|^2, \dots, \|\mathcal{P}_hy\|^2),$$

of the system³. In the present analogy with $I = \mathcal{P}_1 + \mathcal{P}_2$, a unitary state y is associated with the probability distribution

$$P(\mathcal{P}_1) = \|\mathcal{P}_1 y\|^2 = y_1^2 + y_2^2 + y_3^2 + \frac{1}{4}(y_4 + y_5)^2 + \frac{1}{4}(y_6 + y_7)^2 + \frac{1}{4}(y_8 + y_9)^2,$$

$$P(\mathcal{P}_2) = \|\mathcal{P}_2 y\|^2 = \frac{1}{4}(y_4 - y_5)^2 + \frac{1}{4}(y_6 - y_7)^2 + \frac{1}{4}(y_8 - y_9)^2.$$

Moreover, we observe that

$$\mathcal{P}_1 y = (y_1, y_2, y_3, \frac{y_4 + y_5}{2}, \frac{y_4 + y_5}{2}, \frac{y_6 + y_7}{2}, \frac{y_6 + y_7}{2}, \frac{y_8 + y_9}{2}, \frac{y_8 + y_9}{2}),$$

and

$$\mathcal{P}_2 y = (0, 0, 0, \frac{y_4 - y_5}{2}, \frac{y_5 - y_4}{2}, \frac{y_6 - y_7}{2}, \frac{y_7 - y_6}{2}, \frac{y_8 - y_9}{2}, \frac{y_9 - y_8}{2}),$$

identify 9 one-dimensional subspaces associated with the symmetry operation erasing the identity of the particles:

$$\pm y_1, \pm y_2, \pm y_3, \pm(y_4 + y_5), \pm(y_6 + y_7), \pm(y_8 + y_9),$$

derived from \mathcal{P}_1 and

$$\pm(y_4 - y_5), \pm(y_6 - y_7), \pm(y_8 - y_9),$$

derived from \mathcal{P}_2 . It is important to observe that, although these projections \mathcal{P}_1 and \mathcal{P}_2 lead to these subspaces, there is no a priori reason to single out any particular bases for those subspaces. In fact, there are infinitely many such (stable) subspaces, all equivalent to each other in a way that will be made precise in Chapter 2. The *number* (two in this example) of canonical subspaces, on the other hand, is a constant and depends only on the group of symmetries (S_2 in this case).

The projections on one-dimensional subspaces are *pure states*. A further consequence of the symmetries imposed to the system is observed from the fact that these subspaces are characteristic of exactly the two types of one-dimensional representations: those associated with \mathcal{P}_1 are symmetric, that is, $\rho(\tau)\mathcal{P}_1 y = \mathcal{P}_1 y$ for all $\tau \in S_2$; and those associated with \mathcal{P}_2 are anti-symmetric, that is,

$$\rho(\tau)\mathcal{P}_2 y = \begin{cases} \mathcal{P}_2 y & \text{if } \tau = 1 \\ -\mathcal{P}_2 y & \text{if } \tau = (12). \end{cases}$$

As will become clear in the sequence, the usual arithmetic mean and the deviations from the mean are the (only) two symmetry invariants, of dimensions 1 and $n - 1$ respectively, that appear naturally when the data are indexed by $V = \{1, \dots, n\}$ and the symmetries are all the permutations of the indices in V . This is the case, of course, that we are familiar with in univariate statistical sampling x_1, x_2, \dots, x_n .

In general, if there are h canonical projections, then the identity operator, I , in the data space \mathcal{V} decomposes as $I = \mathcal{P}_1 + \mathcal{P}_2 + \dots + \mathcal{P}_h$, where $\mathcal{P}_i \mathcal{P}_j = 0$ for $i \neq j$ and $\mathcal{P}_i^2 = \mathcal{P}_i$, $i = 1, \dots, h$. It then follows that the basic partition

$$\|x\|^2 = (x|x) = (x|\mathcal{P}_1 x) + (x|\mathcal{P}_2 x) + \dots + (x|\mathcal{P}_h x)$$

of the sum of squares for a particular inner product $(\cdot|\cdot)$ of interest (e.g., Euclidean, Hermitian, symplectic) can be obtained. In the sampling case mentioned above, the canonical decomposition is simply

$$\sum_j x_j^2 = n\bar{x}^2 + \sum_j (x_j - \bar{x})^2,$$

a consequence of the fact that in this case $I = \mathcal{A} + \mathcal{Q}$ for exactly two non-zero canonical projections

$$\mathcal{A} = \frac{1}{n} ee', \quad \mathcal{Q} = I - \mathcal{A},$$

where ee' is the $n \times n$ matrix of ones. The reduction $I = \mathcal{A} + \mathcal{Q}$ satisfies $\mathcal{A}^2 = \mathcal{A}$, $\mathcal{Q}^2 = \mathcal{Q}$ and $\mathcal{A}\mathcal{Q} = \mathcal{Q}\mathcal{A} = 0$. Moreover, \mathcal{A} projects $\mathcal{V} = \mathbb{R}^n$ into a subspace \mathcal{V}_a of dimension $\dim \mathcal{V}_a = \text{tr } \mathcal{A} = 1$ generated by $e = e_1 + \dots + e_n = (1, 1, \dots, 1) \in \mathcal{V}$, whereas \mathcal{Q} projects \mathcal{V} into an irreducible subspace \mathcal{V}_q in dimension $n - 1$, the orthogonal complement of \mathcal{V}_a in \mathcal{V} .

³The notation $y' \mathcal{P} y$ has the interpretation of the $\langle y, \mathcal{P} y \rangle$ under the appropriate inner product in \mathcal{H}

The invariants $\mathcal{P}\mathbf{x}$ and $\mathcal{Q}\mathbf{x}$ are respectively the arithmetic mean and the deviations from the mean. In general, for (large-sample) normally distributed data, the Fisher-Cochran theory for the probability distribution of quadratic forms leads to varied forms of analyses of variance, within which parametric hypotheses of the form $\mathbf{H}_j : \mu' \mathcal{P}_j \mu = 0$ based on the expected value μ of \mathbf{x} can be defined and interpreted within a given scientific context.

It is certainly less obvious that the same principle of canonical decomposition when applied to symmetry studies of bilateral systems (eyes, ears, hemispheres) would show that the decomposition

$$\mathbf{A} = \frac{1}{2} \mathbf{A}_{\text{intra class}} + \frac{1}{2} (s_1^2 - s_2^2) \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

for the sample covariance matrix \mathbf{A} obtains, thus saying that matrix \mathbf{A} is an intraclass matrix if and only if the second component in the decomposition vanishes, that is, when the equality $s_1^2 = s_2^2$ of the variances in the *left* and *right* sides of the mechanism holds.

We remark, consequently, that symmetry studies are characterized by a formal and systematic algebraic statistical framework within which properties of the mechanisms under consideration can be suggested (by invariant-related hypotheses) and studied (by statistical inference). The systematic, step-by-step nature of the algebraic component of the argument is clearly remarked in Serre (1977, p.22). The canonical projections are the key elements leading to the explicit calculation and interpretation of the invariants $\mathcal{P}_1\mathbf{x}, \mathcal{P}_2\mathbf{x}, \dots$ in the data, which, once available, would be the basis for designing the corresponding symmetry studies.

Appendix -Maxwell-Boltzmann canonical distribution

Using Stirling's approximation $\ln t! \equiv t \ln t - t$, we have,

$$\ln |\mathcal{O}_s| = \ln \ell! - \sum \ln f_i = \ell \ln \ell - \ell - \sum (f_i \ln f_i - f_i) = \ell \ln \ell - \sum f_i \ln f_i.$$

Equivalently, then, we seek to minimize $\sum f_i \ln f_i$ subject to (1.16). These two conditions lead to

$$(\ell + \sum_i \ln f_i) df_i = 0, \quad \sum_i \mathcal{E}_i df_i = 0.$$

A sufficient condition for the existence of a solution (using Lagrange multipliers argument) is that there are constants α and β satisfying $\sum_i (\mathcal{E}_i + \alpha + \beta \ln f_i) df_i = 0$, in which case the solutions take the form $f_i = \alpha e^{-\beta \mathcal{E}_i}$. The condition $\sum_i f_i = \ell$ implies $\alpha = \ell / \sum_i e^{-\beta \mathcal{E}_i}$, so that

$$(1.23) \quad f_i = \ell \frac{e^{-\beta \mathcal{E}_i}}{\sum_j e^{-\beta \mathcal{E}_j}}.$$

The value of β follows from the condition $\frac{1}{\ell} \sum_i f_i \mathcal{E}_i = \bar{\mathcal{E}}$. That is, β is a solution of

$$\frac{\sum_i e^{-\beta \mathcal{E}_i} \mathcal{E}_i}{\sum_j e^{-\beta \mathcal{E}_j}} = \bar{\mathcal{E}}.$$

From (1.23) we then obtain Maxwell-Boltzmann canonical distribution shown in equation (1.18).

Further reading

In his text on symmetry and science, Joseph Rosen (1995) characterizes symmetry as *immunity to a possible change*. His book includes an accessible introduction to the mathematics of symmetry and leads to the formulation of his Symmetry Principle. See also Rosen (1975).

The classical introductory work of Hermann Weyl (1952) includes the notions of bilateral, translatory, rotational, ornamental and crystal symmetry.

A thing that is symmetrical ... if there is something that you can do to it, so that after you have finished doing it, it still looks the same as it did before you did it.

Hermann Weyl was born on Nov 9th, 1885 in Elmshorn, Germany, and died on Dec 8th, 1955 in Zürich, Switzerland. He was a student of Hilbert at Göttingen, and from 1933 until he retired in 1952 he worked at the Institute for Advanced Study at Princeton. From 1923-38 he evolved the concept of continuous groups using matrix representations and its applications to quantum mechanics. Weyl (1950) is the English translation of the original text *Gruppentheorie und Quantenmechanik*, first published in 1931. Weyl (1953) is a revised and supplemented edition of his 1939 publication on the invariants and representations of the classical groups. Weyl, along with Wigner's *Gruppentheorie und ihre Anwendungen auf die Quantenmechanik der Atomspektren* (Braunschweig: Vieweg, 1931) and van der Waerden's *Die gruppentheoretische Methode in der Quantenmechanik* (Berlin: Springer, 1932) pioneered the methods of group representation to quantum mechanics- *the three W's* of quantum mechanics. The English translation of Wigner's work by J. J. Griffin appeared in 1959 under the title *Group theory and its application to quantum mechanics and atomic spectra*, Academic Press, New York.

The notion of points as labels identifying potential events appears in modern-day physics, in contrast to Newton's views in which points are essentially indistinguishable. A comment in that direction is found in Cartier (2001).

George Pólya, an American mathematician of Hungarian origin, was born in Budapest, Hungary, on December 13, 1887, and died in Palo Alto, USA on September 7, 1985. He worked on a variety of mathematical topics, including series, number theory, combinatorics, and probability. During the first decades of the 1900's he had the company of many leading mathematicians such as Klein, Carathéodory, Hilbert, Runge, Landau, Weyl, Hecke, Courant and Toeplitz. Geometric symmetry and the enumeration of symmetry classes of objects was a major area of interest for Pólya over many years. He added to the understanding of the 17 plane crystallographic groups in 1924 by illustrating each with tilings of the plane. Pólya's work using generating functions and permutation groups to enumerate isomers in organic chemistry was of fundamental importance. In 1978, at the age of 91, he taught a course on combinatorics in the Computer Science Department at Stanford.

The delicate and fascinating theory of measurement by von Neumann and its classical/quantum interpretations are discussed in great detail in the work of Omnès (1994, pp. 60,72). Measurements are sharper judgments, and judgments are broader measurements, remarks de Finetti (1972, p.165).

The text by Martin Aigner (1979) on Combinatorial Theory has a comprehensive discussion on symmetry operations on the set of functions on finite sets.

Vibrational spectroscopy is a perfect example illustrating the objective connection between symmetry and observable measurements. The reader may refer to Harris and Bertolucci (1978), where the authors review the classical symmetry operations applied to molecules and their resulting classification according to the symmetries of point groups. The algebraic methods introduced later on in the sequence are an integral part of the contemporary language with which the theory of vibrational spectroscopy can be explained.

Erwin Schrödinger was born on August 12, 1887, in Vienna. His great discovery, the wave equation in quantum mechanics, was made during the first half of 1926. A colleague of Hermann Weyl and Peter Debye, he was greatly interested in Boltzmann's probability theory. For this work on the atomic spectra as an eigenvalue problem he shared with Dirac the Nobel Prize for 1933. Schrödinger's (1967) text *What is Life? The Physical Aspects of the Living Cell* remains a classic, objective reading on the connections among physics, chemistry, biology and life. Of particular interest is his account of the role of symmetry on molecular stability and transitions between stationary states, e.g., pp. 49-55. He died on the 4th of January, 1961, in Vienna, after a long illness.

The reader may consult Snedecor and Cochran (1989) for the basic notions of classical statistical inference, including the analysis of variance. The statistical aspects of quadratic forms (needed for the second-order analysis associated with the canonical projections) are developed, for example, in Rao (1973), Eaton (1983), Searle (1971), Muirhead (1982). , , ,

The characterization of cyclic symmetries in the study of purine and pyrimidine contents of local nucleotide sequences for evolution of human immunodeficiency virus type 1 is present in the work of Doi (1991) on evolutionary molecular biology.

The literature on the many aspects of symmetry in science and methodology is overly extensive. It ranges from studies considering the role of symmetry in beauty and evolution, e.g., Enquist and Arak (1994), anatomic symmetry between fellow eyes, e.g., Pauleikhoff, Wormald, Wright, Wessing and Bird (1992), parallel visual processes in symmetry perception in normal vision, e.g., Wagemans (1999), symmetry discrimination, e.g., Szlyk, Seiple and Xie (1995), Szlyk, Rock and Fisher (1995), visual signaling by asymmetry, e.g., Swaddle (1999), Tyler (1996), to the works of Graf and Schachman (1996) on random circular permutations of genes and polypeptide chains and of Hellige (1993) on hemispherical asymmetry.

Symmetry-related considerations appear in many aspects of statistics: Symmetry and the covariance structure of symmetrically dependent ordered observations, e.g., Lee and Viana (1999), Viana and Olkin (2000), Lee (1998). ; Invariant measures in groups and their uses in statistics e.g., Wijsman (1990). ; Invariance in factorial models, McCullagh (2000), quotient spaces and statistical models, McCullagh (1999); symmetry models and structured data Dawid (1988), Bailey (1991); Algebraic statistics, e.g., Pistone, Riccomagno and Wynn (2000).

A connection between the notions of symmetry and prior (to experiment) predictions or statements is described in Weyl (1952, p.126) where he argues that all a priori statements in physics have their origin in symmetry. If conditions which uniquely determine their effect possess certain symmetries, then the effect will exhibit the same symmetry. This is also Rosen's symmetry principle (Rosen (1995)). For example, equal weights balance in scales of equal arms, concluded Archimedes a priori; in casting dice which are *perfect* cubes, each side is perceived as equally likely. In contrast, the law of equilibrium for scales with arms of different lengths can only be settled by experience or by physical principles based on experience.

Exercises

Exercise 1.1. Use the planar transformations $\{1, v, h, r, o\}$ to classify the 26 letter symbols A, B, ..., Z and describe the resulting types based on their symmetry groups. Pólya (1954, p.89).

Exercise 1.2. Starting with the list $\{t, v, h, r\}$ of symmetry transformations introduced above, classify the following equations Pólya (1954, p.89):

$$y = x^2, \quad y^2 = x, \quad y = x^3, \quad x^2 + 2y^2 = 1, \quad y = x + x^4.$$

Exercise 1.3. The following are some elementary Euclidean (symmetry) transformations in 3-dimensional space:

(1) Displacements: $I + D$, where D is a diagonal matrix and I is the identity matrix;

(2) Rotation by θ deg around the z axis: $R_\theta = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$;

(3) Plane reflection (fixing the xy plane): $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$;

(4) Line inversion (fixing the z axis): $\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$;

(5) Glide (fixing the xy plane): $\begin{bmatrix} 1+a & 0 & 0 \\ 0 & 1+b & 0 \\ 0 & 0 & -1 \end{bmatrix}$;

(6) Screw (fixing the z axis): $R(\theta) + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & c \end{bmatrix}$;

(7) Spacial dilation: ρI ;

(8) Plane and line projections.

For any two of these Euclidean transformations T_1 and T_2 , classify the resulting product transformation T_1T_2 and its commutativity (Is $T_1T_2 - T_2T_1 = 0$?). Determine, for each transformation T , the existence of its inverse transformation.

Exercise 1.4. Determine the group of symmetries of the parallelepiped with vertices located at $(\pm a, \pm b, \pm c) \in \mathbb{R}^3$ when:

- (1) $a \neq b \neq c$;
- (2) $a = b \neq c$;
- (3) $a = b = c$.

You may assume $a > 0, b > 0, c > 0$.

Exercise 1.5. The following are the basic transformations used to describe a molecular structural symmetries. The notation is that used in the chemistry literature:

- (1) C_n : an axial (clockwise) rotation of $360/n$ deg. A molecule with a C_n rotational axis can be rotated $360/n$ deg along that axis. C_n^m indicates the m -fold (or iterated) rotation and is equivalent to a rotation by $m \times 360/n$ deg;
- (2) i : point symmetry or center of inversion;
- (3) σ_h : a reflection on a plane orthogonal to the principal rotational axis (the C_n axis with largest n);
- (4) σ_v : a reflection on a plane containing the principal rotational axis;
- (5) σ_d : a reflection in a plane of symmetry containing the principal rotational axis and bisecting the angle between two 2-fold axes of symmetry (C_2) which are orthogonal to the principal axis;
- (6) S_n : an improper rotation or n -fold alternating axis of symmetry: this is a combination of a n -fold rotation and a reflection on a plane of symmetry orthogonal to the rotation axis.

Consider the point group $\mathcal{C}_3 = \{E, C_3, C_3^2\}$ and its realization in which the axis of rotation is $e_3 = (0, 0, 1)$. Calculate the multiplication table for \mathcal{C}_3 .

Exercise 1.6. Following the notation of Exercise 1.5, determine the multiplication table for the point group $\mathcal{D}_2 = \{E, C_{2,1}, C_{2,2}, C_{2,3}\}$. The rotation axes of \mathcal{D}_2 are orthogonal, so that you may consider the realization in which the axes of $C_{2,j}$ are e_1, e_2 and e_3 .

Exercise 1.7. Following the notation of Exercise 1.5, determine the multiplication table for the point group $\mathcal{C}_{4v} = \{E, C_4, C_4^2, C_4^3, \sigma_{d,1}, \sigma_{d,2}, \sigma_{v,1}, \sigma_{v,2}\}$. You may consider the realization in which the axis of rotation for C_4 is e_3 and the reflection planes $\sigma_{v,1}, \sigma_{v,2}, \sigma_{d,1}$ and $\sigma_{d,2}$ are defined by

- (1) $\sigma_{v,1}$: e_1 and e_3 ;
- (2) $\sigma_{v,2}$: e_2 and e_3 ;
- (3) $\sigma_{d,1}$: $e_1 + e_2$ and e_3 ;
- (4) $\sigma_{d,2}$: $e_1 - e_2$ and e_3 .

Exercise 1.8. Following the notation of Exercise 1.5, consider the point group

$$\mathcal{D}_{2h} = \{E, C_{21}, C_{22}, C_{23}, \sigma_{h1}, \sigma_{h2}, \sigma_{h3}, i\}$$

and the realization in which the axis of rotation for C_{2j} are e_1, e_2 and e_3 and the reflection plane defining σ_{hj} is the plane orthogonal to e_j through the origin $(0, 0, 0)$. Show that the multiplication

- (2) show that P is left-invariant, that is, $P(s\tau^{-1}) = P(s)$ for all $\tau \in S_\ell$ and $s \in V$;
 (3) conversely, show that all left invariant probability laws in V are linear convex combinations of locally (within-orbit) uniform laws.

Exercise 1.15. Following Exercise 1.14, show that there are only three types of position-symmetry orbits, \mathcal{O}_λ , namely \mathcal{O}_{40} , \mathcal{O}_{31} and \mathcal{O}_{22} , corresponding to the integer partitions $\lambda = 40$, $\lambda = 31$ and $\lambda = 22$ of 4 in length of 2. Consequently, verify that

$$\text{Probability of law type } \mathcal{L}_{40} = P(\mathcal{O}_{40}) = \frac{\binom{4}{4,0}}{|V|} = \frac{4!}{4!0!} \frac{1}{16} = \frac{1}{16},$$

$$\text{Probability of law type } \mathcal{L}_{31} = P(\mathcal{O}_{31}) = \frac{\binom{4}{3,1}}{|V|} = \frac{4!}{3!1!} \frac{1}{16} = \frac{1}{4},$$

$$\text{Probability of law type } \mathcal{L}_{22} = P(\mathcal{O}_{22}) = \frac{\binom{4}{2,2}}{|V|} = \frac{4!}{2!2!} \frac{1}{16} = \frac{6}{16}.$$

Moreover, there are $\binom{2}{1,1} = 2$ orbits of type $\lambda = 40$ or type $\lambda = 31$ and $\binom{2}{2,0} = 1$ orbit of type $\lambda = 22$.

Exercise 1.16. Following Exercise 1.14, let $\Gamma = \{s_k = \underbrace{(1, \dots, 1)}_k, 0, \dots, 0\}; k = 0, \dots, \ell\}$.

- (1) Show that $\Gamma \cap \mathcal{O}_k = \{s_k\}$, $k = 0, \dots, \ell$. That is, Γ is a cross-section in V ;
 (2) If x is a scalar function defined in V and P is a probability law in V , show that

$$\sum_{s \in V} x(s)P(s) = \sum_{s \in \Gamma} \sum_{\tau \in G} x(s\tau^{-1})P(s\tau^{-1})/|\mathcal{O}_k|$$

Exercise 1.17. Suppose that the energy level E of a system is determined by the equation

$$E(n_x, n_y, n_z) = \frac{n_x^2}{a} + \frac{n_y^2}{b} + \frac{n_z^2}{c},$$

where $n_x, n_y, n_z \in \{1, 2, 3\}$ and a, b, c are (real positive) constants of the system.

- (1) Determine all energy levels of the system in each one of the following cases:
 (a) $a = 1/5, b = 1/2, c = 1/4$;
 (b) $a = b = 1/5, c = 1/4$;
 (c) $a = b = c = 1/5$.
 (2) What is the role of symmetry in determining the number of energy levels?

Exercise 1.18. The area of the parallelogram determined by the vectors $v' = (r_1, p_1)$ and $w' = (r_2, p_2)$ in \mathbb{R}^2 is given by the absolute value of the symmetric bilinear form

$$(v, w) = \det \begin{bmatrix} r_1 & r_2 \\ p_1 & p_2 \end{bmatrix} = r_1 p_2 - r_2 p_1,$$

or, equivalently, $(v, w) = v' \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} w$. Let

$$\text{Sp}(2, \mathbb{R}) = \{\text{non-singular } 2 \times 2 \text{ real matrices } A; (Av, Aw) = (v, w)\}.$$

That is, $\text{Sp}(2, \mathbb{R})$ is the set of (symmetry) transformations leaving the form (\cdot, \cdot) invariant. Show that $\text{Sp}(2, \mathbb{R})$ is a subgroup of the group $\text{GL}(2, \mathbb{R})$ of non-singular real 2×2 matrices. In Physics, r and p have the interpretation of position and momentum, respectively, of a particle constrained to move on a line. The set of vectors $(r, p) \in \mathbb{R}^2$ define the particle's *phase space*. Its geometry, determined by the standard form (\cdot, \cdot) , is called *symplectic* geometry and $\text{Sp}(2, \mathbb{R})$ is the (real) symplectic group. It respects the geometry of the phase space and is one of the classic isometry groups.

Exercise 1.19. If an object of mass m is moving in the xy plane according to

$$r(t) = ut \cos(t + \alpha), vt \sin(t + \alpha), 0,$$

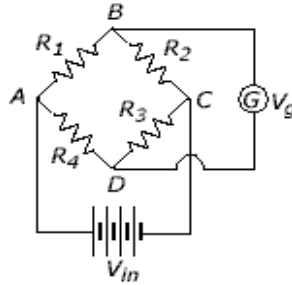
evaluate its momentum $p = mdr/dt$ and show that the angular momentum $r \times p$ remains invariant under rotations in the xy plane around the z axis. Also show that the trajectories of constant angular momentum are elliptic orbits in the xy plane.

Exercise 1.20. Moments of the canonical distribution. Show that the mean $\bar{\mathcal{E}}$, and variance $\text{var}(\mathcal{E})$, of the canonical distribution can be expressed in terms of the partition function $Z = \sum_j e^{-\beta \mathcal{E}_j}$ as

$$\bar{\mathcal{E}} = -\frac{\partial \ln Z}{\partial \beta}, \quad \text{var}(\mathcal{E}) = \frac{\partial^2 \ln Z}{\partial \beta^2}.$$

Exercise 1.21. The diagram of a basic Wheatstone bridge⁵ circuit, shown in Figure 1.1, contains four resistances $\{r_1, r_2, r_3, r_4\}$, a constant voltage input V_{in} , and a voltage V_g , related by

FIGURE 1.1. Wheatstone bridge circuit



$$V_g = \frac{r_1 r_3 - r_2 r_4}{(r_1 + r_2)(r_3 + r_4)} V_{\text{in}}.$$

Given a fixed set of resistors for which $V_g \neq 0$, consider the set

$$K = \{1, (12)(34), (13)(24), (14)(23)\}$$

of permutations of the index set $\{1, 2, 3, 4\}$ and define the function

$$x(\tau) = \frac{r_{\tau 1} r_{\tau 3} - r_{\tau 2} r_{\tau 4}}{(r_{\tau 1} + r_{\tau 2})(r_{\tau 3} + r_{\tau 4})}, \quad \tau \in K.$$

Assume that $V_{\text{in}} = 1$ so that $x(\tau)$ is then the voltage measurement V_g when the resistors in the bridge are permuted according to $\tau \in K$.

- (1) Show that K , together with the operation of composition of functions, is a group, and conclude that x is an example of a scalar function indexed by a finite group;
- (2) Show that x can be written as $x(\tau) = \chi(\tau)x(1)$, where $\chi(\tau) \in \{1, -1\}$ and satisfies

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

for all τ, σ in K . Interpret χ as a one-dimensional representation of K .

Exercise 1.22. Let V indicate the structure defined as the set of all 2×2 matrices with entries in $\{0, 1\}$. There are 16 points in V , called *incidence* or *relation* matrices. Let W indicate the space of all mappings from $\{a, b\}$ to $\{0, 1\}$, with $a \neq b$. Show that there is a one-to-one correspondence between V and $W \times W$. Study the orbits of V under the different actions of S_2 on $W \times W$.

Exercise 1.23. A relation r in a set A is any subset of $A \times A$. The composition $q \star r$ of relations r and q in A is defined as the set of all pairs (x, z) in $A \times A$ such that $(x, y) \in q$ and $(y, z) \in r$ for some $y \in A$. Clearly, any relation in A can be written as a incidence matrix. Indicate by M_r the incidence matrix corresponding to the relation r in A .

Show that $M_{q \star r} = M_q \bullet M_r$, where here \bullet indicates matrix multiplication under Boolean arithmetic (usual multiplication rule and addition rule modified with $1 + 1 = 1$);

⁵e.g., <http://www.efunda.com/designstandards/sensors/methods/>.

Given a permutation τ in A , indicate by $\rho(\tau)$ the permutation matrix associated with τ . Show that $\rho(\tau) \bullet M_r$ is an incidence matrix for all permutation τ in A and all relation r in A ;

If $\varphi(\tau, r) = \rho(\tau) \bullet M_r$, show that

$$\varphi(\tau, \varphi(\sigma, r)) = \varphi(\tau\sigma, r),$$

thus showing that φ is an action of the group of permutations on the structure of all relations in A .

Exercise 1.24. For ω_1 and ω_2 probability laws in $V = S_\ell$, define the convolution $\omega_1 \star \omega_2$ of ω_1 and ω_2 by

$$\omega_1 \star \omega_2 : s \mapsto \sum_{t \in V} \omega_1(ts^{-1})\omega_2(t), \quad s \in V.$$

Show that $\omega_1 \star \omega_2$ is a probability law in V . Is \star associative and commutative? Describe the convolution law when V is restricted to cyclic permutations.

Exercise 1.25. Following Exercise 1.24, given a probability law ω in S_ℓ , define ω' by $\omega'(s) = \omega(s^{-1})$. Show that ω' is a probability law in C^L , and study the properties of the *symmetrized* law $\omega \star \omega'$. Show that the symmetrized uniform law is the triangular law.

Exercise 1.26. Given $x : C^L \mapsto \mathbb{R}$, the (left) symmetrized version \bar{x} of x is the map

$$\bar{x}(s) = \frac{1}{|G|} \sum_{\eta \in G} x(s\eta^{-1}).$$

Let w be a probability law in $V = C^L$ which is permutation symmetric. Then, relative to w , show that $E(\bar{x}) = E(x)$, when the action $s\tau^{-1}$ is transitive (that is, the action generates a single orbit, equal to V), and otherwise,

$$E(\bar{x}) = \sum_i E(x | \mathcal{O}_i) P_w(\mathcal{O}_i),$$

where \mathcal{O}_i are the distinct orbits generated by the action $s\tau^{-1}$. Hint: refer to Matrix (1.9).

Algebraic Methods for Data Analysis

Lecture Notes at EURANDOM - EIDMA, The Netherlands

March 14, 2005

Abstracted from M. Viana *Structured Data - An Introduction to the Study of Symmetry in Applications* ©2005 by M. A. G. Viana

2.1. Introduction

This chapter is an introduction to the elements of linear representations of finite groups that have potential interest to the analysis of structured data. In Chapter 1 we have discussed a number of examples introducing a structure V , such as the mapping space, a group G of symmetries, and a rule φ for composing the symmetries (τ) with the elements $s \in V$. In addition, at each point s of V we measure something, obtaining the data vector $x = (x(s))_{s \in V}$, a point in the data space $\mathcal{V} = \mathbb{R}^V$. The structured data $x, y, \dots \in \mathcal{V}$ are data indexed by V .

Matrix (2.1) describes the frequencies with which the 16 binary sequences in length of four appear in 10 subsequent 200 bp-long regions of BRUCG isolate of the Human Immunodeficiency Virus Type I. The entire virus has a 9229 bp long nucleotide sequence. To locate the sequence in the National Center for Biotechnology Information¹ data base, use the accession number K02013.

(2.1)

| region | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------|----|----|----|----|----|----|----|----|----|----|
| yyyy | 5 | 8 | 3 | 5 | 7 | 8 | 5 | 25 | 16 | 6 |
| uuuu | 52 | 29 | 36 | 35 | 30 | 34 | 44 | 35 | 37 | 17 |
| yuuu | 18 | 16 | 20 | 16 | 20 | 20 | 16 | 18 | 17 | 17 |
| uyuu | 12 | 16 | 19 | 14 | 20 | 14 | 15 | 11 | 16 | 14 |
| uuuu | 15 | 14 | 21 | 17 | 21 | 12 | 13 | 10 | 16 | 12 |
| uuuy | 17 | 16 | 20 | 16 | 20 | 19 | 16 | 18 | 17 | 17 |
| yyuu | 16 | 11 | 11 | 10 | 10 | 14 | 12 | 15 | 11 | 15 |
| yuyu | 6 | 12 | 9 | 11 | 6 | 8 | 8 | 2 | 4 | 10 |
| yuuu | 10 | 11 | 9 | 8 | 10 | 8 | 11 | 8 | 10 | 11 |
| uyuu | 11 | 14 | 10 | 11 | 8 | 12 | 14 | 10 | 11 | 15 |
| uyuy | 9 | 10 | 11 | 14 | 7 | 6 | 6 | 1 | 4 | 9 |
| uuyy | 12 | 14 | 8 | 8 | 8 | 15 | 14 | 16 | 11 | 16 |
| yyuu | 5 | 6 | 5 | 7 | 8 | 8 | 6 | 10 | 7 | 10 |
| yyuy | 1 | 9 | 4 | 8 | 7 | 7 | 7 | 5 | 7 | 10 |
| yuyu | 4 | 6 | 7 | 11 | 8 | 5 | 5 | 4 | 7 | 9 |
| uyyy | 5 | 6 | 5 | 7 | 8 | 8 | 6 | 10 | 7 | 10 |

¹<http://www.ncbi.nlm.nih.gov/>

Here the symbols $\{u, y\}$ represent the classes of purines $\{a, g\}$ and pyrimidines $\{c, t\}$, respectively, translated from the original sequence written with the $\{a, g, c, t\}$ alphabet.

| | | | | | | | | | | | | | | | | | | |
|---------------------|-------------|-------|-------|-------|-------|----|----|----|----|----|---|---|---|---|----|---|-----|--|
| s(1) | y u y | u u u | y y y | u u u | y y y | u | | | | | | | | | | | | |
| s(2) | y u u | y u u | y u u | y y u | y y u | y | | | | | | | | | | | | |
| s(3) | y u u | u y u | u y u | y u y | y u y | y | | | | | | | | | | | | |
| s(4) | y u u | u u y | u u y | u y y | u y y | y | | | | | | | | | | | | |
| label \rightarrow | 1 | 16 | 15 | 14 | 12 | 8 | 13 | 11 | 7 | 10 | 6 | 4 | 9 | 5 | 3 | 2 | fix | |
| $\mu = 1111$ | | | | | | | | | | | | | | | | | | |
| 1 | 1 16 15 | 14 | 12 | 8 | 13 | 11 | 7 | 10 | 6 | 4 | 9 | 5 | 3 | 2 | 16 | | | |
| $\mu = 2110$ | | | | | | | | | | | | | | | | | | |
| (34) | 1 16 15 | 14 | 8 | 12 | 13 | 7 | 11 | 6 | 10 | 4 | 5 | 9 | 3 | 2 | 8 | | | |
| (23) | 1 16 15 | 12 | 14 | 8 | 11 | 13 | 7 | 10 | 4 | 6 | 9 | 3 | 5 | 2 | 8 | | | |
| (24) | 1 16 15 | 8 | 12 | 14 | 7 | 11 | 13 | 4 | 6 | 10 | 3 | 5 | 9 | 2 | 8 | | | |
| (12) | 1 16 14 | 15 | 12 | 8 | 13 | 10 | 6 | 11 | 7 | 4 | 9 | 5 | 2 | 3 | 8 | | | |
| (13) | 1 16 12 | 14 | 15 | 8 | 10 | 11 | 4 | 13 | 6 | 7 | 9 | 2 | 3 | 5 | 8 | | | |
| (14) | 1 16 8 | 14 | 12 | 15 | 6 | 4 | 7 | 10 | 13 | 11 | 2 | 5 | 3 | 9 | 8 | | | |
| $\mu = 3100$ | | | | | | | | | | | | | | | | | | |
| (234) | 1 16 15 | 12 | 8 | 14 | 11 | 7 | 13 | 4 | 10 | 6 | 3 | 9 | 5 | 2 | 4 | | | |
| (243) | 1 16 15 | 8 | 14 | 12 | 7 | 13 | 11 | 6 | 4 | 10 | 5 | 3 | 9 | 2 | 4 | | | |
| (123) | 1 16 14 | 12 | 15 | 8 | 10 | 13 | 6 | 11 | 4 | 7 | 9 | 2 | 5 | 3 | 4 | | | |
| (124) | 1 16 14 | 8 | 12 | 15 | 6 | 10 | 13 | 4 | 7 | 11 | 2 | 5 | 9 | 3 | 4 | | | |
| (132) | 1 16 12 | 15 | 14 | 8 | 11 | 10 | 4 | 13 | 7 | 6 | 9 | 3 | 2 | 5 | 4 | | | |
| (134) | 1 16 12 | 14 | 8 | 15 | 10 | 4 | 11 | 6 | 13 | 7 | 2 | 9 | 3 | 5 | 4 | | | |
| (142) | 1 16 8 | 15 | 12 | 14 | 7 | 4 | 6 | 11 | 13 | 10 | 3 | 5 | 2 | 9 | 4 | | | |
| (143) | 1 16 8 | 14 | 15 | 12 | 6 | 7 | 4 | 13 | 10 | 11 | 5 | 2 | 3 | 9 | 4 | | | |
| $\mu = 2200$ | | | | | | | | | | | | | | | | | | |
| (12)(34) | 1 16 14 | 15 | 8 | 12 | 13 | 6 | 10 | 7 | 11 | 4 | 5 | 9 | 2 | 3 | 4 | | | |
| (13)(24) | 1 16 12 | 8 | 15 | 14 | 4 | 11 | 10 | 7 | 6 | 13 | 3 | 2 | 9 | 5 | 4 | | | |
| (14)(23) | 1 16 8 | 12 | 14 | 15 | 4 | 6 | 7 | 10 | 11 | 13 | 2 | 3 | 5 | 9 | 4 | | | |
| $\mu = 4000$ | | | | | | | | | | | | | | | | | | |
| (1234) | 1 16 14 | 12 | 8 | 15 | 10 | 6 | 13 | 4 | 11 | 7 | 2 | 9 | 5 | 3 | 2 | | | |
| (1243) | 1 16 14 | 8 | 15 | 12 | 6 | 13 | 10 | 7 | 4 | 11 | 5 | 2 | 9 | 3 | 2 | | | |
| (1324) | 1 16 12 | 8 | 14 | 15 | 4 | 10 | 11 | 6 | 7 | 13 | 2 | 3 | 9 | 5 | 2 | | | |
| (1342) | 1 16 12 | 15 | 8 | 14 | 11 | 4 | 10 | 7 | 13 | 6 | 3 | 9 | 2 | 5 | 2 | | | |
| (1432) | 1 16 8 | 15 | 14 | 12 | 7 | 6 | 4 | 13 | 11 | 10 | 5 | 3 | 2 | 9 | 2 | | | |
| (1423) | 1 16 8 | 12 | 15 | 14 | 4 | 7 | 6 | 11 | 10 | 13 | 3 | 2 | 5 | 9 | 2 | | | |
| $ G_s $ | 24 | 24 | 6 | 6 | 6 | 6 | 4 | 4 | 4 | 4 | 4 | 4 | 6 | 6 | 6 | 6 | | |

These 16 sequences, introduced earlier on in Section 1.6.1 of Chapter 1, identify the structure V of interest as the mapping space all binary sequences in length of four, upon which the group S_4 acts on the position of the letters.

The structured data are the frequencies $x(s)$ with which the points $s \in V$ appear in each 200-bp long region of the isolate, as shown in the columns of Matrix (2.1). Figures 2.1, 2.2, 2.3 and 2.4 show the relative frequency distributions of these 16 words along the 45 adjacent 200-bp long regions of the isolate. Figure 2.1 includes the two single-element orbits. The range, $(0 - 0.3)$, of the y-axis is common to all graphs. The x-axis indicates the number of the consecutive 45 regions.

With the language introduced in the previous chapter, each column in Matrix 2.1 is a data vector in $\mathcal{V} = \mathbb{R}^{16}$ indexed by V . The symmetry transformations act on $s \in V$ according to $s\tau^{-1}$

determining the location symmetry orbits, summarized in Table 2.2. For example, when $\tau = (234)$ and $s = \text{yyuu}$ (label 13) we have $s\tau^{-1} = \text{yuyu}$ (label 11). We will refer to Matrix (2.2) several times over again to illustrate many aspects of interest to the analysis of structured data. The symbols λ , G_s and $|\text{fix}|$ in Matrix (2.2) will be defined later on in the chapter.

In this chapter we will introduce the algebraic elements necessary for the analysis of structured data as introduced in Section 1.9 of Chapter 1. These elements will determine the canonical projections in the data space \mathcal{V} and their invariants as determined by the structure V , the symmetries of interest, and the way with which those symmetries act on V .

FIGURE 2.1. Distribution of yyyy and uuuu along the BRU isolate.

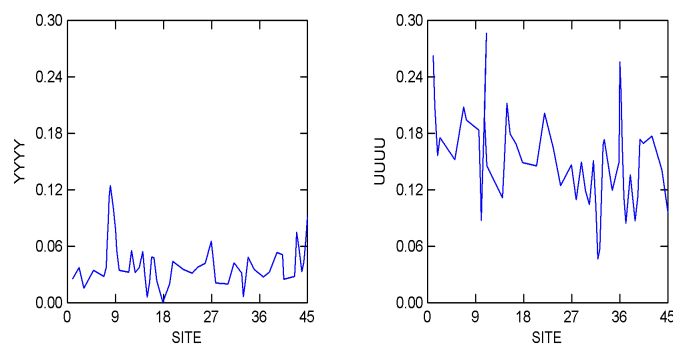


FIGURE 2.2. Distribution of the words in the position symmetry orbit $\mathcal{O}_{\text{yuuu}}$ along the BRU isolate.

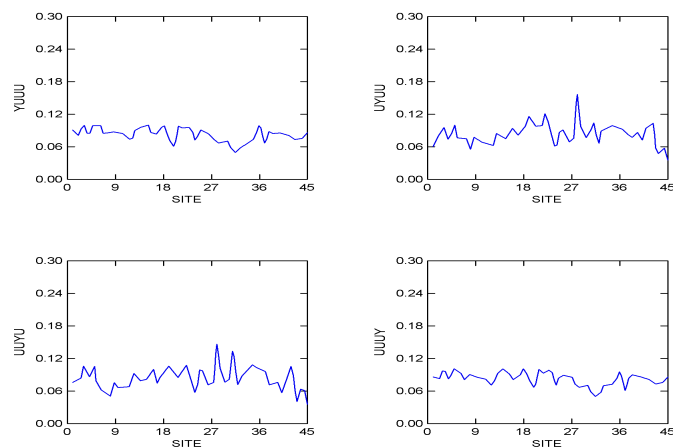


FIGURE 2.3. Distribution of the words in the position symmetry orbit \mathcal{O}_{yuyu} along the BRU isolate.

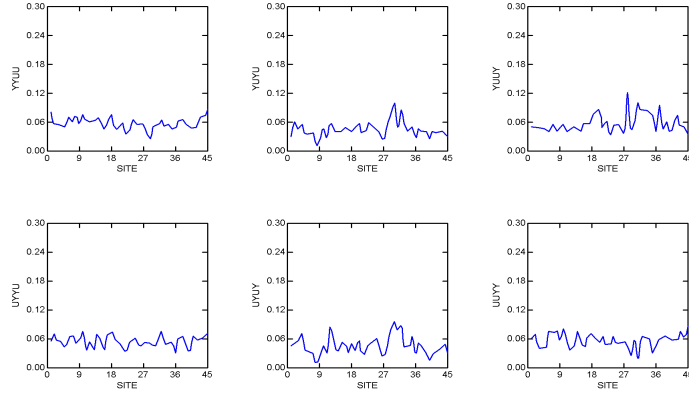
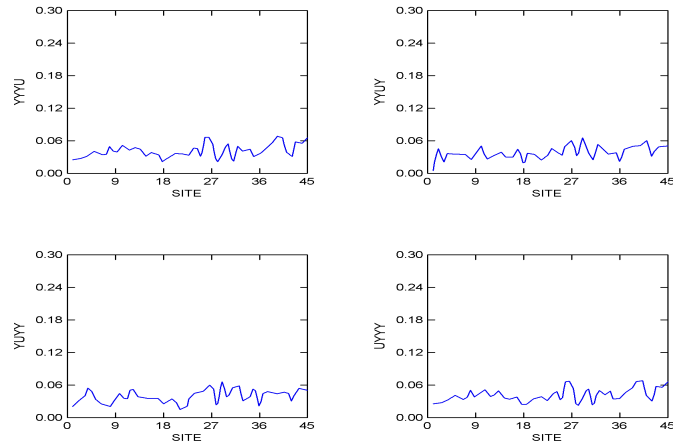


FIGURE 2.4. Distribution of the words in the position symmetry orbit \mathcal{O}_{yyyu} along the BRU isolate.



2.2. Permutations, groups and homomorphisms

Recall, from Section 1.6 in Chapter 1, that we indicate by C^L the set of all mappings s defined on L with values in C , where S and L are finite sets, and by S_L the set of all bijective mappings, or permutations, defined on the set L . In particular, when $L = \{1, 2, \dots, \ell\}$, we write S_ℓ to indicate

these $\ell!$ permutations. In S_3 , the $3! = 6$ permutations are indicated with the notation

$$(12) \equiv \begin{bmatrix} 1 & \rightarrow & 2 \\ 2 & \rightarrow & 1 \\ 3 & \rightarrow & 3 \end{bmatrix}, \quad (13) \equiv \begin{bmatrix} 1 & \rightarrow & 3 \\ 2 & \rightarrow & 3 \\ 3 & \rightarrow & 1 \end{bmatrix}, \quad (23) \equiv \begin{bmatrix} 1 & \rightarrow & 1 \\ 2 & \rightarrow & 3 \\ 3 & \rightarrow & 2 \end{bmatrix},$$

$$(123) \equiv \begin{bmatrix} 1 & \rightarrow & 2 \\ 2 & \rightarrow & 3 \\ 3 & \rightarrow & 1 \end{bmatrix}, \quad (132) \equiv \begin{bmatrix} 1 & \rightarrow & 3 \\ 2 & \rightarrow & 1 \\ 3 & \rightarrow & 2 \end{bmatrix}, \quad 1 \equiv \begin{bmatrix} 1 & \rightarrow & 1 \\ 2 & \rightarrow & 1 \\ 3 & \rightarrow & 3 \end{bmatrix}.$$

The composition $\tau\sigma$ of two permutations τ and σ is the permutation obtained by first applying σ followed by τ , e.g., if $\tau = (23)$ and $\sigma = (13)$, then $\tau\sigma = (123)$. Note that $\sigma\tau = (132)$, so that the composition of permutations is not commutative in general. Permutations such as

$$(12) \equiv 1 \rightarrow 2 \rightarrow 1, \quad \text{or} \quad (132) \equiv 1 \rightarrow 3 \rightarrow 2 \rightarrow 1$$

are called *cyclic* permutations. Every permutation can be decomposed as the product of disjoint cyclic permutations. For example,

$$(2.3) \quad \sigma = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 8 & 7 & 6 & 1 & 2 & 5 & 4 \end{bmatrix} = (1375)(2846).$$

Equivalently, $\sigma = (2846)(1375)$, a consequence of the fact that the composition of any two disjoint (or without common elements) cyclic permutations is commutative. A 2-element cyclic permutation, e.g., (13) , is called a *transposition*. Note that

$$(2.4) \quad (1375) = (15)(17)(13), \quad (2846) = (26)(24)(28).$$

In general, every cyclic permutation decomposes as a (non-commuting) product of transpositions. This is true for all two-cycles. Assuming that it holds for cycles of length $n - 1$, direct evaluation shows that

$$(12 \dots n) = (1n)(12 \dots n - 1),$$

thus proving that the stated decomposition holds for cycles of length n .

From 2.3 and 2.4 we obtain

$$\sigma = (1375)(2846) = (15)(17)(13)(26)(24)(28),$$

and observe that σ , of length 8, decomposes as the product of 2 disjoint cycles. These two numbers are sufficient to characterize the permutation. The difference $8 - 2 = 6$ is called the *decrement* and corresponds to the number of transpositions expressing σ . To see this, note that every cycle of length n is the product of $n - 1$ transpositions,

$$(12 \dots n) = (1n)(1 \ n - 1) \dots (12),$$

so that a permutation $\sigma = C_1 \dots C_h$ of length m with h disjoint cycles of length n_1, \dots, n_h can be written as the product of

$$\sum_{i=1}^h (n_i - 1) = m - h = \text{decrement } (\sigma)$$

transpositions. Since the parity of a permutation is defined as the parity of its decrement, it follows that an even (respectively odd) permutation is the product of an even (respectively odd) number of transpositions.

Proposition 2.2.1. For all permutations σ and transpositions τ ,

$$\text{Decrement } (\tau\sigma) = \text{Decrement } (\sigma) \pm 1.$$

PROOF. Following Bacry (1963), there are two cases to consider: When $\{a, b\}$ belong to a common cycle of σ , the fact that

$$(ab)(a \dots xb \dots y) = (a \dots x)(b \dots y),$$

shows that the number of cycles of σ is increased by one unit, so that the decrement of σ decreases by one unit; Similarly, when $\{a, b\}$ belong to a different cycle, the fact that

$$(a \dots xb \dots y) = (ab)(a \dots x)(b \dots y)$$

shows that the number of cycles is decreased by one unit, and consequently the decrement of σ is increased by one unit. In both cases, because disjoint cycles commute, it is always possible to write the transposition to the left of the cycle(s) of interest. \square

Definition 2.2.1. The sign (Sgn), or signature, of a permutation is given by

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

As a consequence of Proposition 2.2.1, the reader may verify that the parity of the product $\sigma\tau$ of any two permutations σ, τ is given by

$$\left[\begin{array}{c|cc} \sigma \setminus \tau & \text{even} & \text{odd} \\ \hline \text{even} & \text{even} & \text{odd} \\ \text{odd} & \text{odd} & \text{even} \end{array} \right],$$

and that, consequently,

$$(2.5) \quad \text{Sgn}(\sigma\tau) = \text{Sgn}(\sigma)\text{Sgn}(\tau)$$

for any two permutations σ, τ .

Definition 2.2.2. Two permutations σ and η are conjugate when there is a permutation τ such that $\sigma = \tau\eta\tau^{-1}$ for some permutation τ .

The reader may verify that the relation $\sigma \sim \eta$ defined by conjugacy is an equivalence relation. The resulting classes are called *conjugacy classes*. Clearly, Definition 2.2.2 applies to any group in general.

The reader may also observe that the only effect the operation of conjugacy has on the cycle structure of a permutation is that of eventually renaming the elements within each cycle. For example, in S_3 , if $\eta = (12)$ and $\tau = (23)$, then the conjugacy $\tau\eta\tau^{-1}$ transforms the cycle (12) into the cycle (13), that is, $\tau\eta\tau^{-1} = (13)$. We then have:

Proposition 2.2.2. Conjugate permutations have the same cycle structure.

2.2.1. Integer partitions and Young frames. In view of Proposition 2.2.2, the conjugacy classes of permutations groups S_ℓ are naturally associated with the integer partitions

$$\lambda = (n_1, n_2, \dots, n_\ell), \quad n_1 \geq \dots \geq n_\ell \geq 0, \quad n_1 + \dots + n_\ell = \ell$$

of ℓ , describing the permutation's cycle structure, and their corresponding Young frames. To illustrate, consider the case $n = 3$. The partitions, cycle structure and corresponding conjugacy class representatives are: $\lambda = (1, 1, 1)$ indicating the cycle structure with 3 cycles of length 1. We also write $\lambda = 1^3$, with 1 as a representative member; $\lambda = (2, 1, 0)$, or $2^1 1^1$ indicating the cycle structure with 1 cycle of length 2 and one of length 1. A representative is (12); $\lambda = (3, 0, 0)$, or $3^1 0^2$, having one cycle of length 3. A representative is the element (123).

The associated Young frames are

$$(1, 1, 1) = \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array}, \quad (2, 1, 0) = \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \\ \hline \end{array}, \quad (3, 0, 0) = \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \end{array},$$

with n_1, \dots, n_ℓ boxes in rows $1, \dots, \ell$ respectively. Indicating by (m_1, \dots, m_k) the multiplicities with which the k distinct components of each frame occur, each frame can be written, uniquely, as

$$\lambda = (a_1^{m_1}, \dots, a_k^{m_k}), \quad a_1 > a_2 > \dots \geq 0.$$

Definition 2.2.3. A group is a nonempty set G equipped with an associative operation $(\sigma, \tau) \in G \times G \rightarrow \sigma\tau \in G$, and an element $1 \in G$, satisfying:

- (1) $1\tau = \tau 1$, for all $\tau \in G$;
- (2) for every $\tau \in G$, there is an element $\tau^{-1} \in G$ such that $\tau\tau^{-1} = \tau^{-1}\tau = 1$.

A commutative group is one in which the operation is commutative (the term Abelian is also common in the literature). A subset of G which is a group under the group operation of G is called a subgroup of G .

Example 2.2.1 (Permutation groups). The set S_ℓ , together with the operation of mapping composition, defines the group of permutations on the integers $\{1, \dots, \ell\}$. Similarly, S_L together with the operation of composition of functions is a finite group. S_3 is a permutation group of order $3! = 6$ (the number of elements in the group). The resulting multiplication table, or Cayley table, for S_3 is shown in Matrix (2.6), where the permutations are indicated by the symbols on the first column.

(2.6)

| | | | | | | | | |
|-----------|---|---|---|---|---|---|---|---|
| | * | | a | b | c | d | e | f |
| a = 1 | | a | b | c | d | e | f | |
| b = (12) | | b | a | e | f | c | d | |
| c = (13) | | c | f | a | e | d | b | |
| d = (23) | | d | e | f | a | b | c | |
| e = (123) | | e | d | b | c | f | a | |
| f = (132) | | f | c | d | b | a | e | |

In Table (2.6) as in Table (1.1) of Chapter 1, we adopt the convention that when both the algebraic (group multiplication) and analytic (function composition) are to be distinguished, the group operation is then represented by the $*$ symbol, and the function composition by simple juxtaposition. For example, then, $f * b = c$ as group operation of row $*$ column, whereas $c = bf$ indicates the function composition f followed by b . Otherwise, the algebraic interpretation for a notation such as $\sigma\tau$ takes precedence. \square

Example 2.2.2 (Cyclic groups). Indicate by \mathbb{Z} the set of integers and by $p\mathbb{Z}$ the set $\{\dots - p, 0, p, \dots\}$ of integer multiples of p , where p is any positive integer. Note that $p\mathbb{Z}$ is a subgroup of the additive group $(\mathbb{Z}, +)$. The sets

$$\mathcal{O}_k = \{k + z \pmod p; z \in \mathbb{Z}\} \equiv k + p\mathbb{Z}, \quad k = 0, 1, \dots, p-1,$$

of residues modulo p are the *orbits* or *cosets* of $p\mathbb{Z}$ in \mathbb{Z} , and decompose \mathbb{Z} into the disjoint union

$$\mathbb{Z} = \mathcal{O}_0 \cup \dots \cup \mathcal{O}_{p-1}.$$

Direct calculation shows that the set of orbits, together with the operation

$$\mathcal{O}_m + \mathcal{O}_n = \mathcal{O}_{m+n \pmod p}$$

form a group, called the *quotient group*, and denoted by $\mathbb{Z}/p\mathbb{Z}$. The quotient group can be generated by any one of its points by taking successive powers, that is,

$$\mathcal{O}_m, 2\mathcal{O}_m, \dots, (p-1)\mathcal{O}_m$$

generates $\mathbb{Z}/p\mathbb{Z}$, for any \mathcal{O}_m in $\mathbb{Z}/p\mathbb{Z}$. When this property verifies we say that the group is *cyclic* and that \mathcal{O}_m is a group generator. Clearly, then, cyclic groups are commutative. The cyclic group $C_3 = \{1, (123), (132)\}$ is a commutative subgroup of S_3 , of order 3 generated, for example, by (123) . \square

A disjoint decomposition of the form $\mathbb{Z} = \mathcal{O}_0 \cup \dots \cup \mathcal{O}_{p-1}$ is useful in the evaluation of sums (such as expectations) of data $x(z)$ indexed by \mathbb{Z} . If $x: \mathbb{Z} \rightarrow \mathbb{R}$, and $\Gamma = \{0, 1, \dots, p-1\}$ then

$$\sum_{\mathbb{Z}} x(z) = \sum_{k \in \Gamma} \sum_{z \in \mathcal{O}_k} x(z) = \sum_{\Gamma} \sum_{z \in \mathbb{Z}} x(k + pz),$$

provided that the sum over \mathbb{Z} is well-defined. Here, Γ is known as a cross-section in \mathbb{Z} . It intercepts each orbit in exactly one point. In particular, if x is constant in each orbit of the group $p\mathbb{Z}$, taking the value x_k in \mathcal{O}_k , then the expected value $E(x)$ of x can be then evaluated as

$$E(x) = \sum_{\Gamma} \sum_{z \in \mathbb{Z}} x(k + pz)P(k + pz) = \sum_{\Gamma} x_k \sum_{z \in \mathbb{Z}} P(k + pz) = \sum_{\Gamma} x_k P(\mathcal{O}_k),$$

where $P(\mathcal{O}_k)$ is the probability of orbit \mathcal{O}_k under the underlying probability model P for $z \in \mathbb{Z}$. \square

Example 2.2.3 (Matrix groups). The set $GL(\mathbb{F}, n)$ of $n \times n$ nonsingular (i.e., nonzero determinant) matrices with entries in the field \mathbb{F} of scalars is a group under the operation of matrix multiplication. It is usually called the *general linear group*. Equivalently, we may consider the underlying vector space $\mathcal{V} = \mathbb{F}^n$ and write $GL(\mathcal{V})$ to indicate the general linear group. In the present text, it will be understood that $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. Note that for $n \geq 2$ these groups are non-commutative. Some classical subgroups: the set M_n of all $n \times n$ permutation matrices is a finite subgroup of $GL(\mathcal{V})$, of order $n!$; the special linear (or unimodular) group of all $n \times n$ matrices of determinant one, denoted by $SL(n)$; the proper orthogonal group of all $n \times n$ matrices $r \in SL(n)$ for which $rr' = I$, denoted by $O^+(n)$. For $n = 2$,

$$r = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}.$$

The proper affine group in the line, is defined by

$$\left\{ \begin{bmatrix} x & y \\ 0 & 1 \end{bmatrix}, x > 0 \right\}.$$

The step transformation group is the set of all matrices of the form

$$\left\{ \begin{bmatrix} x & y \\ 0 & x \end{bmatrix}, x > 0 \right\}.$$

\square

The convex hull of M_n (all permutation matrices in dimension n) is the set of all doubly stochastic matrices in dimension n . Given a doubly stochastic matrix P and a vector y satisfying $y_1 \geq \dots \geq y_n$ then the new vector $x = Py$ is said to *majorize* y , $x \succ y$, in the sense that $x_1 \geq \dots \geq x_n$ and

$$\sum_{j=1}^n x_j = \sum_{j=1}^n y_j, \quad \sum_{j=1}^m x_j \geq \sum_{j=1}^m y_j, \quad m = 1, \dots, n-1.$$

In fact, it can be shown that $x \succ y \iff x = Py$. For a detailed discussion of majorization and further references see Marshall and Olkin (1979).

Example 2.2.4 (The group of the quaternions). Given the vector $x \in \mathbb{R}^3$, the matrix

$$X = X(x) = \begin{bmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 \end{bmatrix} \in \mathbb{C}^{2 \times 2}$$

is called the matrix associated with the vector x . These matrices have several remarkable properties, e.g., Cartan (1966)[p.43]. In particular, as the reader may verify, for all $x, y \in \mathbb{R}^3$, that

$$\det X = -||x||^2; \quad X^2 = ||x||^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \quad XY + YX = 2x'y;$$

$$i(XY - YX)/2 = \text{matrix associated with (the bivector) } x \times y.$$

In particular if we consider the matrices

$$H_1 = X(e_1) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad H_2 = X(e_2) = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad H_3 = X(e_3) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

associated with the basis vectors $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ and $e_3 = (0, 0, 1)$, we observe that $H_1^2 = H_2^2 = H_3^2$,

$$H_1H_2 + H_2H_1 = H_1H_3 + H_3H_1 = H_2H_3 + H_3H_2 = 0,$$

and $H_1H_2H_3 = iH_0$, where $H_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Moreover, because

$$\sum_{\ell=0}^3 a_\ell H_\ell = \begin{bmatrix} a_0 + a_3 & a_1 - ia_2 \\ a_1 + ia_2 & a_0 - a_3 \end{bmatrix}$$

we conclude that there is no linear relation of the form $\sum_{\ell=0}^3 a_\ell H_\ell = 0$ with complex coefficients unless all these coefficients are zero. That is, any complex 2×2 matrix can be expressed uniquely as the sum of a scalar matrix $a_0 H_0$ and the matrix associated with a vector.

Consequently, we can identify H_0, \dots, H_3 with a basis for a four-dimensional vector space \mathcal{H} on which a vector multiplication has been defined. For all $f, g, h \in \mathcal{H}$ and all scalars α in the field of the vector space, we have $fg \in \mathcal{H}$, $f(g+h) = fg+fh$, $(f+g)h = fh+gh$, $\alpha(fg) = f(\alpha g) = (\alpha f)g$. That is, \mathcal{H} constitute an algebra. When expressed in terms of $I_0 = H_0$, $I_1 = -iH_1$, $I_2 = -iH_2$, $I_3 = -iH_3$, the algebra is known as Hamilton's algebra² of the quaternions. This algebra is associative but not commutative. The elements $1 \equiv H_0$, $i \equiv iH_3$, $j \equiv iH_2$ and $k = iH_1$ satisfy the relations

$$ij = -ji = k, \quad jk = -kj = i, \quad ik = -ki = -j, \quad i^2 = j^2 = k^2 = -1,$$

and $\{\pm 1, \pm i, \pm j, \pm k\}$ form a group known as the group of the quaternions.

We observe that

$$\det\left[\sum_{\ell=0}^3 a_\ell H_\ell\right] = \det \begin{bmatrix} a_0 + a_3 & a_1 - ia_2 \\ a_1 + ia_2 & a_0 - a_3 \end{bmatrix} = a_0^2 - a_1^2 - a_2^2 - a_3^2 = -\|a\|^2,$$

where in the above expression $\|\cdot\|$ indicates the norm defined by the (Lorentz) fundamental form. Whether or not $\|a\|^2 = 0$ has non-zero solutions depends on the field of scalars. In particular, when $a_0 = 0$, there are infinitely many complex (isotropic vectors) solutions to $\|a\|^2 = 0$, whereas only $(0, 0, 0, 0)$ is a solution over the real field. For each non-zero real a_0 the real solutions to $\|a\|^2 = 0$ transform as the full group of rotations in three dimensions. \square

Definition 2.2.4. Given two groups G, H , a homomorphism from G to H is a function $\rho : G \rightarrow H$ preserving the group structure, that is,

$$\rho(\tau\sigma) = \rho(\tau)\rho(\sigma), \quad \text{for all } \tau, \sigma \in G.$$

Note that if ρ is a homomorphism of G then $\rho(1) = 1$ and $\rho(\tau^{-1}) = \rho(\tau)^{-1}$ for all $\tau \in G$. An injective homomorphism is called a monomorphism. An isomorphism is an invertible homomorphism. When $G = H$, the isomorphism ρ is called an automorphism in H . The kernel of the homomorphism ρ is the set of those elements in G mapped into the identity element of H , that is, $\ker_\rho = \{\tau \in G; \rho(\tau) = 1\}$, whereas its range or image is the set $\text{im}_\rho = \{\rho(\tau); \tau \in G\}$.

Note that when $\ker_\rho = \{1\}$ then the homomorphism ρ is an isomorphism onto its image in H . In fact, $\rho(\tau) = \rho(\sigma)$ implies

$$\rho(\tau\sigma^{-1}) = \rho(\tau)\rho(\sigma^{-1}) = \rho(\sigma)\rho(\sigma^{-1}) = \rho(1) = 1,$$

so that $\tau\sigma^{-1} \in \{1\}$, or, $\sigma = \tau$.

Example 2.2.5. The permutation group S_ℓ is isomorphic to the group M_ℓ of $\ell \times \ell$ permutation matrices: to each permutation τ in S_ℓ we associate the permutation matrix $r(\tau)$ in M_ℓ representing the changing

$$\{e_1, e_2, \dots, e_\ell\} \rightarrow \{e_{\tau 1}, e_{\tau 2}, \dots, e_{\tau \ell}\}$$

in the canonical basis $\{e_1, e_2, \dots, e_\ell\}$ of \mathbb{R}^ℓ . We then have $r(\tau * \sigma) = r(\tau)r(\sigma)$, for all $\tau, \sigma \in S_\ell$. In S_3 , for instance, $(123) = (13) * (23)$ and

$$r[(123)] = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = r[(13)]r[(23)].$$

\square

²William Rowan Hamilton, 1805-1865.

Example 2.2.6. The proper affine group in the line described in Example 2.2.3 is isomorphic to the transformations group

$$t \rightarrow xt + y, \quad x > 0,$$

thus justifying its name. More specifically, as the reader may verify, if $F_i = \begin{bmatrix} x_i & y_i \\ 0 & 1 \end{bmatrix}$, $x_i > 0$, $f_i(t) = x_i t + y_i$ and $\xi(f_i) = F_i$, $i = 1, 2$, then $f_1 f_2(t) = f_1(f_2(t)) = x_1 x_2 t + x_1 y_2 + y_1$, so that

$$\xi(f_1 f_2) = \xi(f_1) \xi(f_2).$$

The reader may also verify that ϕ maps the identity transformation $t \rightarrow t$ into the 2×2 identity matrix I . Moreover, by showing that any point $t \rightarrow tx + y$ can be smoothly connected to the identity transformation, one shows that any point in the proper affine group can be smoothly connected to the identity matrix I . The (continuous) group in the neighborhood of I is called a *Lie Group* (See also Exercise 2.4). \square

Example 2.2.7 (Normal subgroups). Fix any member τ of a group G and define the mapping $i_\tau : G \rightarrow G$ by $i_\tau(\sigma) = \tau \sigma \tau^{-1}$. Then, for every $\tau \in G$, the mapping i_τ is an isomorphism in G , and the mapping $\tau \mapsto i_\tau$ is a homomorphism of G , taking values in the set $\text{Aut}(G)$ of automorphisms in G . The mapping i_τ is usually called the *conjugation by τ* , or the *inner automorphism* generated by τ . A subgroup N of G satisfying the property $\tau N \tau^{-1} \subset N$ for all $\tau \in G$ is called a *normal* or *stable* subgroup of G . That is, a normal subgroup contains the complete conjugacy classes of all its elements. $p\mathbb{Z}$ is a normal subgroup of $(\mathbb{Z}, +)$. Homomorphism kernels are normal subgroups. \square

Definition 2.2.5. Given two groups N and H , let α be a homomorphism from H to $\text{Aut}(N)$. For (τ, σ) and (τ_1, σ_1) in $N \times H$, define the operation

$$(\tau, \sigma) \times_\alpha (\tau_1, \sigma_1) = (\tau \alpha(\sigma)(\tau_1), \sigma \sigma_1).$$

Then, $G = N \times H$, together with \times_α , is a group, called the *semidirect product* of N and H under α . The reader may want to verify that $1_G = (1_N, 1_H)$ and that

$$(\tau, \sigma)^{-1} = (\alpha(\sigma^{-1})(\tau^{-1}), \sigma^{-1}).$$

The direct product of two groups is obtained when $\alpha(\sigma) \equiv 1$ for all $\sigma \in H$ in the above definition.

Example 2.2.8. Let's evaluate the direct product $G = C_3 \times C_2$. Write $C_2 = \{1, \tau\}$ and $C_3 = \{1, \sigma, \sigma^2\}$. Direct evaluation shows that the resulting multiplication table for G is

| $C_3 \times C_2$ | 1 | 2 | 3 | 4 | 5 | 6 |
|------------------------|---|---|---|---|---|---|
| $1 = (1, 1)$ | 1 | 2 | 3 | 4 | 5 | 6 |
| $2 = (\sigma, 1)$ | 2 | 3 | 1 | 5 | 6 | 4 |
| $3 = (\sigma^2, 1)$ | 3 | 1 | 2 | 6 | 4 | 5 |
| $4 = (1, \tau)$ | 4 | 5 | 6 | 1 | 2 | 3 |
| $5 = (\sigma, \tau)$ | 5 | 6 | 4 | 2 | 3 | 1 |
| $6 = (\sigma^2, \tau)$ | 6 | 4 | 5 | 3 | 1 | 2 |

where the elements of G are indicated according to the first column of the table. Comparing the first and second rows,

$$\left[\begin{array}{c|cccccc} 1 & 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 2 & 3 & 1 & 5 & 6 & 4 \end{array} \right]$$

of the table, we observe that this corresponds to the permutation (123)(456). Similarly, following with the remaining rows we obtain the equivalence

| | | |
|----------------------|--|--------------|
| (1, 1) | | 1 |
| (σ, 1) | | (123)(456) |
| (σ ² , 1) | | (132)(465) |
| (1, τ) | | (14)(25)(36) |
| (σ, τ) | | (153426) |
| (σ ² , τ) | | (162435) |

The reader now can recognize that these permutations coincide with the cyclic group C₆. We have then, C₃ × C₂ ≃ C₆. It is not difficult to verify that, in general, C_m × C_n ≃ C_{mn}, provided that m and n are relative primes.

Example 2.2.9 (Dihedral groups). The semi-direct product G = C₃ ×_α C₂ defined by the group homomorphism

$$\alpha(\sigma)(\tau) = \begin{cases} \tau & \text{if } \sigma = 1 \\ \tau^{-1} & \text{if } \sigma \neq 1 \end{cases}$$

from C₂ with values in Aut(C₃) follows from the multiplication rule

$$(\tau, \sigma) \times_{\alpha} (\tau_1, \sigma_1) = \begin{cases} (\tau\tau_1, \sigma_1) & \text{when } \sigma = 1 \\ (\tau\tau_1^{-1}, \sigma_1) & \text{when } \sigma = t = (12) \end{cases}$$

in C₃ × C₂, from which we obtain the multiplication table

(2.7)

| × _α | 1 | 2 | 3 | 4 | 5 | 6 |
|--------------------------|---|---|---|---|---|---|
| 1 = (1, 1) | 1 | 2 | 3 | 4 | 5 | 6 |
| 2 = (η, 1) | 2 | 3 | 1 | 5 | 6 | 4 |
| 3 = (η ² , 1) | 3 | 1 | 2 | 6 | 4 | 5 |
| 4 = (1, t) | 4 | 6 | 5 | 1 | 3 | 2 |
| 5 = (η, t) | 5 | 4 | 6 | 2 | 1 | 3 |
| 6 = (η ² , t) | 6 | 5 | 4 | 3 | 2 | 1 |

where t indicates the transposition (12) in C₂. Note that rows 1 and 2 define the permutation (123)(456). Defining the remaining permutations (relative to row 1), we obtain the equivalences

$$\begin{aligned} (1, 1) &\equiv 1 \\ (\eta, 1) &\equiv (123)(456) \equiv (123) \\ (\eta^2, 1) &\equiv (132)(465) \equiv (132) \\ (1, \tau) &\equiv (14)(26)(35) \equiv (23) \\ (\eta, \tau) &\equiv (15)(24)(36) \equiv (12) \\ (\eta^2, \tau) &\equiv (16)(25)(34) \equiv (13). \end{aligned}$$

The sets {1, 4}, {2, 5} and {3, 6} constitute what is known as an imprimitive system and the resulting identification realizes the semi-direct product C₃ ×_α C₂ as the group of symmetry transformations fixing a regular triangle with vertices labeled as {1, 2, 3}. This group is known as the *dihedral group* D₃. Their elements are realized as the rotations {1, (123), (132)} and axial reflections {(12), (23), (13)} fixing the regular triangle.

In general, the dihedral group D_n is obtained as the semi-direct product C_n ×_α C₂. It is the group of symmetries (rotations and the axial reflections) leaving a n-sided regular polygon invariant. Its order is 2n. □

2.3. Group actions and orbits

Definition 2.3.1. Given a set V and a group G , a group action of G on V is a function $\varphi : G \times V \rightarrow V$ satisfying

- (1) $\varphi(1, s) = s$, for all s in V ,
- (2) $\varphi(\sigma, \varphi(\tau, s)) = \varphi(\sigma\tau, s)$, for all $s \in V$, τ, σ in G .

The orbit \mathcal{O}_s of an element $s \in V$ generated by G under the action φ is the set

$$(2.8) \quad \mathcal{O}_s = \{\varphi(\tau, s); \tau \in G\}.$$

We also define the set

$$(2.9) \quad \text{fix}(\tau) = \{s \in V; \varphi(\tau, s) = s\}$$

of elements in V that remain fixed by τ under the action φ , and the set

$$(2.10) \quad G_s = \{\tau \in G; \varphi(\tau, s) = s\},$$

of elements $\tau \in G$ fixing the point $s \in V$. This set is the *stabilizer* of s by G under φ . It is then easy to check that $|G| = |\mathcal{O}_s| |G_s|$. Moreover, note that G_s is a subgroup of G :

- (1) $1 \in G_s$;
- (2) $\tau, \sigma \in G_s$ implies $\varphi(\tau\sigma, s) = \varphi(\tau, \varphi(\sigma, s)) = \varphi(\tau, s) = s$;
- (3) $\tau \in G_s$ implies $s = \varphi(1, s) = \varphi(\tau^{-1}\tau, s) = \varphi(\tau^{-1}, \varphi(\tau, s)) = \varphi(\tau^{-1}, s)$, that is, $\tau^{-1} \in G_s$.

G_s is also called the *isotropy* group of s in G under φ .

When the orbit \mathcal{O}_s of an element $s \in V$ generated by G under the action φ coincides with V we say that the action φ is *transitive*, or that G acts transitively on V .

Example 2.3.1 (Orbits for binary sequences in length of two). Consider the set $V = \{uu, yy, uy, yu\}$ of two-sequences in length of two. That is, V is the set of all mappings s from $\{1, 2\}$ into $\{u, y\}$. The reader can verify that $\varphi_1(\tau, s) = s\tau^{-1}$,

$$\varphi_1 : \{1, 2\} \xrightarrow{\tau^{-1}} \{1, 2\} \xrightarrow{s} \{u, y\}, \quad s \in V, \tau \in S_2 = \{1, (12)\},$$

and $\varphi_2(\sigma, s) = \sigma s$,

$$\varphi_2 : \{1, 2\} \xrightarrow{s} \{u, y\} \xrightarrow{\sigma} \{u, y\}, \quad s \in V, \sigma \in S_2 = \{1, (uy)\},$$

are actions of S_2 on V . Action φ_1 classifies the sequences by symmetries in the positions $\{1, 2\}$ whereas φ_2 classifies the sequences by symmetries in the letters $\{u, y\}$. The evaluations of these actions are summarized in the following matrices:

$$\varphi_1 : \left[\begin{array}{c|cccc} \tau & uu & yy & uy & yu \\ \hline 1 & uu & yy & uy & yu \\ (12) & uu & yy & yu & uy \end{array} \right], \quad \varphi_2 : \left[\begin{array}{c|cccc} \tau \backslash s & uu & yy & uy & yu \\ \hline 1 & uu & yy & uy & yu \\ (12) & yy & uu & yu & uy \end{array} \right].$$

Action φ_1 generates 3 orbits $\{uu\}$, $\{yy\}$ and $\{uy, yu\}$ and φ_2 generates 2 orbits, $\{uu, yy\}$ and $\{uy, yu\}$. The reader may refer to Table (2.2), identify the orbits of S_4 acting on the set V of binary sequences in length of four according to $s\tau^{-1}$ (position symmetry), the isotropy groups and fixed points. \square

Example 2.3.2 (Cyclic orbits for binary sequences in length of four). These orbits are also called *cyclic orbits*. Matrix (2.11) shows the action $s\tau^{-1}$ of $C_4 = \{1, (1234), (13)(24), (1432)\}$ on the mapping space V from Table (2.2).

$$(2.11) \quad \left[\begin{array}{c|cccccccccccc} C_4 \backslash s & 1 & 16 & 15 & 14 & 12 & 8 & 13 & 11 & 7 & 10 & 6 & 4 & 9 & 5 & 3 & 2 \\ \hline 1 & 1 & 16 & 15 & 14 & 12 & 8 & 13 & 11 & 7 & 10 & 6 & 4 & 9 & 5 & 3 & 2 \\ (13)(24) & 1 & 16 & 12 & 8 & 15 & 14 & 4 & 11 & 10 & 7 & 6 & 13 & 3 & 2 & 9 & 5 \\ (1234) & 1 & 16 & 14 & 12 & 8 & 15 & 10 & 6 & 13 & 4 & 11 & 7 & 2 & 9 & 5 & 3 \\ (1432) & 1 & 16 & 8 & 15 & 14 & 12 & 7 & 6 & 4 & 13 & 11 & 10 & 5 & 3 & 2 & 9 \end{array} \right]$$

leads to the orbits (indicating the mappings by their labels)

$$\begin{aligned} \mathcal{O}_0 &= \{1\}, \\ \mathcal{O}_1 &= \{9, 5, 3, 2\}, \\ \mathcal{O}_{21} &= \{13, 7, 10, 4\}, \quad \mathcal{O}_{22} = \{11, 6\}, \\ \mathcal{O}_3 &= \{15, 14, 12, 8\}, \\ \mathcal{O}_4 &= \{16\}. \end{aligned}$$

We note that C_4 splits the original orbit \mathcal{O}_2 under S_4 into two new orbits, \mathcal{O}_{21} and \mathcal{O}_{22} , so that $\mathcal{O}_{21} \cup \mathcal{O}_{22} = \mathcal{O}_2$. Similarly, the action $s\tau^{-1}$ of $G = \{1, (13)(24)\}$ on V is given by

$$(2.12) \quad \left[\begin{array}{c|cccccccccccccccc} G \setminus s & 1 & 16 & 15 & 14 & 12 & 8 & 13 & 11 & 7 & 10 & 6 & 4 & 9 & 5 & 3 & 2 \\ \hline 1 & 1 & 16 & 15 & 14 & 12 & 8 & 13 & 11 & 7 & 10 & 6 & 4 & 9 & 5 & 3 & 2 \\ (13)(24) & 1 & 16 & 12 & 8 & 15 & 14 & 4 & 11 & 10 & 7 & 6 & 13 & 3 & 2 & 9 & 5 \end{array} \right],$$

with corresponding orbits,

$$\begin{aligned} \mathcal{O}_0 &= \{1\}, \\ \mathcal{O}_{11} &= \{9, 3\}, \quad \mathcal{O}_{12} = \{5, 2\}, \\ \mathcal{O}_{211} &= \{13, 4\}, \quad \mathcal{O}_{212} = \{7, 10\}, \quad \mathcal{O}_{221} = \{11\}, \quad \mathcal{O}_{222} = \{6\}, \\ \mathcal{O}_{31} &= \{14, 8\}, \quad \mathcal{O}_{32} = \{15, 12\}, \\ \mathcal{O}_4 &= \{16\}. \end{aligned}$$

The action further splits the original order-4 cyclic orbits into additional, smaller orbits. □

Example 2.3.3 (Dihedral orbits for binary sequences in length of four). Consider the action $s\tau^{-1}$ of the group D_4 , defined earlier on in Example 2.2.9, on the mapping space of binary sequences in length of four. The resulting action is shown in Matrix (2.13):

$$(2.13) \quad \left[\begin{array}{c|cccccccccccccccc} D_4 \setminus s & 1 & 16 & 15 & 14 & 12 & 8 & 13 & 11 & 7 & 10 & 6 & 4 & 9 & 5 & 3 & 2 \\ \hline 1 & 1 & 16 & 15 & 14 & 12 & 8 & 13 & 11 & 7 & 10 & 6 & 4 & 9 & 5 & 3 & 2 \\ (24) & 1 & 16 & 15 & 8 & 12 & 14 & 7 & 11 & 13 & 4 & 6 & 10 & 3 & 5 & 9 & 2 \\ (13) & 1 & 16 & 12 & 14 & 15 & 8 & 10 & 11 & 4 & 13 & 6 & 7 & 9 & 2 & 3 & 5 \\ (12)(34) & 1 & 16 & 14 & 15 & 8 & 12 & 13 & 6 & 10 & 7 & 11 & 4 & 5 & 9 & 2 & 3 \\ (13)(24) & 1 & 16 & 12 & 8 & 15 & 14 & 4 & 11 & 10 & 7 & 6 & 13 & 3 & 2 & 9 & 5 \\ (14)(23) & 1 & 16 & 8 & 12 & 14 & 15 & 4 & 6 & 7 & 10 & 11 & 13 & 2 & 3 & 5 & 9 \\ (1234) & 1 & 16 & 14 & 12 & 8 & 15 & 10 & 6 & 13 & 4 & 11 & 7 & 2 & 9 & 5 & 3 \\ (1432) & 1 & 16 & 8 & 15 & 14 & 12 & 7 & 6 & 4 & 13 & 11 & 10 & 5 & 3 & 2 & 9 \end{array} \right],$$

which shows that D_4 and C_4 generate the same set of orbits under position symmetry. □

Example 2.3.4 (A generic indexing). The set

$$\text{fix}(\tau) = \{s \in V; \varphi(\tau, s) = s\} \subseteq V$$

of elements in V that remain fixed by τ under the action φ of the group G on the set V is useful to define a general method of assigning data to the underlying group acting on V , called the regular indexing. To see this, let x be a scalar measurement in V and to each element $\tau \in G$ associate the evaluation $\underline{x}(\tau)$ of a scalar summary of x defined over $\text{fix}(\tau)$. That is, $\underline{x}(\tau)$ is the summary of the data over those elements in V that share the symmetry of τ . For example, if the summary of interest is the averaging, then

$$\underline{x}(\tau) = \frac{1}{|\text{fix}(\tau)|} \sum_{s \in \text{fix}(\tau)} x(s).$$

In particular, $\underline{x}(\tau) = |\text{fix}(\tau)|$ assigns to τ the volume of points in V with its symmetry: the τ -symmetry content in V .

To illustrate, consider Example 2.3.3 and Matrix (2.13), in which the data are originally indexed by the set V of all binary sequences in length of four. The action is of D_4 on the position of the letters, and the scalar measurements are frequencies $x(s)$ with which these sequences appear

in a parent sequence. Consider the averaging as the summary of interest. Note, from Matrix (2.13) that

$$\text{fix}(\tau) = \begin{cases} V & \text{if } \tau = 1 \\ \{1, 16, 15, 12, 11, 6, 5, 2\} & \text{if } \tau = (24) \\ \{1, 16, 14, 8, 11, 6, 9, 3\} & \text{if } \tau = (13) \\ \{1, 16, 13, 4\} & \text{if } \tau = (12)(34) \\ \{1, 16, 11, 6\} & \text{if } \tau = (13)(24) \\ \{1, 16, 7, 10\} & \text{if } \tau = (14)(23) \\ \{1, 16\} & \text{if } \tau = (1234) \\ \{1, 16\} & \text{if } \tau = (1432), \end{cases}$$

so that, for example,

$$\underline{x}(1) = \frac{1}{16} \sum_{s \in V} x(s),$$

whereas

$$\underline{x}((14)(23)) = \frac{x(1) + x(16) + x(7) + x(10)}{4},$$

and so on. The data \underline{x} are now indexed by D_4 .

Clearly, this construction assumes that $\text{fix}(\tau) \neq \emptyset$ for all τ in the subgroup H of interest. That is, one would search for the largest subgroup in which the property applies. Clearly, when $V = G$ the subgroup H reduces to the identity alone. \square

Example 2.3.5 (Maxwell-Boltzmann and Bose-Einstein counts). Define two mappings s and f in the mapping space $V = C^L$ as equivalent whenever $s\tau^{-1} = f$ for some permutation $\tau \in S_\ell$. That is, s and f differ only by a permutation of the ℓ positions of the their c symbols. The orbits $\{s\tau^{-1}; \tau \in S_\ell\}$ decompose the space V into the disjoint union

$$(2.14) \quad V = \mathcal{O}_{\lambda_1} \cup \dots \cup \mathcal{O}_{\lambda_m},$$

where m is the number of Young frames (Section 2.2.1) or integer partitions $\lambda = (n_1, \dots, n_c)$ of ℓ , and each \mathcal{O}_{λ_i} is a disjoint union of elementary orbits whose members share a particular frame. These elementary orbits define the quotient space V/S_ℓ generated by the action $s\tau^{-1}$ (position symmetry). Permutation orbits generated by the action σs (letter symmetry) are defined similarly.

To illustrate further the orbit decomposition, consider the case in which the mapping space $V = C^L$ represents the possible compositions of an urn with four marbles with labels in the set $L = \{1, 2, 3, 4\}$ and colors in the set

$$C = \{\text{red } (\circ), \text{blue } (\bullet), \text{green } (\diamond)\}.$$

Two urn compositions are defined as equivalent when they differ only by relabeling of the marbles. That is, S_4 acts on V according to $s\tau^{-1}$. Start with the elementary frames: There are $m = 4$ of those, namely,

$$\lambda_1 = (4, 0, 0), \quad \lambda_2 = (3, 1, 0), \quad \lambda_3 = (2, 2, 0), \quad \lambda_4 = (2, 1, 1).$$

Following the notation of Section 2.2.1, we write $\lambda_1 = 40^2$, $\lambda_2 = 310$, $\lambda_3 = 2^20$, $\lambda_4 = 21^2$. In correspondence with equality (2.14), we obtain the decomposition

$$|V| = c^\ell = \sum_{\lambda} \frac{\ell!}{(a_1!)^{m_1} (a_2!)^{m_2} \dots (a_k!)^{m_k}} \frac{c!}{m_1! m_2! \dots m_k!}$$

of the Maxwell-Boltzmann count c^ℓ into the product of the volumes

$$\Omega_\lambda = \frac{\ell!}{(a_1!)^{m_1} (a_2!)^{m_2} \dots (a_k!)^{m_k}}$$

of the elementary orbits in V/S_ℓ and their multiplicities

$$Q_\lambda = \frac{c!}{m_1! m_2! \dots m_k!}.$$

In the above decomposition, λ varies over the m elementary frames, or integer partitions $(a_1^{m_1}, \dots, a_k^{m_k})$ satisfying $m_1 a_1 + \dots + m_k a_k = \ell$ and $m_1 + \dots + m_k = c$. Moreover,

$$\sum_{\lambda} Q_{\lambda} = \binom{c + \ell - 1}{\ell}$$

decomposes the Bose-Einstein count $\binom{c + \ell - 1}{\ell}$ into the sum of the number Q_{λ} of quantal states associated to frame λ . Writing

$$(2.15) \quad v(\lambda) = \Omega_{\lambda} Q_{\lambda}$$

we have $c^{\ell} = \sum_{\lambda} v(\lambda)$. Direct computation leads to

$$|\mathcal{O}_{\lambda_1}| = 3, \quad |\mathcal{O}_{\lambda_2}| = 24, \quad |\mathcal{O}_{\lambda_3}| = 18, \quad |\mathcal{O}_{\lambda_4}| = 36.$$

Matrix (2.16) summarizes the correspondence among frames, orbits, volumes of elementary orbits, multiplicities and urn compositions.

| λ | urn composition | Ω_{λ} | Q_{λ} | $v(\lambda)$ |
|-----------|------------------------------------|--------------------|---------------|--------------|
| 40^2 | $\{\circ \circ \circ \circ\}$ | 1 | 3 | 3 |
| 310 | $\{\circ \circ \circ \bullet\}$ | 4 | 6 | 24 |
| $2^2 0$ | $\{\circ \circ \bullet \bullet\}$ | 6 | 3 | 18 |
| 21^2 | $\{\circ \circ \bullet \diamond\}$ | 12 | 3 | 36 |

We observe that $|V| = 3^4 = 81 = \sum_{\lambda} v(\lambda)$, and that $\sum_{\lambda} Q_{\lambda} = \binom{c + \ell - 1}{\ell} = \binom{6}{4} = 15$. In addition, the volume of the elementary orbit \mathcal{O}_{λ} associated with frame $\lambda = (n_1, \dots, n_c)$ can be decomposed in terms of its subframes of size $c - 1$. In the present example,

$$\begin{aligned} |\mathcal{O}_{\lambda_1}| &= v(4, 0, 0) = \binom{4}{4}v(0, 0) + \binom{4}{0}v(4, 0) = 1 + 2, \\ |\mathcal{O}_{\lambda_2}| &= v(3, 1, 0) = \binom{4}{3}v(1, 0) + \binom{4}{1}v(3, 0) + \binom{4}{0}v(3, 1) = 8 + 8 + 8, \\ |\mathcal{O}_{\lambda_3}| &= v(2, 2, 0) = \binom{4}{2}v(2, 0) + \binom{4}{0}v(2, 2) = 12 + 6, \\ |\mathcal{O}_{\lambda_4}| &= v(2, 1, 1) = \binom{4}{2}v(1, 1) + \binom{4}{1}v(2, 1) = 12 + 24. \end{aligned}$$

The resulting partition of V according to *color-only* attribute is defined by the number $|A_{ij}|$ of configurations in which exactly j marbles have color i . In the present example, $j = 0, 1, 2, 3, 4$, $i \in \{\text{red, blue, green}\}$, and

$$\begin{aligned} |A_{i0}| &= \binom{4}{0}[v(4, 0) + v(3, 1) + v(2, 2)] = 16, \\ |A_{i1}| &= \binom{4}{1}[v(3, 0) + v(2, 1)] = 32, \\ |A_{i2}| &= \binom{4}{2}[v(2, 0) + v(1, 1)] = 24, \\ |A_{i3}| &= \binom{4}{3}v(1, 0) = 8, \\ |A_{i4}| &= \binom{4}{4}v(0, 0) = 1. \end{aligned}$$

If the urn compositions are equally likely, or, equivalently, if the points in V are uniformly distributed, the resulting probabilities

$$w_i(j) = P[\text{exactly } j \text{ marbles have the same color } i]$$

in V/S_4 are

$$(w_i(0), w_i(1), w_i(2), w_i(3), w_i(4)) = \frac{1}{81}(16, 32, 24, 8, 1).$$

□

2.3.1. Burnside's Lemma. The following result evaluates the number of orbits in V generated by the action φ of G as the average number of fixed points of φ . The reader may first refer to Matrix (2.2), which summarizes the action $s\tau^{-1}$ of S_4 on the space V of all two-sequences in length of four and identify, for each $\tau \in G$ the number $|\text{fix}(\tau)|$ of points in V fixed by the action on V , and the number $|G_s|$ of points in G that leave the element $s \in V$ fixed.

Lemma 2.3.1 (Burnside). If a finite group G acts on V according to φ , then

$$\text{Number of orbits in } V = \frac{1}{|G|} \sum_{\tau \in G} |\text{fix}(\tau)|.$$

PROOF. Let $A = \{(\tau, s) \in G \times V; \varphi(\tau, s) = s\}$. First calculate the number $|A|$ of elements in A as

$$(2.17) \quad |A| = \sum_{\tau \in G} |\text{fix}(\tau)|.$$

Secondly, writing $\eta =$ number of orbits in V , evaluate this same number as

$$(2.18) \quad |A| = \sum_{s \in V} |G_s| = \sum_{i=1}^{\eta} |\mathcal{O}_i| |G_{s_i}| = \sum_{i=1}^{\eta} |\mathcal{O}_i| \frac{|G|}{|\mathcal{O}_i|} = \eta \times |G|.$$

From equalities 2.17 and 2.18 the result then follows. \square

The lemma carries the name of William Burnside, Born: 2 July 1852 in London, England. Died: 21 Aug 1927 in West Wickham, London, England. Among his applied mathematics teachers at Cambridge were Stokes, Adams and Maxwell. However, it was actually proved by Frobenius in 1887.

Example 2.3.6. From Matrix (2.2), with $G = S_4$, it follows that

$$\text{Number of orbits of } V = \frac{1}{|G|} \sum_G |\text{fix}(\sigma)| = \frac{120}{24} = 5,$$

namely, indicating the mappings by their labels in Matrix (2.2),

$$\begin{aligned} \mathcal{O}_0 &= \{1\}, \\ \mathcal{O}_1 &= \{9, 5, 3, 2\}, \\ \mathcal{O}_2 &= \{13, 11, 7, 10, 6, 4\}, \\ \mathcal{O}_3 &= \{15, 14, 12, 8\}, \\ \mathcal{O}_4 &= \{16\}. \end{aligned}$$

Note, in each case, that $|\mathcal{O}_i| = |G|/|G_{s_i}|$, where s_i is a representative on \mathcal{O}_i . \square

2.3.2. Contravariant actions. When in Definition 2.3.1, (2) is replaced by

$$\varphi(\sigma, \varphi(\tau, s)) = \varphi(\tau\sigma, s)$$

for all $s \in V$, τ, σ in G , we say that the φ is a *contravariant* action. To illustrate, let $\mathcal{F} = \mathcal{F}(V)$ indicate the vector space of scalar-valued functions, x , defined in the structure V where G acts according to φ . Let also

$$\theta(\tau, x)(s) = x(\varphi(\tau, s)), \quad s \in V, \quad \tau \in G, \quad x \in \mathcal{F}.$$

Then θ is a contravariant action of G on \mathcal{F} . In fact, $\theta(\tau, x) \in \mathcal{F}$,

$$\theta(1, x)(s) = x(\varphi(1, s)) = x(s),$$

for all $s \in V$, that is, $\theta(1, x)$ is the identity function in \mathcal{F} , and

$$\theta(\tau, \theta(\sigma, x))(s) = \theta(\sigma, x)(\varphi(\tau, s)) = x(\varphi(\sigma, \varphi(\tau, s))) = x(\varphi(\sigma\tau, s)) = \theta(\sigma\tau, x)(s),$$

for all $s \in V$. That is, $\theta(\tau, \theta(\sigma, x)) = \theta(\sigma\tau, x)$ for all $\tau, \sigma \in V$ and $x \in \mathcal{F}$.

Example 2.3.7 (Data indexed by conjugacy classes). Applying the above construction, let $V = G$ a finite group and $\varphi(\tau, \sigma) = \tau\sigma\tau^{-1}$, so that $\theta(\tau, x)(\sigma) = x(\tau\sigma\tau^{-1})$. Scalar-valued functions defined on the conjugacy classes of G play an important role in the theory of groups. \square

In addition, the reader may want to verify that

$$\tau^* : x \in \mathcal{F} \mapsto \theta(\tau, x) \in \mathcal{F}$$

is a linear mapping in \mathcal{F} , with inverse $\tau^{-1*} \in \mathcal{F}$ for all $\tau \in G$, and hence $\tau^* \in \text{GL}(\mathcal{F})$. In addition, $\tau \mapsto \tau^*$ is a group homomorphism.

2.3.3. Translations. For all $\sigma, \tau \in G$ we write $S_\tau(\sigma) = \sigma\tau^{-1}$ to indicate the *left translation* action, and $D_\tau(\sigma) = \tau\sigma$ to indicate the *right translation* action. Clearly, $S_\eta(S_\tau(\sigma)) = S_{\eta\tau}(\sigma)$, $D_\eta(S_\tau(\sigma)) = D_{\eta\tau}(\sigma)$, for all $\eta, \tau, \sigma \in G$, and $S_1(\sigma) = D_1(\sigma) = \sigma$ for all $\sigma \in G$.

2.3.4. Cayley's Theorem. Given $\tau \in G$ and an action φ of G on V , note that the evaluations $\tau^*(s) = \varphi(\tau, s)$, $s \in V$, define a permutation in V . In fact, $\varphi(\tau, s) = \varphi(\tau, f)$ implies $s = \varphi(\tau^{-1}, \varphi(\tau, s)) = \varphi(\tau^{-1}, \varphi(\tau, f)) = f$. Moreover, $\tau^*\tau^*(s) = \varphi(\tau, \varphi(\sigma, s)) = \varphi(\tau\sigma, s) = (\tau\sigma)^*(s)$, so that the mapping

$$\tau \in G \mapsto \tau^* \in S_V$$

is a group homomorphism from G to S_V . It is called the permutation representation of G associated with the action φ of G on V . We thus have

Proposition 2.3.1. If G acts on V according to φ , then the evaluations $\tau^*(s) = \varphi(\tau, s)$, $s \in V$, are permutations in V , for all $\tau \in G$ and $\tau \mapsto \tau^*$ is a group homomorphism. Conversely, given a homomorphism $\tau \in G \mapsto \tau^* \in S_V$, the mapping $\varphi(\tau, s) = \tau^*(s)$ defines a group action of G on V .

The argument justifying Proposition 2.3.1, when applied to any group acting on itself by (say) right translation, implies that $\tau \in \ker(\tau \mapsto \tau^*)$ if and only if $\tau^*(\sigma) = D_\tau(\sigma) = \tau\sigma = 1$ for all $\sigma \in G$, so that $\tau = 1$. Consequently, $\tau \mapsto \tau^*$ is a monomorphism from G to S_G . This leads to Cayley's Theorem:

Theorem 2.1 (Cayley, 1878). Every group G is isomorphic to a subgroup of S_G . If G is finite with ℓ elements, then G is isomorphic to a subgroup of S_ℓ .

When $\tau \mapsto \tau^*$ is a monomorphism we say that the corresponding action is *faithful*, or that G acts on V faithfully.

2.4. Linear representations

Proposition 2.3.1 shows that the mapping $\tau \mapsto \tau^*$ defined in G with values in S_V is a homomorphism from G into S_V . Correspondingly, let $\{e_s; s \in V\}$ indicate a basis for a vector space \mathcal{V} , indexed by the elements of V , $\{e_{\tau^*(s)}; s \in V\}$ the new basis determined by τ^* , and

$$\{e_s; s \in V\} \xrightarrow{\rho(\tau)} \{e_{\tau^*(s)}; s \in V\},$$

the nonsingular matrix representing the changing of basis. Then, ρ is a group homomorphism from G into the general linear group $\text{GL}(\mathcal{V})$.

Definition 2.4.1. A linear representation of a group G in a vector space \mathcal{V} is a group homomorphism from G into $\text{GL}(\mathcal{V})$.

Note that every linear representation maps the identity of G into the identity matrix (or operator) of $\text{GL}(\mathcal{V})$, that is, $\rho(1) = I$. Also, it maps the inverse τ^{-1} of τ into the inverse $\rho(\tau)^{-1}$ of the matrix $\rho(\tau)$, that is $\rho(\tau^{-1}) = \rho(\tau)^{-1}$. The dimension of ρ is defined as the dimension of the corresponding vector space.

Also note that if ρ is a linear representation of G in $\text{GL}(\mathcal{V})$ and B is any non-singular matrix of dimension equal to the dimension of ρ , then $\beta : \tau \mapsto B^{-1}\rho(\tau)B$ is also a linear representation of G in $\text{GL}(\mathcal{V})$. Any two such linear representations, obtained one from another by a changing of

basis, are called *equivalent* or *isomorphic* representations. We write $\rho \simeq \beta$ to indicate that ρ and β are equivalent³. Often, for simplicity of notation we may write ρ_τ and $\rho(\tau)$ without distinction.

Example 2.4.1 (One-dimensional representations). The *unity* or symmetric representation assigns the value $\rho_\tau = 1$ for all $\tau \in G$. The *antisymmetric* or *signature* representation of S_ℓ is defined as

$$\text{Sgn}(\tau) = \begin{cases} +1 & \text{if the permutation } \tau \text{ is even;} \\ -1 & \text{if the permutation } \tau \text{ is odd.} \end{cases}$$

From Proposition 2.2.1, we know that $\text{Sgn}(\sigma\tau) = \text{Sgn}(\sigma)\text{Sgn}(\tau)$, for any two permutations σ, τ . When G is the cyclic subgroup C_n of S_n , the reader may verify that

$$\rho_k(\tau^j) = \omega^{jk}, \quad j = 0, \dots, n-1, \quad k = 1, \dots, n,$$

where τ is a generator of C_n and ω is a primitive n -th root of 1, are n distinct one-dimensional representations of C_n . \square

Example 2.4.2 (Regular representations). The (left) regular representation is defined by the (left) translation action $S_\tau(\sigma) = \sigma\tau^{-1}$ of G on itself. A matrix representation of $\tau \in G$ is the matrix ϕ_τ changing the basis $\{e_\sigma; \sigma \in G\}$ into $\{e_{\sigma\tau^{-1}}; \sigma \in G\}$. The dimension of the representation is the number $|G|$ of elements in G . Similarly, the right regular representation is defined by the right translation action D_τ . For simplicity of notation we will refer to either action as the regular action, the context indicating which one is at work, and indicate its linear representation by ϕ .

To illustrate, consider the right translation by $\tau = (132) \in S_3$. From the multiplication table of S_3 ,

| | | | | | | |
|-----------|---|---|---|---|---|---|
| S_3 | a | b | c | d | e | f |
| a = 1 | a | b | c | d | e | f |
| b = (12) | b | a | e | f | c | d |
| c = (13) | c | f | a | e | d | b |
| d = (23) | d | e | f | a | b | c |
| e = (123) | e | d | b | c | f | a |
| f = (132) | f | c | d | b | a | e |

the canonical basis for \mathbb{R}^6 indexed by $\{1, (12), (13), (23), (123), (132)\}$ is changed under $D_\tau(\sigma) = \tau\sigma$ into the basis indexed by $\{(132), (23), (12), (13), 1, (123)\}$. Its representation is then

$$\phi_{(132)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

\square

The regular representation is particular case of a *permutation representation* of G associated with a set V . In the previous example, the action ϕ is the right translation, $V = G = S_3$ and $\tau^* = D_\tau$.

³Most of the theory of linear representation of finite groups have the equivalent result formulated for infinite groups, in which we would look at ρ as linear operators. In the present discussion, we often write or think of ρ as the notation indicating the representation in its matrix form. At times, however, the broader interpretation of ρ as a linear operator also applies.

A permutation representation of S_3 associated with $V = \{1, 2, 3\}$ results from changing the canonical basis indexed by V to the basis $\{\tau 1, \tau 2, \tau 3\}$. For example,

$$\rho_{(132)} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

We will refer to the representation of a subgroup G of S_ℓ acting on the set of indices $V = \{1, \dots, \ell\}$ for the canonical basis for \mathbb{R}^ℓ according to $(\tau, j) = \tau j$ simply as the *permutation representation* of G . The representation is an isomorphism between S_ℓ and M_ℓ .

A representation of S_2 acting on the space of binary sequences in length of two according to

$$\varphi_1 : \left[\begin{array}{c|cccc} \tau \backslash s & uu & yy & uy & yu \\ \hline 1 & uu & yy & uy & yu \\ \hline t = (12) & uu & yy & yu & uy \end{array} \right]$$

as discussed in Example 2.3.1, is given by

$$\rho_1(1) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \rho_1(t) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

The representation has dimension 4. Similarly, a representation of S_2 acting according to

$$\varphi_2 : \left[\begin{array}{c|cccc} \tau \backslash s & uu & yy & uy & yu \\ \hline 1 & uu & yy & uy & yu \\ \hline (12) & yy & uu & yu & uy \end{array} \right],$$

leads to the representation

$$\rho_2(1) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \rho_2(t) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Its dimension is also 4.

Example 2.4.3 (A two-dimensional representation of S_3). The reader may verify that

$$\begin{aligned} \beta_1 &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, & \beta_{(12)} &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, & \beta_{(13)} &= \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix}, \\ \beta_{(23)} &= \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix}, & \beta_{(123)} &= \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}, & \beta_{(132)} &= \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix} \end{aligned}$$

is a two-dimensional representation of S_3 . □

2.4.1. Unitary representations. An inner product in a vector space \mathcal{V} (over $\mathbb{F} = \mathbb{R}$ or \mathbb{C}) is a function $(\cdot, \cdot) : \mathcal{V}^2 \mapsto \mathbb{F}$ such that, for all $x, y, z \in \mathcal{V}$ and $a, b \in \mathbb{F}$,

- (1) $(x, y) = \overline{(y, x)}$, (Hermitian symmetric)
- (2) $(ax + by, z) = a(x, z) + b(y, z)$, (conjugate bilinear)
- (3) $(x, x) \geq 0$, $(x, x) = 0 \iff x = 0$, (positive definite).

The vector space \mathcal{V} , together with (\cdot, \cdot) is called an inner product space. An Euclidian (respectively unitary) space is a real (respectively complex) inner product space.

A linear representation ρ of G in the inner product space \mathcal{V} is unitary if

$$(\rho_\tau x, \rho_\tau y) = (x, y)$$

for all $x, y \in \mathcal{V}$ and $\tau \in G$. If (\cdot, \cdot) is an inner product in \mathcal{V} , direct evaluation shows that then

$$(2.19) \quad (x, y) = \frac{1}{|G|} \sum_{\tau \in G} (\rho_\tau x | \rho_\tau y)$$

is an inner product in \mathcal{V} , relative to which ρ is unitary. Moreover, ρ is equivalent to a representation that is unitary in the initial inner product space. To see this, indicate by $\{e_1, \dots, e_v\}$ an orthonormal basis relative to (\cdot, \cdot) and by $\{f_1, \dots, f_v\}$ an orthonormal basis relative to the invariant inner product (\cdot, \cdot) , and let A be the linear transformation defined by $Ae_i = f_i$. Then $(Ae_i, Ae_j) = \delta_{ij} = (e_i | e_j)$, so that $(Ax, Ay) = (x | y)$. Define $r(\tau) = A^{-1}\rho(\tau)A$, $\tau \in G$. Then r and ρ are equivalent and, because

$$(r(\tau)x | r(\tau)y) = (A^{-1}\rho(\tau)Ax | A^{-1}\rho(\tau)Ay) = (\rho(\tau)Ax, \rho(\tau)Ay) = (Ax, Ay) = (x | y),$$

r is unitary in the original inner product space.

Example 2.4.4. We will construct a representation unitarily equivalent to the two-dimensional representation of S_3 ,

$$\begin{aligned} \beta_1 &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, & \beta_{(12)} &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, & \beta_{(13)} &= \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix}, \\ \beta_{(23)} &= \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix}, & \beta_{(123)} &= \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}, & \beta_{(132)} &= \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix} \end{aligned}$$

illustrated earlier on in Example 2.4.3. The invariant scalar product derived from the Euclidean inner product $(\cdot | \cdot)$ in \mathbb{R}^2 is

$$(x, y) = \sum_{\tau} (\beta_\tau x | \beta_\tau y) = \sum_{\tau} x' \beta'_\tau \beta_\tau y = x' \begin{bmatrix} 8 & 4 \\ 4 & 8 \end{bmatrix} y \equiv x' F y.$$

Next, starting with the canonical basis $e_1 = (1, 0)$, $e_2 = (0, 1)$ for \mathbb{R}^2 , use Gram-Schmidt to construct a basis $\{w_1, w_2\}$ that is orthonormal relative to the invariant inner product:

- (1) $\|e_1\|^2 = e'_1 F e_1 = 8$. Let $w_1 = e_1 / \|e_1\| = (\sqrt{2}/4, 0)$;
- (2) w_2 is the normalized version of $e_2 - w'_1 F w_1 = \begin{bmatrix} -1/2 \\ 1 \end{bmatrix}$, which has norm $\sqrt{6}$. That is
 $w_2 = (-\sqrt{6}/12, \sqrt{6}/6)$.

The resulting new (unitarily equivalent) representation is then $b_\tau = A^{-1}\beta_\tau A$, where

$$A = \begin{bmatrix} 1/4\sqrt{2} & -1/12\sqrt{6} \\ 0 & 1/6\sqrt{6} \end{bmatrix}.$$

We obtain $b_1 = I_2$,

$$b_{12} = \begin{bmatrix} 1/2 & 1/4\sqrt{2}\sqrt{6} \\ 1/4\sqrt{2}\sqrt{6} & -1/2 \end{bmatrix}, \quad b_{13} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$b_{23} = \begin{bmatrix} 1/2 & -1/4\sqrt{2}\sqrt{6} \\ -1/4\sqrt{2}\sqrt{6} & -1/2 \end{bmatrix}, \quad b_{123} = \begin{bmatrix} -1/2 & 1/4\sqrt{2}\sqrt{6} \\ -1/4\sqrt{2}\sqrt{6} & -1/2 \end{bmatrix},$$

and

$$b_{132} = \begin{bmatrix} -1/2 & -1/4\sqrt{2}\sqrt{6} \\ 1/4\sqrt{2}\sqrt{6} & -1/2 \end{bmatrix}.$$

In each case we have $b_\tau b'_\tau = I_2$. □

2.4.2. Regular representations and group algebras. Given a finite group G , consider a vector space (\mathcal{A}) and a basis that is indexed by the elements of G . For example, we may index the canonical basis for \mathbb{R}^2 with the elements of $S_2 = \{1, t\}$ according to $e_1 = (1, 0)$, $e_t = (0, 1)$. The points (x) in this space have the form of symbolic linear combinations

$$x = \sum_{\sigma \in G} x(\sigma)\sigma,$$

with coefficients $x(\tau)$ in \mathbb{R} or \mathbb{C} . This vector space, of dimension equal to the number $|G|$ of elements in G , has an operation of multiplication defined by

$$(2.20) \quad xy = \sum_{\sigma, \eta} x(\sigma)y(\eta)\sigma\eta = \sum_{\tau} \left[\sum_{\sigma\eta=\tau} x(\sigma)y(\eta) \right] \tau \in \mathcal{A},$$

so that for all $x, y, z \in \mathcal{A}$ and all scalars γ in the field of the vector space, we have $xy \in \mathcal{A}$, $x(y + z) = xy + xz$, $(x + y)z = xz + yz$, $\gamma(xy) = x(\gamma y) = (\gamma x)y$. Moreover, because the group operation is associative, we have $x(yz) = (xy)z = xyz$. In this case, we say that the vector space \mathcal{A} , along with the multiplication so defined, constitutes an associative *group algebra*. It then follows that $(\tau, x) \rightarrow \tau x$ defines an action of G on \mathcal{A} in which the invariant subspaces, B , satisfy $\tau B \subset B$ for all $\tau \in G$. Consequently, by the linearity of the multiplication, we note that the invariant subspaces are exactly those subalgebras \mathcal{I} of \mathcal{A} that satisfy $y\mathcal{I} \subset \mathcal{I}$ for all $y \in \mathcal{A}$. These subalgebras are the *left ideals* of \mathcal{A} . In view of this, the determination of the invariant subspaces of the representation corresponds to searching for the left ideals of \mathcal{A} .

2.4.3. Tensor representations. Let ρ indicate the representation of $S_2 = \{1, t\}$ acting on $V = \{1, 2\}$ according to $(\tau, j) = \tau j$, $\tau \in S_2$, $j \in V$. It is given by $\rho(1) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\rho(t) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. It is simple to verify that

$$1 \mapsto I_4 = \rho(1) \otimes \rho(1), \quad t \mapsto \begin{bmatrix} 0 & \rho(t) \\ \rho(t) & 0 \end{bmatrix} = \rho(t) \otimes \rho(t),$$

where \otimes indicates the Kronecker product of two matrices, is a representation of G acting on $V \times V$ according to $\varphi(\tau, (j, k)) = (\tau j, \tau k)$. This new representation of G is called the *tensor* representation of ρ with itself, and is indicated by $\rho \otimes \rho$. Its dimension is $(\dim \rho)^2 = 4$. Similarly, $\varphi(\tau, (s, f)) = (\varphi_1(\tau, s), \varphi_2(\tau, f))$, $\tau \in G$, $(s, f) \in V \times V$ defines a tensor representation of two representations ρ_1 and ρ_2 determined by actions φ_1 and φ_2 of G on V . The same construction applies to tensor representations of three or more representations.

2.4.4. Actions on cosets. Given a subgroup H of G , consider the set G/H of cosets $\{\sigma H; \sigma \in G\}$ of H in G . Then, it is easy to verify that $\varphi: G \times G/H \rightarrow G/H$ defined by $\varphi(\tau, \sigma H) = \tau \sigma H$ is an action of G on G/H . Consequently, from Proposition 2.3.1, we obtain a group homomorphism ρ from G into S_n , where $n = [G : H]$ is the number of cosets σH of H in G . Moreover, if $\tau \in \ker \rho$, we must have $\tau \sigma H = \sigma H$ for all $\sigma \in G$. In particular, $\tau H = H$ so that $\tau \in H$. This proves

Theorem 2.2. If H is a subgroup of G of index $[G : H] = n$, then there is an homomorphism ρ of G into S_n , with $\ker \rho \subseteq H$.

The homomorphism ρ in Theorem 2.2 is a permutation representation of G acting on $V = G/H$. When $H = \{1\}$, Theorem 2.2 leads to Cayley's Theorem of Section 2.3.4.

Example 2.4.5. Let $G = S_3$ and $H = C_3 = \{a, e, f\}$, following the notation in the multiplication table (2.21) for S_3 .

(2.21)

| | | | | | | |
|-----------|---|---|---|---|---|---|
| * | a | b | c | d | e | f |
| a = 1 | a | b | c | d | e | f |
| b = (12) | b | a | e | f | c | d |
| c = (13) | c | f | a | e | d | b |
| d = (23) | d | e | f | a | b | c |
| e = (123) | e | d | b | c | f | a |
| f = (132) | f | c | d | b | a | e |

The cosets of H in G are $aH = eH = fH = H$ and $bH = cH = dH = \{b, c, d\}$. The number of cosets is the index $[G : H] = 2$. The action $(\tau, \sigma H) = \tau\sigma H$ of G on G/H , summarized in the following matrix,

$$\left[\begin{array}{c|cc} S_3 & aH & bH \\ \hline a & aH & bH \\ b & bH & aH \\ c & bH & aH \\ d & bH & aH \\ e & aH & bH \\ f & aH & bH \end{array} \right]$$

leads to a two-dimensional representation ρ of S_3 determined by

$$\rho_a = \rho_e = \rho_f = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \rho_b = \rho_c = \rho_d = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

in which $\ker \rho = H \subseteq H$. □

2.4.5. Matrices with group structure. In multivariate statistics, the following algebraic argument leads to a systematic characterization of certain patterned covariance matrices. See, for example, Andersson (1992) and Gao and Marden (2001).

Proposition 2.4.1. Given a representation ρ of G in \mathbb{R}^n and M a real $n \times n$ matrix, then

$$(2.22) \quad W = \frac{1}{|G|} \sum_{\tau \in G} \rho(\tau) M \rho(\tau)^{-1}$$

has the symmetry of (or is centralized by) ρ in the sense that $\rho(\sigma)W = W\rho(\sigma)$ for all $\sigma \in G$.

PROOF. For any $\sigma \in G$, we have

$$\rho(\sigma)W\rho(\sigma)^{-1} = \frac{1}{|G|} \sum_{\tau \in G} \rho(\sigma)\rho(\tau)M\rho(\tau)^{-1}\rho(\sigma)^{-1} = \frac{1}{|G|} \sum_{\tau \in G} \rho(\sigma\tau)M\rho(\sigma\tau)^{-1} = W,$$

observing that $\sigma\tau$ spans G when $\tau \in G$, for any σ in G . □

In this case, we also say that W commutes with the representation ρ . The set $\text{Hom}(\rho)$ of all linear operators commuting with a representation ρ in \mathcal{V} is a linear subspace of the space of linear operators in \mathcal{V} . Also in $\text{Hom}(\rho)$ are the linear operators of the form

$$\hat{x}(\rho) = \sum_{\tau \in G} x(\tau)\rho(\tau)$$

where x is a scalar function defined in G , with the additional property that x is constant over the conjugacy classes of G . See also Naimark and Štern (1982, p.55) and Simon (1996, p.28).

Example 2.4.6 (Matrices with dihedral structure). Following the notation suggested by Example 2.2.9, we write

$$\begin{aligned} D_4 &= \{1, \eta, \eta^2, \eta^3, \tau, \eta\tau, \eta^2\tau, \eta^3\tau\} \\ &= \{1, (abcd), (ac)(bd), (adcb), (ad)(bc), (bd), (ab)(cd), (ac)\}, \end{aligned}$$

to indicate the group of symmetries of the square. Its permutation representation is given by

$$\rho(\eta^j\tau^k) = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}^j \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}^k, \quad j, k = 0, 1, 2, 3.$$

Direct evaluation shows that the matrices centralized by D_4 have the form

$$\frac{1}{8} \sum_{\sigma \in D_4} \rho(\sigma)W\rho(\sigma)^{-1} = \begin{bmatrix} A & B & C & B \\ B & A & B & C \\ C & B & A & B \\ B & C & B & A \end{bmatrix},$$

with

$$A = \frac{1}{4}\text{tr } W, \quad B = \frac{1}{8}\text{tr } W[\rho(\eta) + \rho(\eta^3)], \quad C = \frac{1}{4}\text{tr } W\rho(\eta^2).$$

□

Example 2.4.7 (Matrices with complex structure). Consider the complex group $G = \{1, i, -1, -i\}$. Its multiplication table is

| | | | | |
|----|----|----|----|----|
| * | 1 | -1 | i | -i |
| 1 | 1 | -1 | i | -i |
| -1 | -1 | 1 | -i | i |
| i | i | -i | -1 | 1 |
| -i | -i | i | 1 | -1 |

Define

$$\rho(i^k) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^k, \quad k = 0, 1, 2, 3.$$

The reader may want to verify that ρ is a representation of the complex group in \mathbb{R}^2 and that

$$\frac{1}{4} \sum_{k=0}^3 \rho(i^k) \begin{bmatrix} a & b \\ c & d \end{bmatrix} \rho(i^k)^{-1} = \frac{1}{2} \begin{bmatrix} d+a & -c+b \\ -b+c & d+a \end{bmatrix}.$$

Matrices of the form

$$M = \begin{bmatrix} A & B \\ -B & A \end{bmatrix}$$

where A and B are $n \times n$ real matrices are said to have *complex structure* and carry the symmetry of the complex group represented by ρ in the sense that

$$(\rho(\tau) \otimes I_n)M = M(\rho(\tau) \otimes I_n), \quad \text{for all } \tau \in G.$$

□

Example 2.4.8 (Matrices with quaternionic structure). Following Example 2.2.4, define

$$\rho(\pm 1) = \pm \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \rho(\pm \mathbf{k}) = \mp \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

$$\rho(\pm\mathbf{j}) = \mp \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \quad \rho(\pm\mathbf{i}) = \mp \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Direct verification shows that ρ is a linear representation in \mathbb{R}^4 of the group of the quaternions. Given a real matrix

$$F = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ p & q & r & s \\ t & u & v & x \end{bmatrix},$$

it then follows that

$$\frac{1}{8} \sum \rho(\tau) F \rho(\tau)^{-1} = \frac{1}{4} \begin{bmatrix} a+f+r+x & b-e-s+v & c+h-p-u & d-g+q-t \\ e-b-v+s & a+f+r+x & g-d+t-q & c+h-p-u \\ p+u-c-h & d-g+q-t & a+f+r+x & e-b-v+s \\ g-d+t-q & p+u-c-h & b-e-s+v & a+f+r+x \end{bmatrix},$$

so that matrices of the form

$$M = \begin{bmatrix} A & B_1 & B_2 & B_3 \\ -B_1 & A & -B_3 & B_2 \\ -B_2 & B_3 & A & -B_1 \\ -B_3 & -B_2 & B_1 & A \end{bmatrix},$$

where A, B_1, B_2 and B_3 are any $n \times n$ real matrices, are said to have a *quaternionic structure*. Those are exactly the matrices with the symmetry of the given representation, in the sense that

$$(\rho(\tau) \otimes I_n)M = M(\rho(\tau) \otimes I_n), \quad \text{for all } \tau \in \mathbb{Q}.$$

□

2.5. Reducibility

Definition 2.5.1. Let ρ be a representation of G in $GL(\mathcal{V})$. A stable subspace of \mathcal{V} is a linear subspace \mathcal{W} of \mathcal{V} with the property that if $x \in \mathcal{W}$ then $\rho(\tau)x \in \mathcal{W}$ for all $\tau \in G$.

Note that $\{0\}$ and \mathcal{V} are stable subspaces of \mathcal{V} .

Example 2.5.1 (The Sym^2 and Alt^2 subspaces). Let $G = S_2 = \{1, t\}$ and ρ the permutation representation of S_2 . Starting with a canonical basis for \mathbb{R}^4 indexed as $\{e_{(1,1)}, e_{(1,2)}, e_{(2,1)}, e_{(2,2)}\}$ for the tensor representation $\rho \otimes \rho$ of G , form the new basis for \mathbb{R}^4 with components

$$v_1 = 2e_{(1,1)}, \quad v_2 = 2e_{(2,2)}, \quad v_3 = e_{(1,2)} + e_{(2,1)}, \quad v_4 = e_{(1,2)} - e_{(2,1)}.$$

The representation, ξ , of S_2 acting on the indices (j, k) is given by

$$\xi(1) = \begin{bmatrix} I_3 & 0 \\ 0 & 1 \end{bmatrix}, \quad \xi(t) = \begin{bmatrix} F & 0 \\ 0 & -1 \end{bmatrix},$$

where $F = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. We observe that the representation ξ *decomposes* as the sum of two representations ξ_1 and ξ_2 , given by

$$\xi_1(1) = I_3, \quad \xi_1(t) = F, \quad \xi_2(1) = 1, \quad \xi_2(t) = -1,$$

and that the corresponding subspaces

$$\mathcal{V}_1 = \langle v_1, v_2, v_3 \rangle, \quad \mathcal{V}_2 = \langle v_4 \rangle$$

are stable subspaces of \mathbb{R}^4 under ξ . The direct-sum decomposition $\mathbb{R}^4 = \mathcal{V}_1 \oplus \mathcal{V}_2$ justifies the notation $\xi = \xi_1 \oplus \xi_2$ to indicate that ξ is *reducible* and that the representations ξ_1 and ξ_2 are its *components*. To indicate that ξ and $\rho \otimes \rho$ are in fact equivalent (or isomorphic) representations we write

$$\rho \otimes \rho \simeq \xi_1 \oplus \xi_2.$$

The subspaces (and corresponding representations) \mathcal{V}_1 and \mathcal{V}_2 are called, respectively, the *symmetric square* (Sym^2) and *alternating square* (Alt^2) subspaces or representations⁴. The study of group representations is concerned with describing all inequivalent, indecomposable representations of a group G . \square

More generally, let ρ indicate a representation of G acting on V with v elements, so that the basis for the tensor representation $\rho \otimes \rho$ of G is indexed by $V \times V$. Let D indicate the main diagonal of $V \times V$ and U its upper triangular part. The representation of G acting on the indices of the basis for the subspace \mathcal{V}_1 indexed as

$$\{e_{(s,f)} + e_{(f,s)}; (s, f) \in D \cup U\}$$

is the Sym^2 (symmetric square) representation, of dimension $v(v+1)/2$, whereas the representation of G acting on the indices of the basis for \mathcal{V}_2 indexed as

$$\{e_{(s,f)} - e_{(f,s)}; (s, f) \in U\}$$

is the Alt^2 (alternating square) representation, of dimension $v(v-1)/2$. Moreover, $\rho \otimes \rho \simeq \text{Sym}^2 \oplus \text{Alt}^2$.

To illustrate, let $G = C_4 = \{1, \tau, \tau^2, \tau^3\}$, and $\rho(\tau^k) = r^k$, $k = 0, 1, 2, 3$, where

$$r = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Correspondingly, $(\rho \otimes \rho)(\tau^k) = r^k \otimes r^k$. The basis for $\rho \otimes \rho$ may be indexed by $V \times V = \{(1, 1), (1, 2), \dots, (4, 3), (4, 4)\}$, from which the two bases for the Sym-square and Alt-square representations can be obtained. The indices for these bases are, respectively,

$$D \cup U = \{(1, 1), (1, 2), (1, 3), (1, 4), (2, 2), (2, 3), (2, 4), (3, 3), (3, 4), (4, 4)\},$$

$$U = \{(1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (3, 4)\}.$$

The matrix generating the two new bases $\{e_{(s,f)} + e_{(f,s)}; (s, f) \in D \cup U\}$ and $\{e_{(s,f)} - e_{(f,s)}; (s, f) \in U\}$ is

$$B = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \end{bmatrix},$$

⁴The representation Alt^2 is also called *exterior square* representation.

from which we obtain,

$$B(\rho \otimes \rho)(\tau)B^{-1} \simeq (\rho \otimes \rho)(\tau) = \begin{bmatrix} \boxed{\text{Alt}^2(\tau)} & 0 \\ 0 & \boxed{\text{Sym}^2(\tau)} \end{bmatrix},$$

or, $\rho \otimes \rho \simeq \text{Alt}^2 \oplus \text{Sym}^2$. Correspondingly, $\mathcal{V} = \mathbb{R}^{16}$ decomposes into the direct sum $\mathcal{V}_1 \oplus \mathcal{V}_2$ of invariant subspaces of dimensions 10 and 6 respectively. The components of the decomposition are given by

$$\text{Sym}^2(\tau^k) = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^k,$$

$$\text{Alt}^2(\tau^k) = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix}^k, \quad k = 0, 1, 2, 3.$$

Note that when $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2$ and \mathcal{V}_1 is a stable subspace of V , of dimension v_1 , under ρ , it is necessary and sufficient that the pattern of $\rho(\tau)$ takes the matrix form

$$\rho(\tau) = \begin{bmatrix} R_1(\tau) & 0 \\ M(\tau) & R_2(\tau) \end{bmatrix},$$

for matrices $R_1(\tau)$, $R_2(\tau)$ and $M(\tau)$ of dimensions $v_1 \times v_1$, $v_2 \times v_2$ and $v_2 \times v_1$, respectively, so that $\tau \mapsto R_1(\tau)$ and $\tau \mapsto R_2(\tau)$ are representations of G , of dimensions v_1 and v_2 , respectively. In this case, we say that ρ is a *reducible* representation.

Definition 2.5.2. We say that a representation ρ of G in $\text{GL}(\mathcal{V})$ is irreducible when the only proper stable linear subspace of \mathcal{V} is the null subspace, that is,

$$\rho(\tau)v \in \mathcal{W} \text{ for all } \tau \in G, \text{ for all } v \in \mathcal{W}, \text{ for some } \mathcal{W} \subsetneq \mathcal{V} \implies \mathcal{W} = \{0\}.$$

Clearly, then, all one-dimensional representations are irreducible.

Example 2.5.2. Consider again the representation of S_2 in $\mathcal{V} = \mathbb{R}^2$ given by $\rho_1 = I_2$ and $\rho_t = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, where $t = (12)$. Changing the canonical basis for \mathcal{V} indexed as $\{e_1, e_t\}$ to the new basis $\{e_t - e_1, e_t\}$, we see that ρ is equivalent to the representation $\xi_\tau = B\rho_\tau B^{-1}$ given by

$$\xi_1 = \begin{bmatrix} \boxed{1} & \boxed{0} \\ \boxed{0} & \boxed{1} \end{bmatrix}, \quad \xi_t = \begin{bmatrix} \boxed{-1} & \boxed{0} \\ \boxed{-1} & \boxed{1} \end{bmatrix},$$

where $B = \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix}$. If $\mathcal{V}_1 = \langle e_t - e_1 \rangle$ and $\mathcal{V}_2 = \langle e_t \rangle$ then \mathcal{V} decomposes as the direct sum $\mathcal{V}_1 \oplus \mathcal{V}_2$ and \mathcal{V}_1 is a stable one-dimensional subspace reduced by the signature representation introduced earlier on in Example 2.4.1. Consequently, ρ is reducible. Note however that \mathcal{V}_2 is not yet a stable complement of \mathcal{V}_1 in \mathcal{V} .

Consider, instead, the basis $\{e_t - e_1, e_t + e_1\}$ for \mathcal{V} , relative to which the equivalent representation is now

$$\beta_1 = \begin{bmatrix} \boxed{1} & 0 \\ 0 & \boxed{1} \end{bmatrix}, \quad \beta_t = \begin{bmatrix} \boxed{-1} & 0 \\ 0 & \boxed{1} \end{bmatrix}.$$

The new component is the unit or symmetric representation, which is also one-dimensional and hence irreducible. In summary, ρ decomposes as the sum $1 \oplus \text{Sgn}$ of two irreducible one-dimensional representations, \mathcal{V} decomposes and the sum $\mathcal{V}_1 \oplus \mathcal{V}_{\text{Sgn}}$ of two stable (irreducible) subspaces:

$$\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_{\text{Sgn}}, \quad \rho \simeq \rho_1 \oplus \rho_{\text{Sgn}}.$$

□

Example 2.5.3. Consider the action

$$\varphi : \left[\begin{array}{c|cccc} \tau \backslash s & uu & yy & uy & yu \\ \hline 1 & uu & yy & uy & yu \\ \hline t = (12) & uu & yy & yu & uy \end{array} \right]$$

of S_2 on the space of binary sequences in length of two according to position symmetry. Its representation is given by

$$\omega(1) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \omega(t) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

We observed that ω is decomposable as $\rho_1 \oplus \rho_1 \oplus \rho$, where ρ_1 is the unit representation and ρ is the permutation representation of S_2 , which further reduces into the (irreducible) one-dimensional components 1 and Sgn . That is, $\omega \simeq 1 \oplus 1 \oplus 1 \oplus \text{Sgn}$. □

Example 2.5.4 (A two-dimensional irreducible representation of S_3). We will construct a two-dimensional irreducible representation of S_3 . Let ρ indicate the permutation representation of S_3 , which is given by

$$\rho_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \rho_{(12)} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \rho_{(13)} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

$$\rho_{(23)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \rho_{(123)} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad \rho_{(132)} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Start with the subspace W_1 generated by the sum $e = e_1 + e_2 + e_3$ of the vectors in the canonical basis $\{e_1, e_2, e_3\}$ of $\mathcal{V} = \mathbb{R}^3$. That is, W_1 is generated by $e' = (1, 1, 1)$. Clearly, W_1 is a stable subspace of ρ , that is,

$$\rho_\tau y \in W_1, \quad \text{for all } y \in W_1, \quad \text{for all } \tau \in S_3.$$

Let $W_0 = \{y \in \mathcal{V}; e'y = 0\}$ be the orthogonal complement of W_1 in \mathcal{V} and $\mathcal{A} = \frac{1}{3}ee'$ the projection on W_1 along W_0 , that is, $\mathcal{V} = W_0 \oplus W_1$, $\mathcal{A}^2 = \mathcal{A}$ and $\mathcal{A}y = 0$ for all $y \in W_0$. Similarly, let

$$(2.23) \quad \mathcal{Q} = I_3 - \mathcal{A} = \frac{1}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

indicate the projection on W_0 along W_1 . The reader can verify that \mathcal{Q} is centralized by ρ , that is, $\rho_\tau \mathcal{Q} = \mathcal{Q} \rho_\tau$ for all $\tau \in S_3$, and that, consequently, if $y \in W_0$ then $y \in \mathcal{Q}z$ for some $z \in \mathcal{V}$, and $\rho_\tau y = \rho_\tau \mathcal{Q}z = \mathcal{Q} \rho_\tau z \in W_0$, for all $\tau \in S_3$. That is, W_0 is a stable two-dimensional complement of W_1 in \mathcal{V} . To construct a 2-dimensional representation (β) in W_0 , note, from the corresponding

projection in (2.23), that a basis $\{v_1, v_2\}$ for $\text{im } \mathcal{Q}$ is $v_1 = 2e_1 - e_2 - e_3$, $v_2 = -e_1 + 2e_2 - e_3$. The resulting representation of $\tau = (12)$, for example, is obtained from the fact that

$$\tau v_1 = 2e_{\tau 1} - e_{\tau 2} - e_{\tau 3} = 2e_2 - e_1 - e_3 = v_2,$$

$$\tau v_2 = -e_{\tau 1} + 2e_{\tau 2} - e_{\tau 3} = -e_2 + 2e_1 - e_3 = v_1,$$

that is, $\beta_{(12)} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Similar calculations (noting, from 2.23, that $-e_1 - e_2 + 2e_3 = -v_1 - v_2$) leads to the linear representation

$$\beta_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \beta_{(12)} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \beta_{(13)} = \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix},$$

$$\beta_{(23)} = \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix}, \quad \beta_{(123)} = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}, \quad \beta_{(132)} = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}.$$

The two-dimensional representation β is irreducible. In fact, if there were a proper one-dimensional stable subspace W , with generator y , then it would verify $\beta_{(12)}y = \lambda y$ for some scalar λ , which implies $y_2 = \lambda y_1$, $y_1 = \lambda y_2$. The non-zero eigenvalue solutions to $y_2 = \lambda^2 y_1$ are $\lambda = \pm 1$, that is, $y = (y_1, y_1)$ or $y = (y_1, -y_1)$. Since the subspace W must also be stable under $\beta_{(13)}$ then we would have

$$\begin{aligned} \beta_{(13)}y &= \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} -2y_1 \\ y_1 \end{bmatrix} \in W \\ &\iff y_1 = 0, \text{ using } y = (y_1, y_1) \text{ or } y = (y_1, -y_1) \implies W = \{0\}. \end{aligned}$$

Because $\{0\}$ is the only proper stable subspace, β is irreducible. Table (2.24) summarizes the irreducible representations of S_3 . It includes the presently derived two-dimensional representation, along with the trivial and signature (one-dimensional) ones. Since the trace $\text{tr } \rho(\tau)$ of a representation, indicated here by $\chi_\rho(\tau)$, is constant over conjugacy classes, it is sufficient to report it for representatives of these classes.

$$(2.24) \quad \left[\begin{array}{c|ccc} \chi & 1 & (12) & (123) \\ \hline \chi_1 & 1 & 1 & 1 \\ \chi_\beta & 2 & 0 & -1 \\ \chi_{\text{Sgn}} & 1 & -1 & 1 \end{array} \right].$$

This table completely describes the representations of S_3 , and will be studied later on in the chapter with more detail. \square

To appreciate the role of the field of scalars in Example 2.5.4, restrict the search for a one-dimensional stable subspace to the cyclic subgroup $C_3 = \{1, (123), (132)\}$ of S_3 . In this case, we have the two-dimensional representation

$$\gamma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \gamma_{(123)} = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}, \quad \gamma_{(132)} = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}.$$

The equations $\gamma_\tau y = \lambda y$ for $\tau \in C_3$ lead to the characteristic equations $(1 - \lambda)^2 = 0$ and $\lambda^2 + \lambda + 1 = 0$. When the field of scalars is \mathbb{C} we find two one-dimensional irreducible representations, corresponding to the roots $\omega = e^{\frac{2\pi i}{3}}$ and ω^2 . If the field of scalars is \mathbb{R} then γ is irreducible. Here is the summary for three irreducible representations of C_3 , over \mathbb{C} :

$$(2.25) \quad \left[\begin{array}{c|ccc} \chi & 1 & (123) & (132) \\ \hline \chi_1 & 1 & 1 & 1 \\ \chi_2 & 1 & \omega & \omega^2 \\ \chi_3 & 1 & \omega^2 & \omega \end{array} \right].$$

Example 2.5.5 (Planar rotations). Let \mathcal{V} indicate the infinite-dimensional real vector space of all trigonometric Fourier series

$$f(t) = \sum_{m=0}^{\infty} a_m \cos(mt) + b_m \sin(mt), \quad -\pi \leq t \leq \pi, \quad a_m, b_m \in \mathbb{R},$$

with basis $\mathcal{B} = \{\cos(mt), \sin(mt) : m = 0, 1, 2, \dots\}$, and consider the subspaces \mathcal{W}_ℓ of \mathcal{V} generated by $\mathcal{B}_\ell = \{\cos(mt), \sin(mt) : m = 0, 1, 2, \dots, \ell - 1\}$. We let the cyclic group $C_\ell = \{\tau^h; h = 0, \dots, \ell - 1\}$ act on the elements $g \in \mathcal{W}_\ell$ according to $\varphi(\tau^h, g)(t) = g\tau^{-h}(t) = g(t - h)$. To see that this is indeed an action note that if the components $\cos(mt)$ and $\sin(mt)$ are in \mathcal{W}_ℓ then the components of

$$\begin{bmatrix} \cos(m(t-h)) \\ \sin(m(t-h)) \end{bmatrix} = \begin{bmatrix} \cos(mh) & -\sin(mh) \\ \sin(mh) & \cos(mh) \end{bmatrix} \begin{bmatrix} \cos(mt) \\ \sin(mt) \end{bmatrix},$$

are also in \mathcal{W}_ℓ , and that $\varphi(\tau^h, \varphi(\tau^k, g)) = \varphi(\tau^{h+k}, g)$. This linear relationship defines a family of representations

$$\rho_m(\tau^h) = \begin{bmatrix} \cos(mh) & -\sin(mh) \\ \sin(mh) & \cos(mh) \end{bmatrix}, \quad h, m = 0, 1, \dots, \ell - 1,$$

where the multiplication is mod ℓ . These representations in turn sum as a representation $\rho = \rho_0 \oplus \rho_1 \oplus \dots \oplus \rho_{\ell-1}$ of C_ℓ in $GL(\mathbb{R}^{2\ell})$ in which the components ρ_m are irreducible (over \mathbb{R}) for $m > 0$ and ρ_0 further reduces as $1 \oplus 1$. To illustrate, when $\ell = 3$,

$$\rho(\tau^h) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \cos(h) & -\sin(h) & 0 & 0 \\ 0 & 0 & \sin(h) & \cos(h) & 0 & 0 \\ 0 & 0 & 0 & 0 & \cos(2h) & -\sin(2h) \\ 0 & 0 & 0 & 0 & \sin(2h) & \cos(2h) \end{bmatrix} = 1 \oplus 1 \oplus \rho_1 \oplus \rho_2.$$

When expressed in trigonometric form, we obtain $\rho(1) = I$,

$$\rho(\tau) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \cos(\alpha) & \sin(\alpha) & 0 & 0 \\ 0 & 0 & -\sin(\alpha) & \cos(\alpha) & 0 & 0 \\ 0 & 0 & 0 & 0 & \cos(2\alpha) & \sin(2\alpha) \\ 0 & 0 & 0 & 0 & -\sin(2\alpha) & \cos(2\alpha) \end{bmatrix},$$

and

$$\rho(\tau^2) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \cos(2\alpha) & \sin(2\alpha) & 0 & 0 \\ 0 & 0 & -\sin(2\alpha) & \cos(2\alpha) & 0 & 0 \\ 0 & 0 & 0 & 0 & \cos(\alpha) & \sin(\alpha) \\ 0 & 0 & 0 & 0 & -\sin(\alpha) & \cos(\alpha) \end{bmatrix}, \quad \alpha = \frac{2\pi}{\ell}.$$

The reader may want to verify that ρ_1 and ρ_2 are indeed irreducible representations of C_3 . The construction of the linear representation ρ , described in terms of three-fold planar rotations in the present example ($\ell = 3$), extends naturally to $\ell = \infty$. The resulting infinite-dimensional representation describes all two-dimensional invariant subspaces for \mathcal{V} .

This example also outlines the general structure of the representation obtained when the planar rotations are replaced by spherical rotations. In that case, the invariant subspaces under

the action of the full three-dimensional rotation group are spanned by the spherical harmonics $Y_{\ell m}(\theta, \phi)$. To each basis indexed by ℓ there corresponds a $2\ell + 1$ -dimensional subspace. See also Riley et al. (2002, p.930). \square

Theorem 2.3. Let $\rho : G \rightarrow \text{GL}(\mathcal{V})$ be a linear representation of G in \mathcal{V} and let W_1 be a vector subspace of \mathcal{V} stable under G . Then there is a complement W_0 of W_1 in \mathcal{V} which is also stable under G .

PROOF. Let \mathcal{P}_1 be a projection on W_1 along some vector space complement of W_1 in \mathcal{V} . Form the average

$$\bar{\mathcal{P}}_1 = \frac{1}{|G|} \sum_{\tau \in G} \rho(\tau) \mathcal{P}_1 \rho(\tau^{-1})$$

of projections on W_1 along that vector space complement. Then, $\text{im } \bar{\mathcal{P}}_1 = \{\bar{\mathcal{P}}_1 z; z \in \mathcal{V}\} = W_1$. To see this, first note that for $z \in \mathcal{V}$ we have $\mathcal{P}_1 \rho_{\tau^{-1}} z \in W_1$, and because W_1 is a stable subspace, $\rho_\tau[\mathcal{P}_1 \rho_{\tau^{-1}} z] \in W_1$, so that $\bar{\mathcal{P}}_1 z \in W_1$, that is, $\text{im } \bar{\mathcal{P}}_1 \subseteq W_1$. Secondly, if $z \in W_1$, which is stable, we have $\rho_{\tau^{-1}} z \in W_1$ for all $\tau \in G$, so that $\mathcal{P}_1 \rho_{\tau^{-1}} z = \rho_{\tau^{-1}} z$. This implies

$$\bar{\mathcal{P}}_1 z = \frac{1}{|G|} \sum_{\tau \in G} \rho_\tau \mathcal{P}_1 \rho_{\tau^{-1}} z = \frac{1}{|G|} \sum_{\tau \in G} \rho_\tau \rho_{\tau^{-1}} z = z,$$

that is, if $z \in W_1$ then $z = \bar{\mathcal{P}}_1 z \in \text{im } \bar{\mathcal{P}}_1$, and hence $W_1 \subseteq \text{im } \bar{\mathcal{P}}_1$. Therefore, $W_1 = \text{im } \bar{\mathcal{P}}_1$. Let then $W_0 = \ker \bar{\mathcal{P}}_1 = \{z \in \mathcal{V}; \bar{\mathcal{P}}_1 z = 0\}$, so that $\mathcal{V} = W_1 \oplus W_0$. To conclude the proof, we must show that W_0 is G -stable: In fact, for all $\tau \in G$,

$$\begin{aligned} \rho_\tau \bar{\mathcal{P}}_1 \rho_{\tau^{-1}} &= \frac{1}{|G|} \sum_{\sigma \in G} \rho_\tau \rho_\sigma \mathcal{P}_1 \rho_{\sigma^{-1}} \rho_{\tau^{-1}} = \frac{1}{|G|} \sum_{\sigma \in G} \rho_{\tau\sigma} \mathcal{P}_1 \rho_{(\tau\sigma)^{-1}} \\ &= \frac{1}{|G|} \sum_{\sigma \in G} \rho_\sigma \mathcal{P}_1 \rho_{\sigma^{-1}} = \bar{\mathcal{P}}_1, \end{aligned}$$

so that $y \in W_0 = \ker \bar{\mathcal{P}}_1$ implies $\bar{\mathcal{P}}_1 y = 0$ and hence $\bar{\mathcal{P}}_1 \rho_\tau y = \rho_\tau \bar{\mathcal{P}}_1 y = 0$, thus showing that $\rho_\tau y \in W_0$, for all $\tau \in G$. Consequently, W_0 is a stable subspace of \mathcal{V} under G . \square

Example 2.5.6 (The invariant subspaces of a group algebra). With the definitions and notation of Section 2.4.2, let \mathcal{I}_1 indicate a left ideal of the group algebra \mathcal{A} . Theorem 2.3 implies that \mathcal{A} decomposes as the direct sum

$$\mathcal{A} = \mathcal{I}_1 \oplus \mathcal{I}_2$$

of \mathcal{I}_1 and a complementary ideal \mathcal{I}_2 . If $x \in \mathcal{A}$ then $x = x_1 + x_2$ with x_1, x_2 in $\mathcal{I}_1, \mathcal{I}_2$ respectively. In particular, the identity $1 \in \mathcal{A}$ can be expressed as

$$1 = e_1 + e_2,$$

so that $x = x e_1 + x e_2$, for all $x \in \mathcal{A}$, thus showing that the subspaces $\mathcal{I}_1, \mathcal{I}_2$ are spanned by e_1 and e_2 , respectively. When $x \in \mathcal{I}_1$, because \mathcal{I}_1 is a left ideal, $x e_1 \in \mathcal{I}_1$ and $x = x(e_1 + e_2) = x e_1$. In particular, for $x = e_1$, $e_1 = e_1^2$. Similarly, $e_2 = e_2^2$. In addition,

$$e_1 = e_1(e_1 + e_2) = e_1^2 + e_1 e_2 = e_1 + e_1 e_2$$

so that $e_1 e_2 = 0$. Similarly, $e_2 e_1 = 0$. Repeating the argument in each component, we obtain a final decomposition of the form

$$\mathcal{A} = \mathcal{I}_1 \oplus \mathcal{I}_2 \oplus \dots \oplus \mathcal{I}_h,$$

$$1 = e_1 + e_2 + \dots + e_h,$$

with $e_i^2 = e_i$ and $e_i e_j = 0$ for $i \neq j$, and such that each ideal cannot be further reduced as a sum of two left ideals. The irreducible left ideals are called the *primitive idempotents* of the group algebra.

Theorem 2.4. Every representation is a direct sum of irreducible representations.

PROOF. Let \mathcal{V} be (the vector space associated to) a linear representation of G . The argument is by induction on the dimension of \mathcal{V} . Suppose $\dim \mathcal{V} \geq 1$. If \mathcal{V} is irreducible, the proof is complete. Otherwise, from Theorem 2.3, $\mathcal{V} = V' \oplus V''$ with $\dim V' < \dim \mathcal{V}$ and $\dim V'' < \dim \mathcal{V}$. By the induction hypothesis, V' and V'' are direct sum of irreducible representations, and then so is \mathcal{V} . \square

2.6. Characters of a representation

Given a representation ρ , the complex-valued function

$$\chi_\rho : \tau \rightarrow \text{tr } \rho_\tau$$

is called the *character* of the representation. It plays an important role in the characterization of the representation. Since $\rho(1) = I_\ell$ and $\ell = \dim \rho$, we note that $\chi_\rho(1) = \dim \rho$.

If λ is an eigenvalue of ρ , then, relative to the invariant inner product, expression (2.19), we have

$$(y, y) = (\rho_\tau y, \rho_\tau y) = (\lambda y, \lambda y) = \lambda \bar{\lambda} (y, y),$$

so that $\lambda \bar{\lambda} = 1$. Let $\lambda_1, \dots, \lambda_m$ indicate the eigenvalues of ρ_τ (over \mathbb{C}). Then

$$(2.26) \quad \chi_\rho(\tau^{-1}) = \text{tr } \rho_{\tau^{-1}} = \text{tr } \rho_\tau^{-1} = \sum \lambda_i^{-1} = \sum \bar{\lambda}_i = \text{tr } \bar{\rho}_\tau = \bar{\chi}_\rho(\tau).$$

Also note, since trace is invariant under similarity, that

$$(2.27) \quad \chi_\rho(\tau \sigma \tau^{-1}) = \chi_\rho(\sigma), \quad \text{for all } \tau, \sigma \in G.$$

Proposition 2.6.1. Let $\rho_i : G \rightarrow \text{GL}(\mathcal{V}_i)$ be a linear representation of G , with corresponding character χ_i , $i = 1, 2$. Then

$$\chi_{\rho_1 \oplus \rho_2} = \chi_1 + \chi_2, \quad \chi_{\rho_1 \otimes \rho_2} = \chi_1 \times \chi_2.$$

PROOF. We have

$$\chi_{\rho_1 \oplus \rho_2} = \text{tr } (\rho_1 \oplus \rho_2) = \text{tr } \begin{bmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{bmatrix} = \text{tr } \rho_1 + \text{tr } \rho_2 = \chi_1 + \chi_2,$$

whereas, noting that the diagonal of $\chi_{\rho_1 \otimes \rho_2}$ is

$$([\rho_1]_{11} \text{diag } \rho_2, [\rho_1]_{22} \text{diag } \rho_2, \dots, [\rho_1]_{n_1 n_1} \text{diag } \rho_2),$$

we obtain

$$\chi_{\rho_1 \otimes \rho_2} = \text{tr } (\rho_1 \otimes \rho_2) = \sum_i [\rho_1]_{ii} \times \sum_j [\rho_2]_{jj} = \text{tr } \rho_1 \times \text{tr } \rho_2 = \chi_1 \times \chi_2. \quad \square$$

In Example 2.5.1 we considered the Sym^2 and Alt^2 representations and showed that $\rho \otimes \rho \simeq \text{Sym}^2 \oplus \text{Alt}^2$. From the decomposition for the tensor representation of C_4 discussed in that example, we obtain the following characters:

| C_4 | $\chi_\rho(\tau)$ | $\chi_{\text{Sym}^2}(\tau)$ | $\chi_{\text{Alt}^2}(\tau)$ |
|----------|-------------------|-----------------------------|-----------------------------|
| 1 | 4 | 10 | 6 |
| (1234) | 0 | 0 | 0 |
| (13)(24) | 0 | 2 | -2 |
| (1432) | 0 | 0 | 0 |

Note that, for all $\tau \in C_4$,

$$\chi_\rho^2(\tau) = \chi_{\rho \otimes \rho}(\tau) = \chi_{\text{Sym}^2}(\tau) + \chi_{\text{Alt}^2}(\tau),$$

and

$$\chi_{\text{Sym}^2}(\tau) = \frac{1}{2}(\chi_\rho^2(\tau) + \chi_\rho(\tau^2)), \quad \chi_{\text{Alt}^2}(\tau) = \frac{1}{2}(\chi_\rho^2(\tau) - \chi_\rho(\tau^2)).$$

It can be shown that these two equalities hold in general for any linear representation ρ of G .

2.7. Schur's Lemma and applications

Lemma 2.7.1 (Schur). Let $\rho_i : G \rightarrow \mathcal{V}_i$ be irreducible representations of G , $i = 1, 2$, and let $f : \mathcal{V}_1 \rightarrow \mathcal{V}_2$ be a non-zero linear mapping satisfying $f\rho_1(\tau) = \rho_2(\tau)f$ for all $\tau \in G$. Then ρ_1 and ρ_2 are isomorphic. If, in addition, $\mathcal{V}_1 = \mathcal{V}_2$ and $\rho_1 = \rho_2$ then f is a scalar multiple of the identity mapping.

PROOF. Let $W_1 = \ker f = \{x; f(x) = 0\}$. If $x \in W_1$ then $f(x) = 0$ and $f\rho_2(\tau)x = \rho_1(\tau)f(x) = 0$, which implies $\rho_1(\tau)x \in W_1$, for all $\tau \in G$. That is, W_1 is a stable subspace. Since ρ is irreducible, we must have $W_1 = \{0\}$ or $W_1 = \mathcal{V}_1$. If $W_1 = \mathcal{V}_1$ then $f = 0$, contrary to the hypothesis, hence $W_1 = \{0\}$. Similarly, we obtain $\text{im } f$ is stable and equal to \mathcal{V}_2 . Hence, f is an isomorphism, and the two representations are equivalent or isomorphic. For the second part, let λ be an eigenvalue of f (the field is \mathbb{C} , so there is at least one) and define $f' = f - \lambda$, understanding that $\lambda \equiv \lambda I$. If $f(x) = \lambda x$ then $(f - \lambda)x = 0$, so that $\ker(f - \lambda) \neq \{0\}$, and equivalently, $f - \lambda$ is not an isomorphism. Moreover,

$$(f - \lambda)\rho(\tau) = f\rho(\tau) - \lambda\rho(\tau) = \rho(\tau)f - \rho(\tau)\lambda = \rho(\tau)(f - \lambda), \quad \text{for all } \tau \in G.$$

From the first part of the Lemma, it follows that $f - \lambda = 0$, or $f = \lambda I$. \square

In the analysis of structured data, it is often of interest to consider the vector space $\mathcal{F}(G)$ of all scalar functions defined on G . An important element in $\mathcal{F}(G)$ is the character $\chi_\rho(\tau) = \text{tr } \rho_\tau$ of a representation ρ of G , introduced earlier on in Section 2.6. In general, note that each entry ρ_{ij} of a linear representation ρ defines a scalar function $\tau \rightarrow \rho_{ij}(\tau)$.

The representations

$$\begin{aligned} \beta_1 &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, & \beta_{(12)} &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, & \beta_{(13)} &= \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix}, \\ \beta_{(23)} &= \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix}, & \beta_{(123)} &= \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}, & \beta_{(132)} &= \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}, \end{aligned}$$

of dimension 2, the unit (1) and the signature (Sgn) representations of S_3 identified in Example 2.5.4 account for $24 + 1 + 1 = 26$ scalar functions defined on G , or 26 points in the vector space $\mathcal{F}(G)$. These functions have a number of characteristic properties. For example, note that

$$\sum_{\tau \in G} 1(\tau)h \text{ Sgn}(\tau^{-1}) = 0, \quad \text{for all scalar } h,$$

$$\sum_{\tau \in G} 1(\tau)H\beta(\tau^{-1}) = 0, \quad \sum_{\tau \in G} \text{Sgn}(\tau)H\beta(\tau^{-1}) = 0,$$

for all linear mappings $H : \mathbb{R}^2 \rightarrow \mathbb{R}$.

Proposition 2.7.1. For every non-equivalent irreducible representations ρ_1, ρ_2 and every linear mapping $H : \mathcal{V}_1 \rightarrow \mathcal{V}_2$, it holds that $\sum_{\tau \in G} \rho_1(\tau)H\rho_2(\tau^{-1}) = 0$.

PROOF. Note that $H_0 = \sum_{\tau \in G} \rho_1(\tau)H\rho_2(\tau^{-1})$ is a linear mapping from \mathcal{V}_1 into \mathcal{V}_2 which intertwines with $\rho_1(\tau)$ and $\rho_2(\tau)$ for all $\tau \in G$, that is, $\rho_1(\tau)H_0 = H_0\rho_2(\tau)$ for all $\tau \in G$. From Schur's Lemma (the representations are non-equivalent irreducible) it follows that $H_0 = 0$. \square

Now take any linear mapping $H = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and consider the two-dimensional irreducible representation β of S_3 reviewed above. Direct evaluation shows that

$$\frac{1}{6} \sum_{\tau \in S_3} \beta_\tau H \beta_{\tau^{-1}} = \frac{a+d}{2} I_2 = \frac{\text{tr } H}{2} I_2.$$

In fact, we have,

Proposition 2.7.2. Let ρ be an irreducible representation of G into $GL(\mathcal{V})$ with $\dim \rho = n$. Then, for any linear mapping H in \mathcal{V} ,

$$\frac{1}{|G|} \sum_{\tau \in G} \rho_{\tau} H \rho_{\tau^{-1}} = \frac{\text{tr } H}{n} I_n.$$

PROOF. Schur's Lemma implies that $H_0 = \frac{1}{|G|} \sum_{\tau \in G} \rho_{\tau} H \rho_{\tau^{-1}} = \lambda I_n$ for some scalar λ . Taking the trace on both sides (and using its invariance under similarity) the result $\lambda = \text{tr } H/n$ obtains. \square

Consider again the irreducible representations 1 , Sgn and β of S_3 , discussed earlier on in Section 2.7. Let $H = (h_{11}, h_{12})$ be any linear mapping from \mathbb{R}^2 into \mathbb{R} . From Schur's Lemma we know that

$$\sum_{\tau \in G} \text{Sgn}(\tau) H \beta(\tau^{-1}) = 0.$$

That is, the linear forms

$$\sum_{\tau \in G} \text{Sgn}(\tau) [h_{11} \beta_{11}(\tau^{-1}) + h_{12} \beta_{21}(\tau^{-1})], \quad \sum_{\tau \in G} \text{Sgn}(\tau) [h_{11} \beta_{12}(\tau^{-1}) + h_{12} \beta_{22}(\tau^{-1})]$$

in h_{11} and h_{12} vanish for all values of h_{11} and h_{12} . Therefore, the corresponding coefficients must be zero, that is,

$$(2.28) \quad \sum_{\tau \in G} \text{Sgn}(\tau) \beta_{11}(\tau^{-1}) = 0, \quad \sum_{\tau \in G} \text{Sgn}(\tau) \beta_{21}(\tau^{-1}) = 0,$$

$$(2.29) \quad \sum_{\tau \in G} \text{Sgn}(\tau) \beta_{21}(\tau^{-1}) = 0, \quad \sum_{\tau \in G} \text{Sgn}(\tau) \beta_{22}(\tau^{-1}) = 0.$$

The reader may verify relations (2.28) and (2.29) from Matrix (2.30).

| τ | 1 | $\text{Sgn}(\tau)$ | $\beta_{11}(\tau)$ | $\beta_{21}(\tau)$ | $\beta_{12}(\tau)$ | $\beta_{22}(\tau)$ |
|--------|---|--------------------|--------------------|--------------------|--------------------|--------------------|
| 1 | 1 | 1 | 1 | 0 | 0 | 1 |
| (12) | 1 | -1 | 0 | 1 | 1 | 0 |
| (13) | 1 | -1 | -1 | 0 | -1 | 1 |
| (23) | 1 | -1 | 1 | -1 | 0 | -1 |
| (132) | 1 | 1 | 0 | -1 | 1 | -1 |
| (123) | 1 | 1 | -1 | 1 | -1 | 0 |

This is the argument that proves

Corollary 2.7.1. For any two non-equivalent irreducible representations ρ, β of G , the relation

$$\sum_{\tau \in G} \rho_{ij}(\tau) \beta_{kl}(\tau^{-1}) = 0$$

holds for all i, j, k, ℓ indexing the entries of these representations.

Consider again the irreducible two-dimensional representation, β , of S_3 discussed in Section 2.7. From Proposition 2.7.1, we know that

$$\frac{1}{|G|} \sum_{\tau \in G} \beta_{\tau} \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \beta_{\tau^{-1}} = \frac{\text{tr } H}{2} I_2,$$

implying that, for all scalars $h_{11}, h_{12}, h_{21}, h_{22}$, we must have

$$\frac{1}{|G|} \sum_{\tau \in G} \sum_{j,k=1}^2 \beta_{ij}(\tau) h_{jk} \beta_{ki}(\tau^{-1}) = \frac{1}{2} h_{11} + \frac{1}{2} h_{22}, \quad i = 1, 2.$$

or, equivalently,

$$\begin{aligned} & \frac{1}{|G|} \left[\sum_{\tau \in G} \beta_{i1}(\tau) \beta_{1i}(\tau^{-1}) \right] h_{11} + \frac{1}{|G|} \left[\sum_{\tau \in G} \beta_{i1}(\tau) \beta_{2i}(\tau^{-1}) \right] h_{12} + \\ & \frac{1}{|G|} \left[\sum_{\tau \in G} \beta_{i2}(\tau) \beta_{1i}(\tau^{-1}) \right] h_{21} + \frac{1}{|G|} \left[\sum_{\tau \in G} \beta_{i2}(\tau) \beta_{2i}(\tau^{-1}) \right] h_{22} = \frac{1}{2} h_{11} + \frac{1}{2} h_{22}, \quad i = 1, 2, \end{aligned}$$

for all scalars $h_{11}, h_{12}, h_{21}, h_{22}$. Consequently, equating the coefficients of the linear forms, the equality $\sum_{\tau \in G} \beta_{ij}(\tau) \beta_{kl}(\tau^{-1}) = \frac{1}{2}$ when $i = l, j = k$ (and 0 otherwise) must obtain. This is the argument proving the following result:

Proposition 2.7.3. For any n -dimensional irreducible representation, ρ , of G we have

$$\frac{1}{|G|} \sum_{\tau \in G} \rho_{ij}(\tau) \rho_{kl}(\tau^{-1}) = \begin{cases} 1/n & \text{if } i = l, j = k; \\ 0 & \text{otherwise.} \end{cases}$$

Matrix (2.30) provides the numerical values for applying Proposition 2.7.3 to the irreducible representations of S_3 .

2.7.1. Orthogonality relations for characters. Following Section 2.4.1, we observe that

$$(2.31) \quad (f | g) = \frac{1}{|G|} \sum_{\tau \in G} f(\tau) \overline{g(\tau)}$$

is an inner product in the vector space $\mathcal{F}(G)$ of complex-valued functions defined in G . In particular, if χ_1 and χ_2 are characters of a representation of G , then $\chi_1, \chi_2 \in \mathcal{F}(G)$, and because $\chi(\tau^{-1}) = \overline{\chi(\tau)}$, we have

$$(\chi_1 | \chi_2) = \frac{1}{|G|} \sum_{\tau \in G} \chi_1(\tau) \overline{\chi_2(\tau)} = \frac{1}{|G|} \sum_{\tau \in G} \chi_1(\tau) \chi_2(\tau^{-1}).$$

From Section 2.4.1 we may assume that the representation ρ is unitary so that Proposition 2.7.3 can then be expressed as

$$(2.32) \quad \frac{1}{|G|} \sum_{\tau \in G} \rho_{ij}(\tau) \rho_{kl}(\tau^{-1}) = \frac{1}{|G|} \sum_{\tau \in G} \rho_{ij}(\tau) \overline{\rho_{kl}(\tau)} = (\rho_{ij} | \rho_{kl}) = \begin{cases} 1/n & \text{if } i = l, j = k, \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, Corollary 2.7.1 becomes

$$(2.33) \quad (\rho_{ij} | \beta_{kl}) = 0, \quad \text{for all } i, j, k, l,$$

where ρ and β are two non-equivalent irreducible representations of G .

Theorem 2.5. (a) If χ is the character of an irreducible representation then $(\chi | \chi) = 1$; (b) If χ_1 and χ_2 are the characters of two non-equivalent irreducible representations of a group G , then $(\chi_1 | \chi_2) = 0$.

PROOF. From expression (2.32), we have

$$(\chi | \chi) = \frac{1}{|G|} \sum_{\tau \in G} \left(\sum_{i=1}^n \rho_{ii}(\tau) \mid \sum_{j=1}^n \rho_{jj}(\tau) \right) = \sum_{i=1}^n (\rho_{ii} | \rho_{ii}) = \sum_{i=1}^n \frac{1}{n} = 1,$$

whereas, from Expression (2.33), similarly, we obtain $(\chi_1 | \chi_2) = 0$. □

We refer to the character of an irreducible representation as an *irreducible character*.

Example 2.7.1. The following matrix shows three irreducible characters of S_3 , corresponding to the irreducible representations 1, Sgn , and β discussed earlier on in Section 2.7:

$$(2.34) \quad \begin{array}{c|cccc} \tau & \chi_\rho & \chi_1 & \chi_{Sgn} & \chi_\beta \\ \hline 1 & 3 & 1 & 1 & 2 \\ (12) & 1 & 1 & -1 & 0 \\ (13) & 1 & 1 & -1 & 0 \\ (23) & 1 & 1 & -1 & 0 \\ (123) & 0 & 1 & 1 & -1 \\ (132) & 0 & 1 & 1 & -1 \end{array};$$

It also shows the character χ_ρ of the permutation representation ρ of S_n . □

The reader may verify, from Matrix 2.34, that

$$(\chi_1 | \chi_\rho) = (\chi_\beta | \chi_\rho) = (\chi_{Sgn} | \chi_\rho) = 1.$$

In fact, $(\chi_\theta | \chi_\rho)$ is the number of irreducible representations isomorphic to θ in *any* decomposition of ρ . We have, then,

Proposition 2.7.4. If ρ is a linear representation of G with character ρ and χ_1, \dots, χ_h are the irreducible characters of G , then $(\chi_i | \chi)$ is the number of representations in any decomposition of ρ that are isomorphic to ρ_i .

We remark that the notation

$$\rho \simeq m_1 \rho_1 \oplus \dots \oplus m_h \rho_h,$$

with $m_i = (\chi_i | \chi)$, indicates that there is a basis in \mathcal{V} relative to which

$$\rho(\tau) = \text{Diag} (I_{m_1} \otimes \rho_1(\tau), \dots, I_{m_h} \otimes \rho_h(\tau)), \quad \tau \in G.$$

Example 2.7.2. Let $S_2 = \{1, t\}$ act on $V = \{uu, yy, uy, yu\}$ according to $s\tau^{-1}$ (location symmetry). A representation in $\mathcal{V} = \mathbb{R}^4$ is given by

$$\rho_1 = I_4, \quad \rho_t = \begin{bmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}} \end{bmatrix},$$

which, relative to the basis for \mathcal{V} indexed by $\{uu, yy, uy + yu, uy - yu\}$, can be expressed as $1 \oplus 1 \oplus 1 \oplus Sgn$. That is $\rho \simeq 1 \oplus 1 \oplus 1 \oplus Sgn$, or

$$\rho(\tau) = \text{Diag} (1, 1, 1, Sgn(\tau)).$$

In fact, from the character table

$$\left[\begin{array}{c|ccc} \tau & \chi_\rho & \chi_1 & \chi_{Sgn} \\ \hline 1 & 4 & 1 & 1 \\ t & 2 & 1 & -1 \end{array} \right],$$

we obtain

$$(\chi_1 | \chi_\rho) = \frac{1}{2}(\chi_1(1)\chi_\rho(1) + \chi_1(t)\chi_\rho(t)) = \frac{1}{2}(4 + 2) = 3,$$

which is the multiplicity of the unit representation in this decomposition of ρ . The signature representation appears with multiplicity

$$(\chi_{Sgn} | \chi_\rho) = \frac{1}{2}(1 \times 4 + (-1) \times 2) = 1.$$

Similarly, when S_2 acts on V according to σ_s (letter symmetry), we have,

$$\rho_1 = I_4, \quad \rho_t = \begin{bmatrix} \boxed{\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}} & 0 \\ 0 & \boxed{\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}} \end{bmatrix},$$

so that

$$\rho \simeq 1 \oplus 1 \oplus \text{Sgn} \oplus \text{Sgn}.$$

The multiplicities for 1 and Sgn are, respectively,

$$(\chi_1 | \chi_\rho) = \frac{1}{2}(\chi_1(1)\chi_\rho(1) + \chi_1(t)\chi_\rho(t)) = \frac{1}{2}(4 + 0) = 2,$$

and

$$(\chi_{\text{Sgn}} | \chi_\rho) = \frac{1}{2}(\chi_{\text{Sgn}}(1)\chi_\rho(1) + \chi_{\text{Sgn}}(t)\chi_\rho(t)) = \frac{1}{2}(4 + 0) = 2.$$

Note that the multiplicity of a given irreducible component does not depend on the underlying choice of basis. Moreover, two representations with the same character are isomorphic, because they contain each irreducible component with exactly the same multiplicity. These arguments reflect the importance of characters in the study of linear representations. It is in that sense that irreducible representations are the building blocks of generic representations. \square

We may then restrict our attention to the set χ_1, \dots, χ_h of *distinct* irreducible characters of G , and write,

$$\mathcal{V} = m_1 \mathcal{V}_1 \oplus \dots \oplus m_h \mathcal{V}_h,$$

or, equivalently, $\rho \simeq m_1 \rho_1 \oplus \dots \oplus m_h \rho_h$. In this case, we have

$$(2.35) \quad \chi_\rho = m_1 \chi_1 + \dots + m_h \chi_h.$$

The multiplicities m_i are given by the integers $(\chi_\rho | \chi_i) \geq 0$, $i = 1, \dots, h$. In the previous example, under location symmetry,

$$\chi_\rho = 3\chi_1 + \chi_{\text{Sgn}}.$$

Consequently, the orthogonality relations among the irreducible components imply that $(\chi_\rho | \chi_\rho) = \sum_{i=1}^h m_i^2$. The following result is a useful characterization of the irreducible representations.

Theorem 2.6. $(\chi_\rho | \chi_\rho) = 1$ if and only if ρ is irreducible.

PROOF. We have $(\chi_\rho | \chi_\rho) = \sum_{i=1}^h m_i^2 = 1$ if and only if exactly one of the m_i 's is equal to 1 and all the others are equal to 0, in which case ρ is isomorphic to that irreducible component. \square

Example 2.7.3. Consider the irreducible representations 1, β and Sgn of S_3 , along with the tensor $\beta \otimes \beta$ representation. Matrix 2.36 shows the corresponding characters:

$$(2.36) \quad \begin{array}{c|cccc} \tau & \beta & \beta \otimes \beta & 1 & \text{Sgn} \\ \hline 1 & 2 & 4 & 1 & 1 \\ (12) & 0 & 0 & 1 & -1 \\ (13) & 0 & 0 & 1 & -1 \\ (23) & 0 & 0 & 1 & -1 \\ (123) & -1 & 1 & 1 & 1 \\ (132) & -1 & 1 & 1 & 1 \end{array}.$$

The reader may verify that

$$(\chi_\beta | \chi_\beta) = (\chi_1 | \chi_1) = (\chi_{\text{Sgn}} | \chi_{\text{Sgn}}) = 1;$$

As for the tensor representation, $(\chi_{\beta \otimes \beta} | \chi_{\beta \otimes \beta}) = 18/6 = 3$, so it must be reducible. On the other hand,

$$(\chi_{\beta \otimes \beta} | \chi_\beta) = (\chi_{\beta \otimes \beta} | \chi_1) = (\chi_{\beta \otimes \beta} | \chi_{\text{Sgn}}) = 1,$$

so that these representations appear in the decomposition of the tensor representation with single multiplicity. In fact, $\beta \otimes \beta \simeq 1 \oplus \beta \oplus \text{Sgn}$, with the corresponding character decomposition. \square

Of particular interest in the study of group representations is the *regular* representation, introduced earlier on in Example 2.4.2. It is defined by the action $\varphi(\tau, \sigma) = \tau\sigma$ in $G \times G$. Its dimension is $|G|$. Since, for all $\sigma \in G$, $\varphi(\tau, \sigma) = \varphi(\eta, \sigma)$ if and only if $\tau = \sigma$, and $\varphi(\tau, 1) = \tau$ for all $\tau \in G$, it follows that its character is given by

$$\chi_{\text{reg}}(\tau) = \begin{cases} 0 & \text{if } \tau \neq 1; \\ |G| & \text{if } \tau = 1. \end{cases}$$

Consequently, for any irreducible representation ρ of G with character χ_ρ , we have

$$(2.37) \quad (\chi_{\text{reg}}, \chi_\rho) = \frac{1}{|G|} \sum_{\tau \in G} \chi_{\text{reg}}(\tau) \chi_\rho(\tau^{-1}) = \chi_\rho(1) = \dim \rho,$$

that is, from 2.35, every irreducible representation is contained in the regular representation with multiplicity $(\chi_{\text{reg}} | \chi_\rho)$ equal to its dimension.

Proposition 2.7.5. The dimensions n_1, \dots, n_h of the h distinct irreducible representations of G , satisfy the relation $|G| = \sum_{i=1}^h n_i^2$.

PROOF. From relation 2.35, we have $\chi_{\text{reg}}(\tau) = \sum_{i=1}^h m_i \chi_i(\tau)$, where, from 2.37, $m_i = (\chi_{\text{reg}} | \chi_i) = \dim \rho_i = n_i$, so that

$$\chi_{\text{reg}}(\tau) = \sum_{i=1}^h n_i \chi_i(\tau),$$

for all $\tau \in G$. Taking $\tau = 1$, the proposed equality obtains. \square

Note that for $\tau \neq 1$, the defining property of χ_{reg} implies that $\sum_{i=1}^h n_i \chi_i(\tau) = 0$. This equality together with Proposition 2.7.5 show that

$$(2.38) \quad \frac{1}{|G|} \sum_{i=1}^h n_i \chi_i(\sigma^{-1}\tau) = \delta_{\sigma\tau}.$$

Example 2.7.4. Let $G = S_3$. The irreducible non-equivalent representations $1, \beta$ and Sgn are contained in the regular representation with multiplicities $1, 2, 1$, respectively. Because $|G| = 6 = 1^2 + 2^2 + 1^2$, these must be all the distinct irreducible non-equivalent representations of S_3 . \square

Definition 2.7.1. A scalar-valued function h defined on G and satisfying $h(\tau\sigma\tau^{-1}) = h(\sigma)$, for all $\sigma, \tau \in G$ is called a *class function*.

Clearly, class functions are constant within each conjugacy class of G . We indicate by \mathcal{C} the set of class functions on G . Note that \mathcal{C} is a linear subspace of the vector space $\mathcal{F}(G)$ of scalar-valued functions defined on G . All characters belong to \mathcal{C} . From Example 2.3.7 we observe that \mathcal{C} is a stable subspace of $\mathcal{F}(G)$ under the representation $\sigma \xrightarrow{\phi} \sigma^*$, that is,

$$x \in \mathcal{C} \implies \phi(\sigma)x = x, \quad \text{for all } \sigma \in G.$$

More precisely, \mathcal{C} is the subspace of $\mathcal{F}(G)$ of functions invariant under this conjugation action. For each class function, x , and any representation ρ , define the linear mapping

$$(2.39) \quad \widehat{x}(\rho) = \sum_{\tau \in G} x(\tau) \rho(\tau).$$

Note that $\widehat{x}(\rho)$ commutes with $\rho(\tau)$ for all $\tau \in G$. In fact,

$$\begin{aligned} \rho_\tau \widehat{x}(\rho) \rho_{\tau^{-1}} &= \rho_\tau \sum_{\sigma} x(\sigma) \rho_\sigma \rho_{\sigma^{-1}} = \sum_{\sigma} x(\sigma) \rho_\tau \rho_\sigma \rho_{\sigma^{-1}} = \sum_{\sigma} x(\sigma) \rho_{\tau\sigma\tau^{-1}} \\ &= \sum_{\sigma} x(\tau\sigma\tau^{-1}) \rho_{\tau\sigma\tau^{-1}} = \sum_{\sigma} x(\sigma) \rho_\sigma = \widehat{x}(\rho). \end{aligned}$$

Therefore, if ρ is an irreducible representation, it follows from Schur's Lemma that $\widehat{x}(\rho) = \lambda I$. To evaluate λ we take the trace in each side of the above equality, to obtain

$$\begin{aligned} \text{tr } \widehat{x}(\rho) &= \sum_{\tau \in G} x(\tau) \text{tr } \rho(\tau) = \sum_{\tau \in G} x(\tau) \chi_{\rho}(\tau) = \sum_{\tau \in G} x(\tau) \overline{\chi}_{\rho}(\tau^{-1}) \\ &= |G|(x, \overline{\chi}_{\rho}) = \text{tr } \lambda I_n = n\lambda, \end{aligned}$$

so that $\lambda = |G|(x, \overline{\chi}_{\rho})/n$. This proves

Proposition 2.7.6. If ρ is an n -dimensional irreducible representation of G and $x \in \mathcal{C}$ then

$$\widehat{x}(\rho) = \frac{|G|}{n}(x | \overline{\chi}_{\rho}) I_n.$$

Theorem 2.7. The distinct irreducible characters form an orthonormal basis for \mathcal{C} .

PROOF. From Theorem 2.5 we know that the set of distinct irreducible characters form an orthonormal set of functions in \mathcal{C} . We need to show that this set generates \mathcal{C} . Suppose that $x \in \mathcal{C}$ and that x is orthogonal to $\overline{\chi}_1, \dots, \overline{\chi}_h$. Therefore, for any irreducible n -dimensional representation ρ of G , we have

$$\widehat{x}(\rho) = \frac{|G|}{n}(x | \overline{\chi}_{\rho}) I_n = 0.$$

Because every representation decomposes isomorphically as a sum of irreducible components, it follows that $\widehat{x}(\rho) = 0$ for every representation ρ . In particular, $\widehat{x}(\rho_{\text{reg}}) = 0$, $\{e_{\tau} : \tau \in G\}$ is a basis for \mathcal{V} , and

$$0 = \widehat{x}(\rho_{\text{reg}})e_1 = \sum_{\tau \in G} x(\tau) \rho_{\text{reg}}(\tau)e_1 = \sum_{\tau \in G} x(\tau)e_{\tau},$$

which implies $x(\tau) = 0$ for all $\tau \in G$. That is, $x = 0$. \square

Note that the dimension of the subspace \mathcal{C} of class functions is determined both by the number of distinct irreducible representations of G and by the number of orbits, or *conjugacy classes*, of G under the action $\sigma\tau\sigma^{-1}$, in which the class functions can be arbitrarily defined. Consequently, the number of distinct irreducible representations coincide with the number of conjugacy classes of G .

Example 2.7.5. If G is a commutative group, then G has $|G|$ conjugacy classes and hence $|G|$ distinct irreducible representations. Moreover, because

$$|G| = \sum_j \dim^2 \rho_j,$$

we conclude that these representations are all one-dimensional. In particular, if G is cyclic, the irreducible representations are given by $\rho_j(\tau^k) = e^{2\pi ijk/|G|}$. \square

Proposition 2.7.7. If χ_1, \dots, χ_h are the distinct irreducible characters of group G , then

$$\sum_i \overline{\chi}_i(\eta) \chi_i(\tau) = \begin{cases} \frac{|G|}{|\mathcal{O}_{\tau}|} & \text{if } \eta \in \mathcal{O}_{\tau}; \\ 0 & \text{if } \eta \notin \mathcal{O}_{\tau}, \end{cases}$$

where $|\mathcal{O}_{\tau}|$ is the number of elements in the conjugacy class $\mathcal{O}_{\tau} = \{\sigma\tau\sigma^{-1}, \sigma \in G\}$ of $\tau \in G$.

PROOF. Define

$$x_{\tau}(\eta) = \begin{cases} 1 & \text{if } \eta \in \mathcal{O}_{\tau}; \\ 0 & \text{if } \eta \notin \mathcal{O}_{\tau}. \end{cases}$$

Then x_{τ} is a class function and, consequently, can be expressed as a linear combination $\sum_i c_i \chi_i$ of the distinct irreducible characters χ_1, \dots, χ_h of G . The reader may verify that, in this case, $c_i = (x_{\tau} | \chi_i) = |\mathcal{O}_{\tau}| \overline{\chi}_i(\tau)/|G|$, so that

$$x_{\tau}(\eta) = \sum_i \frac{|\mathcal{O}_{\tau}|}{|G|} \overline{\chi}_i(\tau) \chi_i(\eta) = \begin{cases} 1 & \text{if } \eta \in \mathcal{O}_{\tau} \\ 0 & \text{if } \eta \notin \mathcal{O}_{\tau}, \end{cases}$$

from which the result follows. \square

Example 2.7.6. Matrix (2.40) shows the irreducible characters $\chi_1, \chi_{\text{Sgn}}, \chi_\beta$ of S_3 , along with the characters $\chi_\rho, \chi_{\beta \otimes \beta}, \chi_{\text{reg}}$ of the permutation, tensor $\beta \otimes \beta$ and regular representations, respectively.

(2.40)

| τ | χ_ρ | χ_1 | χ_{Sgn} | χ_β | $\chi_{\beta \otimes \beta}$ | χ_{reg} |
|--------|-------------|----------|---------------------|--------------|------------------------------|---------------------|
| 1 | 3 | 1 | 1 | 2 | 4 | 6 |
| (12) | 1 | 1 | -1 | 0 | 0 | 0 |
| (13) | 1 | 1 | -1 | 0 | 0 | 0 |
| (23) | 1 | 1 | -1 | 0 | 0 | 0 |
| (123) | 0 | 1 | 1 | -1 | 1 | 0 |
| (132) | 0 | 1 | 1 | -1 | 1 | 0 |

S_3 has three conjugate orbits (and hence three distinct irreducible representations),

$$\mathcal{O}_1 = \{1\}, \quad \mathcal{O}_t = \{(12), (13), (23)\}, \quad \mathcal{O}_c = \{(123), (132)\}.$$

We obtain

$$\bar{\chi}_1(\tau)\chi_1(\tau) + \bar{\chi}_{\text{Sgn}}(\tau)\chi_{\text{Sgn}}(\tau) + \bar{\chi}_\beta(\tau)\chi_\beta(\tau) = \begin{cases} 4 + 1 + 1 = 6 = |G|/|\mathcal{O}_1|, & \text{if } \tau \in \mathcal{O}_1; \\ 0 + 1 + 1 = 2 = |G|/|\mathcal{O}_t|, & \text{if } \tau \in \mathcal{O}_t; \\ 1 + 1 + 1 = 3 = |G|/|\mathcal{O}_c|, & \text{if } \tau \in \mathcal{O}_c, \end{cases}$$

whereas

$$\begin{aligned} & \bar{\chi}_1(\tau)\chi_1(\eta) + \bar{\chi}_{\text{Sgn}}(\tau)\chi_{\text{Sgn}}(\eta) + \bar{\chi}_\beta(\tau)\chi_\beta(\eta) = \\ & \begin{cases} 2 \times 0 + 1 \times (-1) + 1 \times 1 = 0, & \text{if } \tau = 1, \eta = (12); \\ 2 \times (-1) + 1 \times 1 + 1 \times 1 = 0, & \text{if } \tau = 1, \eta = (123); \\ 0 \times (-1) + (-1) \times 1 + 1 \times 1 = 0, & \text{if } \tau = (12), \eta = (123). \end{cases} \end{aligned}$$

To decompose, say, the character of $\beta \otimes \beta$, we write $\chi_{\beta \otimes \beta} = c_1\chi_1 + c_{\text{Sgn}}\chi_{\text{Sgn}} + c_\beta\chi_\beta$, in which the coefficients are determined by

$$c_1 = (\chi_{\beta \otimes \beta} | \chi_1) = 6/6 = 1, \quad c_{\text{Sgn}} = (\chi_{\beta \otimes \beta} | \chi_{\text{Sgn}}) = 6/6 = 1,$$

and $c_\beta = (\chi_{\beta \otimes \beta} | \chi_\beta) = 6/6 = 1$. In fact, $\chi_{\beta \otimes \beta} = \chi_1 + \chi_{\text{Sgn}} + \chi_\beta$. \square

2.8. The canonical decomposition

Consider again the representations of $S_2 = \{1, t\}$ acting on the space V of binary sequences in length of two according to position and letter symmetry, introduced earlier on in Section 2.4, see also Example 2.3.1. We examine first the position symmetry. The representation is isomorphic to

$$\xi_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \xi_t = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

taking values in $GL(\mathbb{R}^4)$. Equivalently, writing $\mathcal{V} = \mathbb{R}^4$, we observe that the ξ determines the decomposition

$$\mathcal{V} = \mathcal{W}_1 \oplus \mathcal{W}_2 \oplus \mathcal{W}_3,$$

where the stable subspaces \mathcal{W}_1 and \mathcal{W}_2 reduce isomorphically as a subspace \mathcal{U}_1 associated with the unit representation and \mathcal{W}_3 reduces as the sum of an isomorphic copy of \mathcal{U}_1 and a subspace \mathcal{U}_{Sgn} associated with the sign representation. Therefore,

$$\mathcal{V} = \underbrace{\mathcal{U}_1 \oplus \mathcal{U}_1 \oplus \mathcal{U}_1}_{\mathcal{V}_1} \oplus \mathcal{U}_{\text{Sgn}} = \mathcal{V}_1 \oplus \mathcal{V}_{\text{Sgn}},$$

showing a decomposition of \mathcal{V} into a direct sum of the irreducible representations of S_2 in which we collected together the isomorphic copies. This is the *canonical decomposition* of \mathcal{V} . Next, we will construct projections \mathcal{P}_1 and \mathcal{P}_{Sgn} of \mathcal{V} on the irreducible subspaces \mathcal{V}_1 and \mathcal{V}_{Sgn} . To do this, define

$$\mathcal{P}_\beta = \frac{n_\beta}{|\mathbf{G}|} \sum_{\tau \in \mathbf{G}} \bar{\chi}_\beta(\tau) \xi_\tau,$$

where χ_β is the irreducible character of the irreducible representation β of \mathbf{G} , and n_β its dimension. In the present case ($\mathbf{G} = S_2$), $\beta \in \{1, \text{Sgn}\}$ with corresponding characters

$$\left[\begin{array}{c|cc} & 1 & t \\ \hline \chi_1 & 1 & 1 \\ \chi_{\text{Sgn}} & 1 & -1 \end{array} \right],$$

so that

$$\mathcal{P}_1 = \frac{1}{2}[\xi_1 + \xi_t], \quad \mathcal{P}_{\text{Sgn}} = \frac{1}{2}[\xi_1 - \xi_t].$$

When the projections are evaluated relative to a basis of $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_{\text{Sgn}}$ on which

$$\xi_\tau = \text{Diag}(\mathbf{I}_3 \otimes \mathbf{1}_\tau, \text{Sgn}_\tau)$$

we obtain

$$\mathcal{P}_1 = \frac{1}{2}[\text{Diag}(\mathbf{I}_3 \otimes \mathbf{1}, 1) + \text{Diag}(\mathbf{I}_3 \otimes \mathbf{1}, -1)] = \text{Diag}(\mathbf{I}_3 \otimes \mathbf{1}, 0) = \text{Diag}(1, 1, 1, 0),$$

and

$$\mathcal{P}_{\text{Sgn}} = \frac{1}{2}[\text{Diag}(\mathbf{I}_3 \otimes \mathbf{1}, 1) - \text{Diag}(\mathbf{I}_3 \otimes \mathbf{1}, -1)] = \text{Diag}(\mathbf{I}_3 \otimes \mathbf{0}, 1) = \text{Diag}(0, 0, 0, 1).$$

It then follows that

- (1) $\mathcal{P}_1^2 = \mathcal{P}_1$, $\mathcal{P}_{\text{Sgn}}^2 = \mathcal{P}_{\text{Sgn}}$;
- (2) $\mathcal{P}_1 \mathcal{P}_{\text{Sgn}} = \mathcal{P}_{\text{Sgn}} \mathcal{P}_1 = 0$;
- (3) $\mathbf{I}_4 = \mathcal{P}_1 + \mathcal{P}_{\text{Sgn}}$,

so that \mathcal{P}_β is a projection on a subspace isomorphic to \mathcal{V}_β . Note that properties (1),(2) and (3) above remain valid if \mathcal{P}_β is evaluated relative to any representation $M\xi_\tau M^{-1}$ equivalent to ξ . In this case, \mathcal{P}_β transforms as $M\mathcal{P}_\beta M^{-1}$ and (1),(2) and (3) remain unchanged. For example, relative to the basis for \mathcal{V} indexed by $\{uu, yy, uy, yu\}$, we have

$$\mathcal{P}_1 = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} & \\ & & & \end{bmatrix}, \quad \mathcal{P}_{\text{Sgn}} = \begin{bmatrix} 0 & & & \\ & 0 & & \\ & & \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix} & \\ & & & \end{bmatrix}.$$

Here is an outline of the same construction when ξ (letter symmetry) is given by

$$\xi_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \xi_t = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

In this case $\mathcal{V} = \mathcal{W}_1 \oplus \mathcal{W}_2$ and in each one of these two stable subspaces ξ reduces isomorphically as the sum of the unit and the sign representations. Collecting the isomorphic copies of the corresponding irreducible subspaces \mathcal{U}_1 and \mathcal{U}_{Sgn} we have then

$$\mathcal{V} = \underbrace{\mathcal{U}_1 \oplus \mathcal{U}_1}_{\mathcal{V}_1} \oplus \underbrace{\mathcal{U}_{\text{Sgn}} \oplus \mathcal{U}_{\text{Sgn}}}_{\mathcal{V}_{\text{Sgn}}}.$$

This is the canonical decomposition of \mathcal{V} . The corresponding projections are:

$$\mathcal{P}_1 = \frac{1}{2}[\xi_1 + \xi_t], \quad \mathcal{P}_{\text{Sgn}} = \frac{1}{2}[\xi_1 - \xi_t].$$

When the projections are evaluated relative to a basis of $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_{\text{Sgn}}$ on which

$$\xi_\tau = \text{Diag}(\mathbf{I}_2 \otimes 1_\tau, \mathbf{I}_2 \otimes \text{Sgn}_\tau)$$

we obtain

$$\mathcal{P}_1 = \frac{1}{2}[\text{Diag}(\mathbf{I}_2 \otimes 1, \mathbf{I}_2 \otimes 1) + \text{Diag}(\mathbf{I}_2 \otimes 1, \mathbf{I}_2 \otimes -1)] = \text{Diag}(\mathbf{I}_2 \otimes 1, \mathbf{I}_2 \otimes 0) = \text{Diag}(1, 1, 0, 0),$$

and

$$\mathcal{P}_{\text{Sgn}} = \frac{1}{2}[\text{Diag}(\mathbf{I}_2 \otimes 1, \mathbf{I}_2 \otimes 1) - \text{Diag}(\mathbf{I}_2 \otimes 1, \mathbf{I}_2 \otimes -1)] = \text{Diag}(\mathbf{I}_2 \otimes 0, \mathbf{I}_2 \otimes 1) = \text{Diag}(0, 0, 1, 1).$$

It then follows that \mathcal{P}_β is a projection on a subspace isomorphic to \mathcal{V}_β and, regardless of the chosen basis for \mathcal{V} ,

- (1) $\mathcal{P}_1^2 = \mathcal{P}_1$, $\mathcal{P}_{\text{Sgn}}^2 = \mathcal{P}_{\text{Sgn}}$;
- (2) $\mathcal{P}_1 \mathcal{P}_{\text{Sgn}} = \mathcal{P}_{\text{Sgn}} \mathcal{P}_1 = 0$;
- (3) $\mathbf{I}_4 = \mathcal{P}_1 + \mathcal{P}_{\text{Sgn}}$.

In particular,

$$\mathcal{P}_1 = \frac{1}{2} \begin{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} & \\ & \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \end{bmatrix}, \quad \mathcal{P}_{\text{Sgn}} = \frac{1}{2} \begin{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} & \\ & \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \end{bmatrix}.$$

The arguments illustrated in the above example will now be applied to prove the following theorem.

Theorem 2.8 (Canonical Decomposition). Let ρ be a linear representation of G into $\text{GL}(\mathcal{V})$, ρ_1, \dots, ρ_h the distinct non-isomorphic irreducible representations of G , with corresponding characters χ_1, \dots, χ_h and dimensions n_1, \dots, n_h . Then,

$$\mathcal{P}_i = \frac{n_i}{|G|} \sum_{\tau \in G} \bar{\chi}_i(\tau) \rho(\tau),$$

is a projection of \mathcal{V} onto a subspace \mathcal{V}_i , sum of m_i isomorphic copies of the irreducible subspaces associated with ρ_i , $i = 1, \dots, h$. Moreover, $\mathcal{P}_i \mathcal{P}_j = 0$, for $i \neq j$, $\mathcal{P}_i^2 = \mathcal{P}_i$ and $\sum_i \mathcal{P}_i = \mathbf{I}_v$, where $v = \dim \mathcal{V} = \sum_{i=1}^h m_i n_i$.

PROOF. From Proposition 2.7.4 we know that $\rho \simeq \sum_{j=1}^h m_j \rho_j$, where ρ_1, \dots, ρ_h are the distinct irreducible representations of G . That is, there is a basis in \mathcal{V} relative to which

$$\rho = \text{Diag}(\mathbf{I}_{m_1} \otimes \rho_1, \dots, \mathbf{I}_{m_h} \otimes \rho_h).$$

Therefore,

$$\mathcal{P}_i = \frac{n_i}{|G|} \text{Diag}(\mathbf{I}_{m_1} \otimes \sum_{\tau} \bar{\chi}_i(\tau) \rho_1(\tau), \dots, \mathbf{I}_{m_h} \otimes \sum_{\tau} \bar{\chi}_i(\tau) \rho_h(\tau)).$$

Applying Proposition 2.7.6 with $x = \chi_i$, so that $\hat{x}(\rho_j) = \sum_{\tau} \chi_i(\tau) \rho_j(\tau)$, we have

$$\sum_{\tau} \bar{\chi}_i(\tau) \rho_j(\tau) = \frac{|G|}{n_i} (\chi_j | \chi_i) \mathbf{I}_{n_i},$$

and consequently

$$\mathcal{P}_i = \text{Diag}(\delta_{i1} \mathbf{I}_{m_1} \otimes \mathbf{I}_{n_1}, \dots, \delta_{ih} \mathbf{I}_{m_h} \otimes \mathbf{I}_{n_h}).$$

It is then clear that $\mathcal{P}_i^2 = \mathcal{P}_i$, so that \mathcal{P}_i is a projection of \mathcal{V} into the subspace \mathcal{V}_i direct sum of m_i copies of the irreducible subspaces associated with ρ_i , $i = 1, \dots, h$. It is also clear that, in addition, $\mathcal{P}_i \mathcal{P}_j = 0$ for $j \neq i$ and that

$$\sum_{i=1}^h \mathcal{P}_i = \text{Diag}(\mathbf{I}_{m_1} \otimes \mathbf{I}_{n_1}, \dots, \mathbf{I}_{m_h} \otimes \mathbf{I}_{n_h}) = \mathbf{I}_v,$$

concluding the proof. \square

Note that $\text{tr } \mathcal{P}_i = n_i m_i = \dim \mathcal{V}_i$.

2.9. The standard decomposition

In this section we will characterize the canonical decomposition applied to the permutation representation ρ of S_n . This is the reduction that is naturally associated with data that are indexed by $V = \{1, \dots, n\}$, such as in statistical sampling. Recall that ρ is defined by the action of S_n on the set of indices $V = \{1, \dots, n\}$ for the canonical basis for \mathbb{R}^n according to $(\tau, j) = \tau j$. The resulting representation is an isomorphism between S_n and M_n , the group of $n \times n$ permutation matrices.

It will be useful to adopt the notation

$$\mathcal{A} = \frac{1}{n} ee', \quad \mathcal{Q} = I - \mathcal{A}$$

from now on, where ee' is the $n \times n$ matrix of ones. Clearly, the reduction

$$\boxed{I = \mathcal{A} + \mathcal{Q}}$$

satisfies $\mathcal{A}^2 = \mathcal{A}$, $\mathcal{Q}^2 = \mathcal{Q}$ and $\mathcal{A}\mathcal{Q} = \mathcal{Q}\mathcal{A} = 0$. Moreover, \mathcal{A} projects $\mathcal{V} = \mathbb{R}^n$ into a subspace \mathcal{V}_a of dimension $\dim \mathcal{V}_a = \text{tr } \mathcal{A} = 1$ generated by $e = e_1 + \dots + e_n = (1, 1, \dots, 1) \in \mathcal{V}$, whereas \mathcal{Q} projects \mathcal{V} into the subspace \mathcal{V}_q in dimension $n - 1$, the orthogonal complement of \mathcal{V}_a in \mathcal{V} . We will show that the reduction $\mathcal{V} = \mathcal{V}_a + \mathcal{V}_q$ is exactly the canonical reduction determined by ρ . We refer to this decomposition as the *standard* decomposition or standard reduction.

To illustrate the argument, consider first the case $n = 3$. The joint character table for ρ and the irreducible representations of S_3 is

$$\left[\begin{array}{c|ccc} \chi & 1 & (12) & (123) \\ \hline \chi_\rho & 3 & 1 & 0 \\ \chi_1 & 1 & 1 & 1 \\ \chi_\beta & 2 & 0 & -1 \\ \chi_{\text{Sgn}} & 1 & -1 & 1 \end{array} \right],$$

where β is the two-dimensional irreducible representation derived earlier on in Example 2.5.4. Recall also that there are 3 elements in the class of (12) and two elements in the class of (123). It then follows that $(\chi_1 | \chi_\rho) = 1$, $(\chi_{\text{Sgn}} | \chi_\rho) = 0$ and $(\chi_\beta | \chi_\rho) = 1$, so that $\rho \simeq 1 \oplus \beta$ and $\chi_\beta = \chi_\rho - 1$. In general, we have:

Proposition 2.9.1. $\chi_\beta = \chi_\rho - 1$ is an irreducible character of S_n . Its dimension is $n - 1$.

PROOF. Write $\chi_\rho = \chi$ to indicate the character of the permutation representation of S_n . To evaluate

$$(\chi_\beta | \chi_\beta) = \frac{1}{|G|} \sum_{\tau} (\chi(\tau) - 1)^2 = \frac{1}{|G|} \sum_{\tau} (\chi^2(\tau) - 2\chi(\tau) + 1)$$

and verify the irreducibility criteria $(\chi_\beta | \chi_\beta) = 1$ of Proposition 2.6, we need the first two moments

$$\sum_{\tau} \chi(\tau)/|G|, \quad \sum_{\tau} \chi^2(\tau)/|G|,$$

of χ . The argument is as follows: Consider the action $(\tau i, \tau j)$ of S_n on the product space $V^2 = \{1, \dots, n\}^2$. Its character is χ^2 and the number of orbits is clearly two, namely

$$\mathcal{O}_0 = \{(i, i), i = 1, \dots, n\}, \quad \mathcal{O}_1 = \{(i, j), i, j = 1, \dots, n, i \neq j\}.$$

Now apply Burnside's Lemma to write

$$2 = \text{Number of orbits in } V^2 = \frac{1}{|G|} \sum_{\tau} \chi^2(\tau).$$

Similarly, since S_n acts transitively on $\{1, \dots, n\}$, we have

$$1 = \text{Number of orbits in } V = \frac{1}{|G|} \sum_{\tau} \chi(\tau).$$

Consequently,

$$(\chi_\beta | \chi_\beta) = \frac{1}{|G|} \sum_{\tau} (\chi^2(\tau) - 2\chi(\tau) + 1) = (2 - 2 + 1) = 1,$$

thus showing that χ_β is an irreducible character of S_n . Its dimension is $\chi_\beta(1) = \chi(1) - 1 = n - 1$, concluding the proof. \square

Consequently, the multiplicities $(\chi_1 | \chi_\rho) = 1$ and $(\chi_\beta | \chi_\rho) = 1$ apply to the permutation representation of S_n , and because its dimension is n we conclude that $\rho \simeq 1 \oplus \beta$ is an irreducible decomposition of ρ .

The implication for the canonical decomposition of the permutation representation of S_n is as follows: Because $\rho \simeq 1 \oplus \beta$ there are only two (non-null) projections, namely \mathcal{P}_1 associated with the symmetric character, and \mathcal{P}_β associated with the irreducible character χ_β of dimension $n - 1$. Clearly,

$$\mathcal{P}_1 = \frac{1}{|G|} \sum_{\tau} \rho(\tau) = \mathcal{A};$$

Moreover, $I = \mathcal{A} + \mathcal{Q} = \mathcal{A} + \mathcal{P}_\beta$ so we must have $\mathcal{P}_\beta = \mathcal{Q}$. That is, \mathcal{A} and \mathcal{Q} are the only canonical projections associated with the permutation representation of S_n , which is the characterization we had in mind.

The results summarized in the following proposition are useful in obtaining new orthogonal decompositions from existing ones. Its proof is by direct verification that in each case the appropriate identity matrix decomposes as a sum of pairwise orthogonal idempotents components.

Proposition 2.9.2. If $I_m = \sum_i \mathcal{P}_i$ and $I_n = \sum_j \mathcal{T}_j$ are canonical reductions of dimensions m and n respectively, then

$$I_{mn} = \sum_{i,j} \mathcal{P}_i \otimes \mathcal{T}_j$$

and $\mathcal{P}_i \otimes \mathcal{T}_j$ are orthogonal projections. In particular

$$(\mu \otimes e)' (\mathcal{P} \otimes \mathcal{T}) (\mu \otimes e) = \begin{cases} n\mu' \mathcal{P} \mu & \text{if } \mathcal{T} = \mathcal{A} \\ 0 & \text{if } \mathcal{T} = \mathcal{Q} \end{cases}.$$

If L_1, \dots, L_h is a disjoint partition of L then the amalgamated components of

$$I_m = \left(\sum_{i \in L_1} \mathcal{P}_i \right) + \dots + \left(\sum_{i \in L_h} \mathcal{P}_i \right)$$

are orthogonal projections. If, in addition, $m = n$ and the components \mathcal{P}_i and \mathcal{T}_j all commute, then

$$I_n = \sum_{i,j} \mathcal{P}_i \mathcal{T}_j$$

and $\mathcal{P}_i \mathcal{T}_j$ are orthogonal projections. In particular, the components \mathcal{A} and \mathcal{Q} of the standard reduction commute with every symmetric matrix of same dimension. \square

The following result describes the matrices that are centralized by the permutation representation of S_n . In multivariate analysis, these matrices play a significant role in describing the (intra)class covariance structure of permutation symmetric random variables.

Proposition 2.9.3. If ρ is the permutation representation of S_n , then, for every real or complex $n \times n$ matrix H ,

$$\frac{1}{n!} \sum_{\tau \in S_n} \rho_\tau H \rho_{\tau^{-1}} = a_0 e e' + a_1 I_n,$$

where the coefficients a_0 and a_1 are scalars defined by the relations $n(a_0 + a_1) = \text{tr } H$ and $n(n - 1)a_0 = e' H e - \text{tr } H$, in which $e' H e$ is the sum of the entries in H .

PROOF. Let $M = \frac{1}{n!} \sum_{\tau \in S_n} \rho_\tau H \rho_{\tau^{-1}}$ and let $J = PHP^{-1}$ where

$$(2.41) \quad P = \begin{bmatrix} 1 & 1 & \dots & 1 & 1 \\ n-1 & -1 & \dots & -1 & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & -1 & \dots & n-1 & -1 \end{bmatrix}.$$

It is simple to verify that the irreducible decomposition $\rho \simeq 1 \oplus \beta$ of the permutation representation is realized by $P\rho P^{-1}$. Consequently, applying Proposition 2.7.2, we have

$$PMP^{-1} = \frac{1}{n!} \sum_{\tau} (P\rho_\tau P^{-1})J(P\rho_{\tau^{-1}} P^{-1}) = \begin{bmatrix} J_{11} & 0 \\ 0 & \frac{\text{tr } J_{22}}{n-1} I_{n-1} \end{bmatrix},$$

from which we obtain

$$M = P^{-1} \begin{bmatrix} J_{11} & 0 \\ 0 & \frac{\text{tr } J_{22}}{n-1} I_{n-1} \end{bmatrix} P.$$

Direct evaluation, using the definition of the matrix P , shows that M is the matrix with entries

$$M_{ij} = \begin{cases} \text{tr } H/n & \text{if } i = j; \\ (e'He - \text{tr } H)/(n-1) & \text{if } i \neq j, \end{cases}$$

which is the proposed result. \square

2.10. Inference

We conclude this chapter with examples of prototypic applications of canonical decompositions and their connection with the Fisher-Cochran's Theorem for quadratic forms.

The aspects of statistical inference associated with the canonical reduction are those of the distribution of the corresponding quadratic forms. These results are known as the Fisher-Cochran theorem. See, for example, Rao (1973), Eaton (1983), Searle (1971) or Muirhead (1982). The corollary that relates to a canonical reduction can be formulated as follows:

Proposition 2.10.1. If the components of $y' = (y_1, \dots, y_n)$ are independent and normally distributed with mean μ_i and unit variance, then the components $y'\mathcal{P}_i y$ of the canonical reduction

$$y'y = \sum_i y'\mathcal{P}_i y$$

are independent and distributed as χ^2 with $\text{tr } \mathcal{P}_i$ degrees of freedom and noncentrality parameter $\mu'\mathcal{P}_i \mu$.

PROOF. Let X indicate the matrix in which the columns are the normalized characteristic vectors from $\mathcal{P}_1, \mathcal{P}_2, \dots$. The orthogonality of the associated subspaces, implied by $\mathcal{P}_i \mathcal{P}_j = 0$ for any $i \neq j$, allows that $X'X = I$. Moreover, because $\mathcal{P}_i^2 = \mathcal{P}_i$, the characteristic roots of \mathcal{P}_i are either 0 or 1 and, and consequently, in the new base $z = X'y$ we have

$$y'\mathcal{P}_i y = z'X'\mathcal{P}_i Xz = z' \text{diag} (0, \dots, 0, \underbrace{1, \dots, 1}_{\mathcal{W}_i}, 0, \dots, 0)z = \sum_{\mathcal{W}_i} z_i^2$$

where \mathcal{W}_i is the corresponding invariant subspace in the decomposition $\mathcal{W}_1 \oplus \mathcal{W}_2 \oplus \dots$ of the full vector space. From the parametric assumptions on the components of y , we obtain the law of $y'\mathcal{P}_i y$ as χ^2 with degrees of freedom equal to the dimension $\text{tr } \mathcal{P}_i$ of \mathcal{W}_i , independently of the $y'\mathcal{P}_j y$, $j \neq i$, with noncentrality parameter equal to $\mu'\mathcal{P}_i \mu$. \square

The following example illustrates the construction described in Proposition 2.10.1. Consider the regular reduction of D_4 . The projections are $\mathcal{P}_1 = ee'/8$,

$$\mathcal{P}_2 = 1/8 \begin{bmatrix} 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \end{bmatrix},$$

$$\mathcal{P}_3 = 1/8 \begin{bmatrix} 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \end{bmatrix},$$

$$\mathcal{P}_4 = 1/8 \begin{bmatrix} 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \end{bmatrix},$$

$$\mathcal{P}_5 = 1/4 \begin{bmatrix} 2 & 0 & -2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & -2 & 0 & 0 & 0 & 0 \\ -2 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & -2 \\ 0 & 0 & 0 & 0 & -2 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & -2 & 0 & 2 \end{bmatrix},$$

so that the invariants and corresponding subspaces are

$$\left[\begin{array}{l} \mathcal{W}_1 \mid m_1 + m_2 + m_3 + m_4 + m_5 + m_6 + m_7 + m_8 \\ \mathcal{W}_2 \mid m_1 + m_2 + m_3 + m_4 - m_5 - m_6 - m_7 - m_8 \\ \mathcal{W}_3 \mid m_1 - m_2 + m_3 - m_4 + m_5 - m_6 + m_7 - m_8 \\ \mathcal{W}_4 \mid m_1 - m_2 + m_3 - m_4 - m_5 + m_6 - m_7 + m_8 \\ \mathcal{W}_5 \mid 2m_1 - 2m_3 \\ \quad \mid 2m_2 - 2m_4 \\ \quad \mid 2m_5 - 2m_7 \\ \quad \mid 2m_6 - 2m_8 \end{array} \right],$$

from which we obtain the matrix of orthogonal normalized characteristic vectors

$$X = \begin{bmatrix} 1/4\sqrt{2} & 1/4\sqrt{2} & 1/4\sqrt{2} & 1/4\sqrt{2} & 1/2\sqrt{2} & 0 & 0 & 0 \\ 1/4\sqrt{2} & 1/4\sqrt{2} & -1/4\sqrt{2} & -1/4\sqrt{2} & 0 & 1/2\sqrt{2} & 0 & 0 \\ 1/4\sqrt{2} & 1/4\sqrt{2} & 1/4\sqrt{2} & 1/4\sqrt{2} & -1/2\sqrt{2} & 0 & 0 & 0 \\ 1/4\sqrt{2} & 1/4\sqrt{2} & -1/4\sqrt{2} & -1/4\sqrt{2} & 0 & -1/2\sqrt{2} & 0 & 0 \\ 1/4\sqrt{2} & -1/4\sqrt{2} & 1/4\sqrt{2} & -1/4\sqrt{2} & 0 & 0 & 1/2\sqrt{2} & 0 \\ 1/4\sqrt{2} & -1/4\sqrt{2} & -1/4\sqrt{2} & 1/4\sqrt{2} & 0 & 0 & 0 & 1/2\sqrt{2} \\ 1/4\sqrt{2} & -1/4\sqrt{2} & 1/4\sqrt{2} & -1/4\sqrt{2} & 0 & 0 & -1/2\sqrt{2} & 0 \\ 1/4\sqrt{2} & -1/4\sqrt{2} & -1/4\sqrt{2} & 1/4\sqrt{2} & 0 & 0 & 0 & -1/2\sqrt{2} \end{bmatrix}.$$

We have $X'X = I$ and

$$y'P_i y = z'X'P_i X z = \begin{cases} z_1^2 & \text{in } \mathcal{W}_1 \\ z_2^2 & \text{in } \mathcal{W}_2 \\ z_3^2 & \text{in } \mathcal{W}_3 \\ z_4^2 & \text{in } \mathcal{W}_4 \\ z_5^2 + z_6^2 + z_7^2 + z_8^2 & \text{in } \mathcal{W}_5. \end{cases}$$

The degrees of freedom are, respectively, 1, 1, 1, 1 and 4, and the noncentrality parameters

$$\mu'P_i \mu = \begin{cases} (m_1 + m_2 + m_3 + m_4 + m_5 + m_6 + m_7 + m_8)^2/8 & \text{in } \mathcal{W}_1 \\ (m_1 + m_2 + m_3 + m_4 - m_5 - m_6 - m_7 - m_8)^2/8 & \text{in } \mathcal{W}_2 \\ (m_1 - m_2 + m_3 - m_4 + m_5 - m_6 + m_7 - m_8)^2/8 & \text{in } \mathcal{W}_3 \\ (m_1 - m_2 + m_3 - m_4 - m_5 + m_6 - m_7 + m_8)^2/8 & \text{in } \mathcal{W}_4 \\ [(2m_1 - 2m_3)^2 + (2m_2 - 2m_4)^2 + (2m_5 - 2m_7)^2 + (2m_6 - 2m_8)^2]/4 & \text{in } \mathcal{W}_5. \end{cases}$$

Related results. Direct evaluation, e.g., Searle (1971) of the cumulants of $x'P x$ shows that, for $x \sim N(\mu, \Sigma)$,

$$E(x'P x) = \text{tr } P\Sigma + \mu'P\mu, \quad \text{Var}(x'P x) = 2\text{tr } P\Sigma + 4\mu'P\Sigma P\mu.$$

Moreover, e.g. Muirhead (1982), for any symmetric projection P , the distribution of $x'P x$ is $\chi_k^2(\delta)$, where $k = \text{tr } P$ and $\delta = \mu'P\mu$, if and only if $P\Sigma$ is idempotent. If P_1 and P_2 are symmetric projections such that $P_1\Sigma P_2 = 0$ then $x'P_1 x$ and $x'P_2 x$ are independent.

Example 2.10.1. Consider the simple structure

$$V = \{(0, 0), (1, 0), (0, 1), (2, 0), (0, 2), (1, 1)\} = \{\alpha, x, y, X, Y, \gamma\},$$

where S_2 acts according to (τ_i, τ_j) , $\tau \in S_2$. The resulting permutation table is,

$$\left[\begin{array}{c|cccccc} S_2 & \alpha & x & y & X & Y & \gamma \\ \hline 1 & \alpha & x & y & X & Y & \gamma \\ \hline \tau & \alpha & y & x & Y & X & \gamma \end{array} \right],$$

from which we obtain the canonical reduction

$$P_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & -1/2 & 0 & 0 & 0 \\ 0 & -1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & -1/2 & 0 \\ 0 & 0 & 0 & -1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Note that the invariants on

$$\mu = (\alpha, x, y, X, Y, \gamma)$$

are

$$P_1 \mu = 1/2 [2\alpha \quad x+y \quad x+y \quad X+Y \quad X+Y \quad 2\gamma],$$

and

$$P_2 \mu = 1/2 [0 \quad x-y \quad -x+y \quad X-Y \quad -X+Y \quad 0],$$

of dimensions 4 and 2 respectively. Correspondingly, we have,

$$\mu' \mathcal{P}_1 \mu = 1/2 [2\alpha^2 + (x + y)^2 + (X + Y)^2 + 2\gamma^2],$$

and

$$\mu' \mathcal{P}_2 \mu = 1/2 [(x - y)^2 + (X - Y)^2].$$

Suppose that $n = 3$ independent, identically distributed normal observations are obtained in each point of V , resulting in the data

$$f' = [3 \ 4 \ 3 \ 4 \ 5 \ 5 \ 6 \ 7 \ 3 \ 9 \ 7 \ 4 \ 10 \ 9 \ 9 \ 5 \ 5 \ 9] \in \mathbb{R}^{18},$$

respectively, as in V (the first three data points are from α , and so on). Following Proposition 2.9.2 with $e' = (1, 1, 1)$, we know that

$$(\mu \otimes e)' (\mathcal{P} \otimes T) (\mu \otimes e) = \begin{cases} n\mu' \mathcal{P} \mu & \text{if } T = \mathcal{A} \\ 0 & \text{if } T = \mathcal{Q} \end{cases}$$

is the expected value of

$$f' (\mathcal{P} \otimes T) f.$$

Here is the resulting decomposition after tensoring $\mathcal{P}_1 + \mathcal{P}_2$ with the standard reduction:

| | | | | |
|---|-------------------------------------|--------------------|--------------------------|--|
| | \mathcal{P} | $f' \mathcal{P} f$ | $\text{tr } \mathcal{P}$ | |
| [| $\mathcal{P}_1 \otimes \mathcal{A}$ | 687.66 | 4 | |
| | $\mathcal{P}_2 \otimes \mathcal{A}$ | 11.33 | 2 | |
| | $\mathcal{P}_1 \otimes \mathcal{Q}$ | 24.33 | 8 | |
| | $\mathcal{P}_2 \otimes \mathcal{Q}$ | 9.66 | 4 | |
|] | total | 733 | 18 | |

Combining the error terms $\mathcal{P}_1 \otimes \mathcal{Q} + \mathcal{P}_2 \otimes \mathcal{Q}$ we obtain the analysis of variance table

| | | | | |
|---|-------------------------------------|--------|----|--|
| [| source | ss | df | |
| | $\mathcal{P}_1 \otimes \mathcal{A}$ | 687.66 | 4 | |
| | $\mathcal{P}_2 \otimes \mathcal{A}$ | 11.33 | 2 | |
| | residual | 34 | 12 | |
|] | total | 733 | 18 | |

Under the hypothesis $H : x = y, X = Y$, we have

$$\mu' \mathcal{P}_2 \mu = 1/2 [(x - y)^2 + (X - Y)^2] = 0,$$

so that

$$E(f' (\mathcal{P}_2 \otimes \mathcal{A}) f) = E(f' (\mathcal{P}_2 \otimes \mathcal{Q}) f) = 0.$$

Therefore,

$$F = \frac{f' (\mathcal{P}_2 \otimes \mathcal{A}) f / \text{tr } (\mathcal{P}_2 \otimes \mathcal{A})}{f' (\mathcal{P}_2 \otimes \mathcal{Q}) f / \text{tr } (\mathcal{P}_2 \otimes \mathcal{Q})}$$

has a F distribution with degrees of freedom $df_1 = \text{tr } (\mathcal{P}_2 \otimes \mathcal{A})$ and $df_2 = \text{tr } (\mathcal{P}_2 \otimes \mathcal{Q})$, and can be used to assess the hypothesis. In the present example, $F = 1.99$ with degrees of freedom 2 and 12. □

Example 2.10.2. Consider the simple set product structure $V = L_1 \times L_2$ with $L_1 = \{1, 2\}$ and $L_2 = \{1, 2, 3\}$, which is the index set for a 2×3 data table such as

$$y = \begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{bmatrix}.$$

Let $S_2 \times S_3$ act on V according to $(\tau i, \sigma j)$, where $(\tau, \sigma) \in S_2 \times S_3$ and $(i, j) \in V$. The data are indicated by $y' = (u_1, u_2, u_3, v_1, v_2, v_3) \in \mathcal{V}$. The character tables of S_2 and S_3 are, respectively,

$$\left[\begin{array}{c|cc} \chi_2 \backslash \tau & 1 & t \\ \hline \chi_{21} & 1 & 1 \\ \chi_{22} & 1 & -1 \end{array} \right], \quad \left[\begin{array}{c|ccc} \chi_3 \backslash \tau & 1 & t & r \\ \hline \chi_{31} & 1 & 1 & 1 \\ \chi_{32} & 2 & 0 & -1 \\ \chi_{33} & 1 & -1 & 1 \end{array} \right],$$

in which 1 indicates the appropriate identity, t the corresponding (conjugacy class of) transpositions and r the (class of) order 3 cyclic permutations. Indicate by ρ and η the resulting permutation representations of S_2 and S_3 , respectively. From Proposition 2.9.1 we know that these reductions are exactly the corresponding standard reductions, indicated here by $\mathcal{A}_2, \mathcal{Q}_2$ and $\mathcal{A}_3, \mathcal{Q}_3$. The proposed reduction follows from Proposition 2.9.2, by tensoring. That is:

$$\mathcal{P}_1 = \mathcal{A}_2 \otimes \mathcal{A}_3 = \frac{1}{6} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix},$$

$$\mathcal{P}_2 = \mathcal{A}_2 \otimes \mathcal{Q}_3 = \frac{1}{6} \begin{bmatrix} 2 & -1 & -1 & 2 & -1 & -1 \\ -1 & 2 & -1 & -1 & 2 & -1 \\ -1 & -1 & 2 & -1 & -1 & 2 \\ 2 & -1 & -1 & 2 & -1 & -1 \\ -1 & 2 & -1 & -1 & 2 & -1 \\ -1 & -1 & 2 & -1 & -1 & 2 \end{bmatrix},$$

$$\mathcal{P}_3 = \mathcal{Q}_2 \otimes \mathcal{A}_3 = \frac{1}{6} \begin{bmatrix} 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 & 1 & 1 \end{bmatrix},$$

$$\mathcal{P}_4 = \mathcal{Q}_2 \otimes \mathcal{Q}_3 = \frac{1}{6} \begin{bmatrix} 2 & -1 & -1 & -2 & 1 & 1 \\ -1 & 2 & -1 & 1 & -2 & 1 \\ -1 & -1 & 2 & 1 & 1 & -2 \\ -2 & 1 & 1 & 2 & -1 & -1 \\ 1 & -2 & 1 & -1 & 2 & -1 \\ 1 & 1 & -2 & -1 & -1 & 2 \end{bmatrix}.$$

Table 2.1 shows the dimensions ($d = \text{tr } \mathcal{P}$) of the corresponding subspaces and indices for the respective bases. These indices carry the first-order interpretation of the data summarized in the subspaces generated by these bases. Suppose that $n = 3$ independent and identically distributed are obtained at each

TABLE 2.1. Canonical subspaces of $\rho \otimes \eta$, respective dimensions ($d = \text{tr } \mathcal{P}$) and corresponding bases for invariant subspaces.

| \mathcal{P} | d | basis | interpretation |
|-----------------|---|---|----------------------|
| \mathcal{P}_1 | 1 | $u_1 + u_2 + u_3 + v_1 + v_2 + v_3$ | baseline average |
| \mathcal{P}_2 | 2 | $2u_1 - u_2 - u_3 + 2v_1 - v_2 - v_3, -u_1 + 2u_2 - u_3 - v_1 + 2v_2 - v_3$ | column effect |
| \mathcal{P}_3 | 1 | $u_1 + u_2 + u_3 - v_1 - v_2 - v_3$ | row effect |
| \mathcal{P}_4 | 2 | $2u_1 - u_2 - u_3 - 2v_1 + v_2 + v_3, -u_1 + 2u_2 - u_3 + v_1 - 2v_2 + v_3$ | remainder ϵ |

point of the initial structure. The new underlying structure is then $V = L_1 \times L_2 \times L_3$ with $L_1 = \{1, 2\}$, $L_2 = \{1, 2, 3\}$ and $L_3 = \{1, \dots, n\}$. The data space \mathcal{V} has dimension $\ell_1 \times \ell_2 \times n$. The unsung reduction is now obtained by an additional tensoring with the standard reduction $\mathcal{A}_2, \mathcal{Q}_3$. The data are written as $y' = (u_1, u_2, u_3, v_1, v_2, v_3) \in \mathcal{V}$, with the understanding that each entry is a vector in \mathbb{R}^3 . From Proposition 2.9.2 we know that

$$I = \mathcal{P}_1 \otimes \mathcal{A} + \dots \mathcal{P}_4 \otimes \mathcal{A} + \mathcal{P}_1 \otimes \mathcal{Q} + \dots \mathcal{P}_4 \otimes \mathcal{Q}$$

is a canonical reduction. Here is one numerical example, with the resulting decomposition followed by the standard analysis of variance:

| | | | |
|---|----------------|---------|-------|
| $y = \left[\begin{array}{c c c} 3.1, 3.5, 3.2 & 4.5, 4, 4.7 & 6.7, 4.5, 6.8 \\ \hline 4.1, 4, 4.4 & 2.3, 1.9, 1.5 & 7.9, 7.7, 8 \end{array} \right] :$ | V | ss | dim V |
| | 1 | 380.870 | 1 |
| | c | 49.958 | 2 |
| | r | 0.035 | 1 |
| | cr | 15.779 | 2 |
| | e ₁ | 0.086 | 2 |
| | e ₂ | 0.260 | 2 |
| | e ₃ | 3.380 | 2 |
| | e ₄ | 0.086 | 2 |
| | e ₅ | 0.319 | 2 |
| e ₆ | 0.045 | 2 | |
| total | 450.840 | 18 | |

| Source | Sum-of-Squares | df | Mean-Square | F-ratio |
|---------|----------------|----|-------------|---------|
| ROW | 0.036 | 1 | 0.036 | 0.102 |
| COL | 49.963 | 2 | 24.982 | 71.718 |
| ROW*COL | 15.781 | 2 | 7.891 | 22.652 |
| Error | 4.180 | 12 | 0.348 | |

□

Example 2.10.3 (Reducing the standard 2^p factorial data). The canonical reduction for the 2^p factorial data is simply the p -fold tensor of the standard reduction in S_2 , that is,

$$I_{2^p} = \underbrace{(\mathcal{A} + \mathcal{Q}) \otimes \dots \otimes (\mathcal{A} + \mathcal{Q})}_{p \text{ times}}$$

For $p = 2$, denoting

$$u = \mathcal{A} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad t = \mathcal{Q} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix},$$

we obtain the canonical reduction

$$\begin{aligned} uu = u \otimes u &= \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, & ut = u \otimes t &= \frac{1}{4} \begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix}, \\ tu = t \otimes u &= \frac{1}{4} \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}, & tt = t \otimes t &= \frac{1}{4} \begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}. \end{aligned}$$

These projections act on observations indexed by the high (1)-low (0) labels (00, 10, 01, 11) in V . To illustrate, consider the case in which $n = 2$ observations are obtained at each of the 8 labels of a 2^3 factorial experiment, that is,

$$V = \begin{bmatrix} a & b & c \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \quad y = \begin{bmatrix} 15 & 779 \\ 999 & 990 \\ 499 & 212 \\ 286 & 611 \\ 438 & 239 \\ 926 & 787 \\ 871 & 303 \\ 891 & 663 \end{bmatrix} \in \mathcal{V}.$$

The data reduce according to

$$I = \text{uuu} \otimes \mathcal{A} + \dots \text{ttt} \otimes \mathcal{A} + \text{uuu} \otimes \mathcal{Q} + \dots \text{ttt} \otimes \mathcal{Q},$$

leading to the decomposition

| trait | \mathcal{P} | $y'(\mathcal{P} \otimes \mathcal{A})y$ | $\text{tr } \mathcal{P} \otimes \mathcal{A}$ |
|-------|---------------|--|--|
| 1 | uuu | 5651317.56 | 1 |
| c | uut | 488950.56 | 1 |
| b | utu | 43785.56 | 1 |
| bc | utt | 173264.06 | 1 |
| a | tuu | 33033.06 | 1 |
| ac | tut | 76.56 | 1 |
| ab | ttu | 143073.06 | 1 |
| abc | ttt | 7788.06 | 1 |
| | | $y'(\mathcal{P} \otimes \mathcal{Q})y$ | $\text{tr } \mathcal{P} \otimes \mathcal{Q}$ |
| e | uuu | 7267.56 | 1 |
| e | uut | 3570.06 | 1 |
| e | utu | 86289.06 | 1 |
| e | utt | 173264.06 | 1 |
| e | tuu | 232083.06 | 1 |
| e | tut | 19670.06 | 1 |
| e | ttu | 4192.56 | 1 |
| e | ttt | 76314.06 | 1 |
| total | I | 7143939.00 | 16 |

All subspaces are one-dimensional so that the decomposition is irreducible and the analysis is complete. Here is the condensed standard analysis:

Analysis of Variance

| Source | Sum-of-Squares | df | Mean-Square | F-ratio |
|--------|----------------|----|-------------|---------|
| a | 33033.063 | 1 | 33033.063 | 0.439 |
| b | 43785.563 | 1 | 43785.563 | 0.581 |
| c | 488950.563 | 1 | 488950.563 | 6.491 |
| a*b | 143073.062 | 1 | 143073.062 | 1.899 |
| a*c | 76.563 | 1 | 76.563 | 0.001 |
| b*c | 173264.062 | 1 | 173264.062 | 2.300 |
| a*b*c | 7788.063 | 1 | 7788.063 | 0.103 |
| Error | 602650.500 | 8 | 75331.313 | |

□

Example 2.10.4 (Fractional factorial experiments). The projections defining a half fraction (2^{3-1}) of the 2^3 experiment described in Example 2.10.3 can be obtained as

$$\mathcal{P}_1 = (\mathbf{u} + \mathbf{t}) \otimes \mathbf{u} \otimes \mathbf{u} = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix},$$

$$\mathcal{P}_2 = (\mathbf{u} + \mathbf{t}) \otimes \mathbf{u} \otimes \mathbf{t} = \frac{1}{4} \begin{bmatrix} 1 & -1 & 1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 1 & -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 1 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 1 & -1 & 1 \end{bmatrix},$$

$$\mathcal{P}_3 = (\mathbf{u} + \mathbf{t}) \otimes \mathbf{t} \otimes \mathbf{t} = \frac{1}{4} \begin{bmatrix} 1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 & 1 \end{bmatrix},$$

$$\mathcal{P}_4 = (\mathbf{u} + \mathbf{t}) \otimes \mathbf{t} \otimes \mathbf{u} = \frac{1}{4} \begin{bmatrix} 1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ -1 & -1 & 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & -1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & -1 & -1 & 1 & 1 \end{bmatrix},$$

all in dimension of 2. Each projection generates two sets of one-dimensional orthogonal invariants, obtained by symbolically multiplying the projection matrices by the labels

$$\mathbf{v}' = (000, 100, 010, 110, 001, 101, 011, 111).$$

Collecting one from each projection, we obtain the invariants

$$L_1 = 000 + 100 + 010 + 110,$$

$$L_2 = 000 - 100 + 010 - 110,$$

$$L_3 = 000 - 100 - 010 + 110,$$

$$L_4 = 000 + 100 - 010 - 110,$$

thus showing that only the half fraction of the original 8 labels in V are needed in the fractional experiment.

More generally, the fractional experiments for the 2^4 factorial experiment are obtained as the solutions to the equations

$$(2.42) \quad \mathcal{P}_1 + \mathcal{P}_2 = \mathbf{I}, \quad 2^{-3} \text{ fraction},$$

$$(2.43) \quad \mathcal{P}_1 + \mathcal{P}_2 + \mathcal{P}_3 + \mathcal{P}_4 = \mathbf{I}, \quad 2^{-2} \text{ fraction},$$

$$(2.44) \quad \mathcal{P}_1 + \mathcal{P}_2 + \mathcal{P}_3 + \mathcal{P}_4 + \mathcal{P}_5 + \mathcal{P}_6 + \mathcal{P}_7 + \mathcal{P}_8 = \mathbf{I}, \quad 2^{-1} \text{ fraction},$$

where the unknown are pairwise disjoint projection matrices in the original 2^4 -dimensional space. Equation (2.42) has three sets of non-isomorphic solutions, each one corresponding to a 2^{4-3} fractional experiment, namely:

$$\begin{aligned} \mathcal{P}_{3,11} &= \mathbf{u} \otimes (\mathbf{u} + \mathbf{t}) \otimes \mathbf{1} \otimes \mathbf{1}, & \mathcal{P}_{3,12} &= \mathbf{t} \otimes (\mathbf{u} + \mathbf{t}) \otimes \mathbf{1} \otimes \mathbf{1}, \\ \mathcal{P}_{3,21} &= \mathbf{u} \otimes (\mathbf{u} + \mathbf{t}) \otimes (\mathbf{u} + \mathbf{t}) \otimes \mathbf{1}, & \mathcal{P}_{3,22} &= \mathbf{t} \otimes (\mathbf{u} + \mathbf{t}) \otimes (\mathbf{u} + \mathbf{t}) \otimes \mathbf{1}, \\ \mathcal{P}_{3,31} &= \mathbf{u} \otimes (\mathbf{u} + \mathbf{t}) \otimes (\mathbf{u} + \mathbf{t}) \otimes (\mathbf{u} + \mathbf{t}), & \mathcal{P}_{3,32} &= \mathbf{t} \otimes (\mathbf{u} + \mathbf{t}) \otimes (\mathbf{u} + \mathbf{t}) \otimes (\mathbf{u} + \mathbf{t}). \end{aligned}$$

Equation (2.43) leads to two sets of non-isomorphic 2^{4-2} fractional experiments, given by

$$\begin{aligned} \mathcal{P}_{2,11} &= \mathbf{u} \otimes \mathbf{u} \otimes (\mathbf{u} + \mathbf{t}) \otimes (\mathbf{u} + \mathbf{t}), & \mathcal{P}_{2,12} &= \mathbf{u} \otimes \mathbf{t} \otimes (\mathbf{u} + \mathbf{t}) \otimes (\mathbf{u} + \mathbf{t}), \\ \mathcal{P}_{2,13} &= \mathbf{t} \otimes \mathbf{u} \otimes (\mathbf{u} + \mathbf{t}) \otimes (\mathbf{u} + \mathbf{t}), & \mathcal{P}_{2,14} &= \mathbf{t} \otimes \mathbf{t} \otimes (\mathbf{u} + \mathbf{t}) \otimes (\mathbf{u} + \mathbf{t}), \\ \mathcal{P}_{2,21} &= \mathbf{u} \otimes \mathbf{u} \otimes (\mathbf{u} + \mathbf{t}) \otimes \mathbf{1}, & \mathcal{P}_{2,22} &= \mathbf{u} \otimes \mathbf{t} \otimes (\mathbf{u} + \mathbf{t}) \otimes \mathbf{1}, \\ \mathcal{P}_{2,23} &= \mathbf{t} \otimes \mathbf{u} \otimes (\mathbf{u} + \mathbf{t}) \otimes \mathbf{1}, & \mathcal{P}_{2,24} &= \mathbf{t} \otimes \mathbf{t} \otimes (\mathbf{u} + \mathbf{t}) \otimes \mathbf{1}. \end{aligned}$$

Equation (2.44) has one set of solutions, defining the 2^{4-1} fractional experiment, given by

$$\begin{aligned} \mathcal{P}_{1,1} &= \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{u} \otimes (\mathbf{u} + \mathbf{t}), & \mathcal{P}_{1,2} &= \mathbf{t} \otimes \mathbf{u} \otimes \mathbf{u} \otimes (\mathbf{u} + \mathbf{t}), \\ \mathcal{P}_{1,3} &= \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{t} \otimes (\mathbf{u} + \mathbf{t}), & \mathcal{P}_{1,4} &= \mathbf{t} \otimes \mathbf{u} \otimes \mathbf{t} \otimes (\mathbf{u} + \mathbf{t}), \\ \mathcal{P}_{1,5} &= \mathbf{u} \otimes \mathbf{t} \otimes \mathbf{u} \otimes (\mathbf{u} + \mathbf{t}), & \mathcal{P}_{1,6} &= \mathbf{t} \otimes \mathbf{t} \otimes \mathbf{u} \otimes (\mathbf{u} + \mathbf{t}), \\ \mathcal{P}_{1,7} &= \mathbf{u} \otimes \mathbf{t} \otimes \mathbf{t} \otimes (\mathbf{u} + \mathbf{t}), & \mathcal{P}_{1,8} &= \mathbf{t} \otimes \mathbf{t} \otimes \mathbf{t} \otimes (\mathbf{u} + \mathbf{t}). \end{aligned}$$

□

Example 2.10.5 (Cyclic symmetries). In this example we consider the set product space $V = C \times L$ subject to permutation action of C_c and C_ℓ on C and L , respectively. There are $c\ell$ one-dimensional irreducible representations of $C_c \times C_\ell$ with projection matrices given by

$$(2.45) \quad \mathcal{P}_{mn} = \frac{1}{c\ell} \sum_{i,j} \omega_c^{mi} \omega_\ell^{nj} (\rho_c^i \otimes \rho_\ell^j), \quad n = 1, \dots, c, \quad m = 1, \dots, \ell,$$

where $\omega_f = e^{2\pi i/f}$ and ρ_f is the permutation representation of the generating cyclic permutation (12...f). As commented earlier on in the chapter, it is important to distinguish the field of scalars defining the vector space \mathcal{V} , where these projections operate on. Consider, to illustrate this point, the case of C_3 . The resulting canonical projections are given by

$$\mathcal{P}_1 = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad \mathcal{P}_2 = \frac{1}{3} \begin{bmatrix} 1 & \omega & \omega^2 \\ \omega^2 & 1 & \omega \\ \omega & \omega^2 & 1 \end{bmatrix}, \quad \mathcal{P}_3 = \frac{1}{3} \begin{bmatrix} 1 & \omega^2 & \omega \\ \omega & 1 & \omega^2 \\ \omega^2 & \omega & 1 \end{bmatrix},$$

where $\omega = \omega_3 = e^{2\pi i/3}$. In general, these matrices are in $\text{GL}(\mathbb{C}^3)$, so that the resulting linear operations upon the vectors in \mathcal{V} then require that \mathcal{V} be regarded as a complex vector space. In the real vector space case the irreducible decomposition is $\mathbf{I} = \mathcal{T}_1 + \mathcal{T}_2$, with $\mathcal{T}_1 = \mathcal{P}_1$ of $\dim = 1$ and

$$\mathcal{T}_2 = \mathcal{P}_2 + \mathcal{P}_3 = \frac{1}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

of $\dim = 2$. Note that $\mathcal{P}_2 = \mathcal{P}'_3$, so that $y'\mathcal{P}_2y = (y'\mathcal{P}_2y)' = y'\mathcal{P}_3y$, leading to the reduction

$$y'y = y'\mathcal{T}_1y + y'\mathcal{T}_2y = y'\mathcal{P}_1y + 2y'\mathcal{P}_2y.$$

Similarly, with C_4 , the reduction over the real field is $I = \mathcal{T}_1 + \mathcal{T}_2 + \mathcal{T}_3$, with

$$\mathcal{T}_1 = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \quad \mathcal{T}_2 = \frac{1}{4} \begin{bmatrix} 1 & \omega^2 & 1 & \omega^2 \\ \omega^2 & 1 & \omega^2 & 1 \\ 1 & \omega^2 & 1 & \omega^2 \\ \omega^2 & 1 & \omega^2 & 1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix},$$

both of dimension one, and

$$\begin{aligned} \mathcal{T}_3 = \mathcal{P}_3 + \mathcal{P}_4 &= \frac{1}{4} \left(\begin{bmatrix} 1 & \omega & \omega^2 & \omega^3 \\ \omega^3 & 1 & \omega & \omega^2 \\ \omega^2 & \omega^3 & 1 & \omega \\ \omega & \omega^2 & \omega^3 & 1 \end{bmatrix} + \begin{bmatrix} 1 & \omega^3 & \omega^2 & \omega \\ \omega & 1 & \omega^3 & \omega^2 \\ \omega^2 & \omega & 1 & \omega^3 \\ \omega^3 & \omega^2 & \omega & 1 \end{bmatrix} \right) \\ &= \frac{1}{4} \begin{bmatrix} 2 & 0 & -2 & 0 \\ 0 & 2 & 0 & -2 \\ -2 & 0 & 2 & 0 \\ 0 & -2 & 0 & 2 \end{bmatrix}, \end{aligned}$$

of dimension two. Here, $\mathcal{P}_3 = \mathcal{P}'_4$, and the sum of squares reduces as

$$y'y = y'\mathcal{T}_1y + y'\mathcal{T}_2y + y'\mathcal{T}_3y = y'\mathcal{T}_1y + y'\mathcal{T}_2y + 2y'\mathcal{P}_3y.$$

We conclude this example with the evaluation of the product action of $C_4 \times C_7$ on the data set shown in Table 2.2. These data are discussed in Wit and McCullagh (2001). To decompose these

TABLE 2.2. Frequency of mining disasters between 1851-1962.

| | Mon | Tue | Wed | Thu | Fri | Sat | Sun | <i>total</i> |
|--------------|-----|-----|-----|-----|-----|-----|-----|--------------|
| Autumn | 7 | 10 | 5 | 5 | 6 | 7 | 1 | 41 |
| Winter | 5 | 9 | 10 | 10 | 11 | 7 | 0 | 52 |
| Spring | 3 | 7 | 10 | 12 | 13 | 9 | 2 | 56 |
| Summer | 4 | 8 | 8 | 9 | 5 | 6 | 2 | 42 |
| <i>total</i> | 19 | 34 | 33 | 36 | 35 | 29 | 5 | 191 |

data, we apply (2.45) with $\ell = 4$ and $c = 7$. There are 28 one-dimensional projections defined in \mathcal{V} regarded as a complex vectors space, decomposing the total sum of squares, 1607. The results are summarized in the following matrix in which the (m,n) entry corresponds to the projection \mathcal{P}_{mn} .

$$[\mathcal{P}_{mn}] = \begin{bmatrix} & \begin{array}{c|cccccc|c} & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline 1 & 13.378 & 3.325 & 2.419 & 0.539 & 2.218 & 16.762 & 11.607 \\ 2 & 2.784 & 1.165 & 0.014 & 0.014 & 1.165 & 2.784 & 0.321 \\ 3 & 16.762 & 2.218 & 0.539 & 2.419 & 3.325 & 13.378 & 11.607 \\ 4 & 68.450 & 23.760 & 5.510 & 5.510 & 23.760 & 68.450 & 1302.900 \end{array} \\ \hline \end{bmatrix}.$$

The corresponding 15 components of the irreducible (in \mathbb{R}) reduction of the original ($x'x$) sum of squares and the transformed ($u'u$) sum of squares based on the multinomial vector $u' =$

$\sqrt{v}(x - np)/\sqrt{n}$ are:

| | m, n | m', n' | $x'(\mathcal{P}_{mn} + \mathcal{P}_{m'n'})x$ | dim | $u'(\mathcal{P}_{mn} + \mathcal{P}_{m'n'})u$ |
|------------------|------|--------|--|---------|--|
| weekly-quarterly | 1, 1 | 3, 6 | 26.756 | 2 | 3.922 |
| | 1, 2 | 3, 5 | 6.650 | 2 | 0.975 |
| | 1, 3 | 3, 4 | 4.838 | 2 | 0.709 |
| | 1, 4 | 3, 3 | 1.078 | 2 | 0.157 |
| | 1, 5 | 3, 2 | 4.436 | 2 | 0.650 |
| | 1, 6 | 3, 1 | 33.524 | 2 | 4.914 |
| | 2, 1 | 2, 6 | 5.568 | 2 | 0.816 |
| | 2, 2 | 2, 5 | 2.330 | 2 | 0.341 |
| | 2, 3 | 2, 4 | 4.838 | 2 | 0.004 |
| weekly | 4, 1 | 4, 6 | 136.900 | 2 | 20.064 |
| | 4, 2 | 4, 5 | 47.520 | 2 | 6.963 |
| | 4, 3 | 4, 4 | 11.020 | 2 | 1.612 |
| quarterly | 1, 7 | 3, 7 | 23.214 | 2 | 3.403 |
| | 2, 7 | | 0.321 | 1 | 0.047 |
| | 4, 7 | | 1302.900 | 1 | (0) |
| total | | | $x'x = 1607.000$ | 28 (27) | $u'u = 44.584$ |

It is not difficult to conclude that the probability law of $u'(\mathcal{P}_{mn} + \mathcal{P}_{m'n'})u$ is approximately χ_2^2 , so that the reduction identifies

- (1) the *DC component* $m = 4, n = 7$: note that under the multinomial transformation, as expected, this component is zero;
- (2) the weekly cycle $n = 1$, with $u'(\mathcal{P}_{mn} + \mathcal{P}_{m'n'})u = 20.064$ and the weekly cycle $n = 2$, with $u'(\mathcal{P}_{mn} + \mathcal{P}_{m'n'})u = 6.963$, corresponding to angular phases $\theta = 2\pi/7$ and $\theta = 4\pi/7$, respectively;
- (3) a suppressed, not significant, quarterly cycle $m=1$, with $u'(\mathcal{P}_{mn} + \mathcal{P}_{m'n'})u = 3.403$;
- (4) a significant quarterly-weekly cycle $m = 1, n = 6$ (equivalently $m = 1, n = 1$), with $u'(\mathcal{P}_{mn} + \mathcal{P}_{m'n'})u = 4.9146$ and angular phase $\theta = 2\pi/28$.

□

Example 2.10.6 (Latin squares). The reduction of a Latin square experiment has the form

$$I = \underbrace{[(\mathcal{A} + \mathcal{Q}) \otimes (\mathcal{A} + \mathcal{Q})]}_{\text{rows}} \underbrace{(\mathcal{A} + \mathcal{Q})}_{\text{columns}} \underbrace{(\mathcal{A} + \mathcal{Q})}_{\text{letters}}.$$

To illustrate, consider the following experiment described in Youden (1951, p.96), in which the data

$$\left[\begin{array}{c|cccc} & \text{I} & \text{II} & \text{III} & \text{IV} \\ \hline 1 & \text{A} & \text{B} & \text{C} & \text{D} \\ 2 & \text{C} & \text{D} & \text{A} & \text{B} \\ 3 & \text{B} & \text{C} & \text{D} & \text{A} \\ 4 & \text{D} & \text{A} & \text{B} & \text{C} \end{array} \right] \rightarrow \begin{bmatrix} 36 & 38 & 36 & 30 \\ 17 & 18 & 26 & 17 \\ 30 & 39 & 41 & 34 \\ 30 & 45 & 38 & 33 \end{bmatrix}$$

are the melting point temperature readings of 4 chemical cells (1, 2, 3, 4) obtained from 4 thermometers (I, II, III, IV) in 4 different days (A, B, C, D). The numerical entries are the readings converted to degrees Centigrade. Only the fourth decimal places are given, as the readings agreed up to the last two places. The experimental background is such that there is no reason to assume an interaction between cells and thermometers. Write the data as

$$y' = (36, 38, 36, 30, 17, 18, 26, 17, 30, 39, 41, 34, 30, 45, 38, 33)$$

and first evaluate the four projections associated with $(\mathcal{A} + \mathcal{Q}) \otimes (\mathcal{A} + \mathcal{Q})$, where $I_4 = \mathcal{A} + \mathcal{Q}$ is the standard reduction in dimension 4. We obtain

$$\begin{aligned} \text{SS total} &= y'y = 17230, \\ \text{SS constant} &= y'(\mathcal{A} \otimes \mathcal{A})y = 16129, \\ \text{SS thermometers} &= y'(\mathcal{A} \otimes \mathcal{Q})y = 182.50, \\ \text{SS cells} &= y'(\mathcal{Q} \otimes \mathcal{A})y = 805, \\ \text{SS days} + \text{SS residual} &= y'(\mathcal{Q} \otimes \mathcal{Q})y = 113.5 \end{aligned}$$

To further reduce $y'(\mathcal{Q} \otimes \mathcal{Q})y$ and determine the component due to eventual day-to-day variability, we apply the standard reduction (indicated here by \mathcal{A}_\bullet and \mathcal{Q}_\bullet) to aggregate the data from corresponding days. That is,

$$\mathcal{A}_\bullet = 1/4 \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

and $\mathcal{Q}_\bullet = I_{16} - \mathcal{A}_\bullet$. We obtain

$$\begin{aligned} \text{SS days} &= y'([\mathcal{Q} \otimes \mathcal{Q}]\mathcal{A}_\bullet)y = 70, \\ \text{SS residual} &= y'([\mathcal{Q} \otimes \mathcal{Q}]\mathcal{Q}_\bullet)y = 43.5. \end{aligned}$$

The following table summarizes the results:

| \mathcal{P} | $ss = y'\mathcal{P}y$ | $df = \text{tr } \mathcal{P}$ | $mss = ss/df$ |
|---|-----------------------|-------------------------------|---------------|
| $\mathcal{A} \otimes \mathcal{A}$ (constant) | 16129.00 | 1 | 16129.00 |
| $\mathcal{A} \otimes \mathcal{Q}$ (thermometers) | 182.50 | 3 | 60.83 |
| $\mathcal{Q} \otimes \mathcal{A}$ (cells) | 805.00 | 3 | 268.33 |
| $[\mathcal{Q} \otimes \mathcal{Q}]\mathcal{A}_\bullet$ (days) | 70.00 | 3 | 23.33 |
| $[\mathcal{Q} \otimes \mathcal{Q}]\mathcal{Q}_\bullet$ (residual) | 43.50 | 6 | 7.25 |
| I (total) | 17230.00 | 16 | 1076.90 |

The F ratio for cells is quite significant ($268.33/7.25 = 37.01$). The F ratio for thermometers, $60.83/7.25 = 8.39$, also points to a difference among the thermometers whereas the F ratio for days is only suggestive of a day-to-day effect. The estimated standard deviation for a single measurement is $\sqrt{7.25} = 2.69$, which shows an improvement in the error of comparison. In fact, if the effect of days on the readings is not eliminated, the standard deviation would then be $\sqrt{113.5/9} = 3.55$. \square

Example 2.10.7 (One-way analysis of variance). The canonical reduction for the one-way ANOVA is simply

$$I_n = \mathcal{A} + \mathcal{Q}(D_{\mathcal{A}} + D_{\mathcal{Q}}),$$

where $\mathcal{A} + \mathcal{Q}$ is the standard reduction in dimension $n = n_1 + \dots + n_k$,

$$D_{\mathcal{A}} = \text{diag}(\mathcal{A}_{n_1}, \dots, \mathcal{A}_{n_k}),$$

and $D_Q = I - D_A$. Here is one illustration with $n_1 = n_2 = 3$ and $n_3 = 5$, so that

$$D_A = \text{diag} (\mathcal{A}_3, \mathcal{A}_3, \mathcal{A}_5).$$

The reader may verify that indeed

$$I_{11} = \mathcal{A}_{11} + Q_{11}D_A + Q_{11}D_Q \equiv \mathcal{P}_1 + \mathcal{P}_2 + \mathcal{P}_3$$

is a canonical reduction.

Given the data

$$y' = \underbrace{(12, 14, 11)}_{\text{group 1}}, \underbrace{(10, 9, 11, 8, 12, 15, 14, 12)}_{\text{group 2}}, \underbrace{(\quad)}_{\text{group 3}},$$

the resulting decomposition

| | | | |
|-----------------------------|------------------|--------------------------|---|
| \mathcal{P} | $y'\mathcal{P}y$ | $\text{tr } \mathcal{P}$ | $y'\mathcal{P}y/\text{tr } \mathcal{P}$ |
| \mathcal{P}_2 (treatment) | 11.07 | 2 | 5.53 |
| \mathcal{P}_3 (residual) | 35.46 | 8 | 4.43 |
| \mathcal{P}_1 (constant) | 1489.45 | 1 | 1489.45 |
| I (total) | 1536.0 | 11 | 139.63 |

correspond to the standard ANOVA table

Analysis of Variance

| Source | Sum-of-Squares | df | Mean-Square | F-ratio |
|-----------|----------------|----|-------------|---------|
| treatment | 11.079 | 2 | 5.539 | 1.249 |
| Error | 35.467 | 8 | 4.433 | |

□

Example 2.10.8 (Two-way analysis of variance). The reduction of a two-way ANOVA with r row levels, c columns levels, and n observation in each cell is given by

$$I_{rcn} = (\mathcal{A}_r + Q_r) \otimes (\mathcal{A}_c + Q_c) \otimes (\mathcal{A}_n + Q_n).$$

The resulting 8 projections define

- (1) $\mathcal{P}_1 = \mathcal{A}_r \otimes \mathcal{A}_c \otimes \mathcal{A}_n$, the constant;
- (2) $\mathcal{P}_r = Q_r \otimes \mathcal{A}_c \otimes \mathcal{A}_n$, the row effect;
- (3) $\mathcal{P}_c = \mathcal{A}_r \otimes Q_c \otimes \mathcal{A}_n$, the column effect;
- (4) $\mathcal{P}_{rc} = Q_r \otimes Q_c \otimes \mathcal{A}_n$, the interaction term;
- (5) $\mathcal{P}_e = \mathcal{A}_r \otimes \mathcal{A}_c \otimes Q_n + Q_r \otimes \mathcal{A}_c \otimes Q_n + \mathcal{A}_r \otimes Q_c \otimes Q_n + Q_r \otimes Q_c \otimes Q_n$, the error term.

□

Example 2.10.9 (Linear regression). The reduction of the linear regression structure is obtained from

$$I = \mathcal{A} + QP + QP^\perp \equiv \mathcal{P}_1 + \mathcal{P}_2 + \mathcal{P}_3,$$

where $P = X(X'X)^{-1}X'$, $P^\perp = I - P$, and X is the usual $n \times p$ design matrix. The reader may verify that $I = \mathcal{P}_1 + \mathcal{P}_2 + \mathcal{P}_3$ is in fact a canonical reduction (use Proposition 2.9.2 and the fact that P is a symmetric matrix). The data reduce according to

$$y'y = \underbrace{y'\mathcal{P}_1y}_{\text{constant}} + \underbrace{y'\mathcal{P}_2y}_{\text{regression}} + \underbrace{y'\mathcal{P}_3y}_{\text{residual}}.$$

□

2.11. Summary

In this chapter we introduced the elements of algebra for the analysis of structured data. Structured data are data that are indexed by a set V of indices or labels upon which certain symmetry relations can be defined. Some of the basic structures are the set of all mappings $V = C^L$, the set product $V = L \times C$ and the set $V = L \times \Omega$, where $\Omega = \{\omega; \omega^c = 1\}$. Corneal surface curvature data and Shack-Hartmann wave-front sensor data are typically indexed by a structure V of the type $L \times \Omega$. The points or labels in $L \times \Omega$ are at the intersection of ℓ concentric rings and c equally spaced semi meridians, and provide the index for a surface curvature or for a point spread function value. In many applications, the data are indexed by a group (G) of symmetries. The refractive group described in Campbell (1997) is an example of a set of labels or indices for refractive data. See, for example, Lakshminarayanan and Viana (2005), in addition to Viana (2003a), Viana (2004) and Viana (2003b).

We have illustrated the notion that summarizing and analyzing the structured data $(x(s))_{s \in V}$ can be facilitated by the symmetries in the data space (\mathcal{V}) that are induced by the symmetries in the underlying structure. These symmetries when applied to the labels in V according to a group action φ reduce the structure V into disjoint similarity orbits

$$\mathcal{O}_1 \cup \dots \cup \mathcal{O}_m.$$

The group action leads to a linear group representation in the data space which associates to each $\sigma \in G$ the matrix $\rho(\sigma)$ changing the canonical basis $\{e_s; s \in V\}$ of \mathcal{V} indexed by V into the basis $\{e_{\varphi(\sigma,s)}; s \in V\}$. The resulting factorization $\mathcal{V} = \mathcal{V}_1 \oplus \dots \oplus \mathcal{V}_h$ in the data vector space is the consequence of defining a set of orthogonal projections $\mathcal{P}_1, \dots, \mathcal{P}_h$, each one of these a linear combination of the matrices $\rho(\sigma)$ over G , with scalar coefficients the characters of the corresponding irreducible representation of G . If there are h irreducible characters then the identity operator I in \mathbb{R}^V reduces as

$$I = \mathcal{P}_1 + \mathcal{P}_2 + \dots + \mathcal{P}_h$$

and the operators satisfy the properties $\mathcal{P}_i \mathcal{P}_j = 0$ for $i \neq j$ and $\mathcal{P}_i^2 = \mathcal{P}_i$, $i = 1, \dots, h$. It then follows that the basic decomposition

$$\|x\|^2 = (x|x) = (x|\mathcal{P}_1 x) + (x|\mathcal{P}_2 x) + \dots + (x|\mathcal{P}_h x)$$

for the sum of squares for a particular inner product $(\cdot|\cdot)$ of interest (e.g., Euclidean, Hermitian, symplectic) can be obtained.

The canonical decomposition, we remark, establishes the formal, unifying, connection between the symmetries in the structured data and statistical inference. For example, the statistical (Fisher-Cochran) theory of quadratic forms can be applied to obtain new forms of analysis of variance, within which symmetry-related hypotheses can be defined and interpreted.

In particular, when the data are indexed by a group of symmetries (the case $V = G$ mentioned above), varied forms of spectral analysis for the structured data are then obtained. We observed that a data set

$$\{x(\tau); \tau \in G\}$$

indexed by a finite group G can be identified with the elements $\sum_{\tau \in G} x_s(\tau)\tau$ of the group algebra A_G associated with G . There are several experimental conditions in which data $\{x(\tau); \tau \in G\}$ are naturally indexed by group symmetries: For example, the symmetry perception clinical studies described by Szlyk, Seiple and Xie (1995) and Szlyk, Rock and Fisher (1995) in which the data are naturally indexed by rotational and axial symmetries (dihedral experimental designs). These dihedral designs are potentially useful to describe and suggest interpretations to the (rotational, axial) symmetries present in human visual field data or in two-dimensional wave-front aberration data from the Shack-Hartmann wave-front sensor, e.g., Salmon, Thibos and Bradley (1998). Similarly, these symmetries are visibly present in the maxilla-mandible axial and rotational symmetries in data indexed by the points in the dental arch system, often used in anthropological science, orthodontics and oral biology morphologic studies, e.g., Lestrel, Takahashi and Kanazawa (2004), Oliveira, Silveira, Kusnoto and Viana (2004).

The group algebra interpretation given to data indexed by a group of symmetries suggests a different mechanism to indexing the data, namely as the inverse solution to a Fourier transform

$$F = \hat{x}(\beta) = \sum_{\tau \in G} x(\tau)\beta(\tau),$$

at the (irreducible) representation β of G . In applications, we start with a given optical linear operator F and evaluations of the Fourier transform $\hat{x}(\rho)$ at the remaining irreducible representations ρ to pass the data to the group. Once this is obtained we give to G the interpretation of a set structure as indicated above, followed with the canonical decomposition of the structured data.

Further reading

- (1) The presentation of the material in this chapter closely follows the program of Serre (1977). All the basic results of functions on groups can be seen in the classic text Naimark and Štern (1982). The basic facts about projections and vector spaces are found in Halmos (1987)'s classic text. See also Rotman (1995) on general facts about the theory of groups, and Simon (1996) for a more contemporary text on representations of finite and compact groups. The reader will enjoy reading the historical account, by Lam (1998), of representations of finite groups in the past century;
- (2) On permutation groups, Cameron (1999) or Dixon and Mortimer (1996);
- (3) On combinatorics, Cameron (1994) or Stanton and White (1986);
- (4) Matrix groups, e.g., Curtis (1984);

- (5) Random matrices, functions, and permutations, Diaconis and Shahshahani (1986), Tuljapurkar (1986), Kolchin (1971), Diaconis and Freedman (1999), Arratia and Simon (92), Lalley (96) ;

Exercises

Exercise 2.1. Signed matrices. Show that the set of all $n \times n$ matrices

$$\{\text{Diag}(\pm 1, \dots, \pm 1)\},$$

with the operation of matrix multiplication defines a finite group of order 2^n . It can be identified with the mapping space of all two-sequences in length of n .

Exercise 2.2. Show that

$$\rho_i = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad \rho_{-1} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \rho_{-i} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \rho_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

is two-dimensional representation of the complex group.

Exercise 2.3. Show that, for any member τ of group G , the mapping $\tau^* : G \rightarrow G$ by $\tau^*(\sigma) = \tau\sigma\tau^{-1}$ is an isomorphism in G , and the mapping $\tau \mapsto \tau^*$ is a homomorphism of G , taking values in the set $\text{Aut}(G)$ of isomorphisms in G .

Exercise 2.4. For each of the 2×2 matrix groups, with elements indicated by X , introduced in Example 2.2.3, evaluate the *infinitesimal group element* $\Omega(X) = X^{-1}dX$ and show that Ω is invariant under the group of transformations $X \rightarrow AX$, for all A in a sufficiently smooth subgroup (the Lie Group) of $GL(n)$ in a neighborhood of the identity matrix I . The terminology follows from the fact that when $w = vX$, for a row vector v , then $dw = vdX = wX^{-1}dX$, or, equivalently, $dw = w\Omega$. See, for example, Flanders (1989).

Exercise 2.5. Following Definition 2.2.5, show that $G \times H$, together with \times_η , is a group in which;

- (1) The identity is $(1_G, 1_H)$;
- (2) The inverse $(\tau, \sigma)^{-1}$ of (τ, σ) is given by $(\alpha(\sigma^{-1})(\tau^{-1}), \sigma^{-1})$.

Exercise 2.6. Following Exercise 2.5, show that $G \times \{1_H\}$ is a normal subgroup of $G \times_\eta H$ (recall that N is a normal subgroup of G whenever $\tau N = N\tau$ for all $\tau \in G$).

Exercise 2.7. Consider the semi-direct product $S_2 \times_\alpha S_2$

$$(\sigma, \tau)(\eta, \theta) = (\sigma\alpha(\eta, \tau), \tau\theta) = (\sigma\tau\eta\tau^{-1}, \tau\theta)$$

of S_2 with S_2 . Show that the resulting group, with elements in $S_2 \times S_2$, has its multiplication table given by

| | | | | |
|-----------------|---|---|---|---|
| \times_α | a | b | c | d |
| a | a | b | c | d |
| b | b | a | d | c |
| c | c | d | a | b |
| d | d | c | b | a |

where $a = (1, 1)$, $b = (1, t)$, $c = (t, 1)$ and $d = (t, t)$.

Exercise 2.8. Following the notation of Example 2.2.9, let $B = \{(1, 1), (1, t)\}$ and verify that

$$\{B, (\eta, 1) \times_\alpha B, (\eta^2, 1) \times_\alpha B\}$$

is a partition of $G = C_3 \times_\alpha C_2$. It is called the imprimitive system generated by B (see also Rotman (1995, p. 257)).

is an action of G on $\mathcal{L}(V)$.

Exercise 2.9. Indicate by G the set of all non-singular $n \times n$ real doubly-stochastic matrices. If $A \in G$ then $Ae = e$ and $e'A = e'$, where e indicates the n -component vector of ones. Given $A \in G$ and $a = (a_1, a_2) \in \mathbb{R}^2$ define the $n \times n$ matrix $[a, A] = a_1ee' + a_2A$. The equality

$$[a, A][b, B] = (na_1b_1 + a_1b_2 + a_2b_1)ee' + a_2b_2AB,$$

and the fact that $AB \in G$, suggests the operation $ab = (na_1b_1 + a_1b_2 + a_2b_1, a_2b_2)$ in $\mathbb{R}^2 \times \mathbb{R}^2$, so that $[a, A][b, B] = [ab, AB]$. Show that \mathbb{R}^2 together with the product ab and the usual sum of vectors is an algebra.

Exercise 2.10. Show that $W = \{(a_1, a_2) \in \mathbb{R}^2; a_2 \neq 0, na_1 + a_2 \neq 0\}$, together with the product ab of Exercise 2.9, is a commutative group in which the unit is $(0, 1) \in W$ and, for $a \in W$,

$$a^{-1} = \left(\frac{-a_1}{a_2(na_1 + a_2)}, \frac{1}{a_2} \right) \in W$$

and $aa^{-1} = a^{-1}a = (0, 1)$.

Exercise 2.11. Show that $WG = \{[a, A]; a \in W, A \in G\}$, together with the operation $(a, A), (b, B) \rightarrow (ab, AB)$, is a group.

Exercise 2.12. With the notation of Exercise 2.11, show that when $G = \{I_n\}$, WG is the subgroup of all equicorrelated covariance matrices; When $W = \{(0, 1)\}$ and $G = S_n$, WG generates the group of $n \times n$ permutation matrices; When $W = \{(0, 1)\}$ and

$$G = \{w_0I_n + w_1g + w_2g^2 + \dots + w_{n-1}g^{n-1}; \sum_{i=0}^{n-1} w_i = 1, w_i \in \mathbb{R}\},$$

where g is a primitive element of order n in S_n , WG generates the subgroup of stochastic circulants with first row $w' = (w_0, \dots, w_{n-1})$. For example, take $n = 4$ and let F be a stochastic circulant with first row w' . Then $FF' = w_0I + w_1g^3 + w_2g^2 + w_3g \in G$ and

$$FF' = \alpha_0I + \alpha_1g + \alpha_2g^2 + \alpha_3g^3 = \begin{bmatrix} \alpha_0 & \alpha_1 & \alpha_2 & \alpha_3 \\ \alpha_1 & \alpha_0 & \alpha_1 & \alpha_2 \\ \alpha_2 & \alpha_1 & \alpha_0 & \alpha_1 \\ \alpha_3 & \alpha_2 & \alpha_1 & \alpha_0 \end{bmatrix}$$

is a symmetric stochastic circulant with first row determined by $\alpha_i = w'g^i w$.

Exercise 2.13. Consider the group algebra defined by the cyclic group $C_2 = \{1, \tau, \tau^2\}$ over the finite field $F = \{0, 1\}$ of two elements. Evaluate the addition and multiplication tables for the group algebra and verify that this algebra has non-null elements x, y such that $xy = 0$, that is, the algebra has a divisor of zero. See, for example, Dean (1966, p. 204).

Exercise 2.14. Based on Example 2.3.1 calculate $|\text{fix}(\tau)|$, $|G_s|$, and $|\mathcal{O}_s|$ and verify that

$$\text{number of orbits in } V = \frac{1}{|G|} \sum_{\tau \in G} |\text{fix}(\tau)|.$$

Exercise 2.15. From Matrix 2.2, calculate the isotropy group for $\{yyuu, yuyu\}$.

Exercise 2.16. Refer to Example 2.3.1 and consider the action σ_s of $S_2 = \{1, t\}$ on the set of binary sequences in length of two and let $\mathcal{O} = \{uy, yu\}$, where S_2 acts transitively. Apply the action σ_s componentwise on \mathcal{O}^m and show that the number k of orbits in \mathcal{O}^m is $k = 2^m - 1$.

Exercise 2.17. Let ρ be a representation of G (with g elements) on a finite dimensional vector space \mathcal{V} , in which a scalar product $(\cdot | \cdot)$ is defined e.g., Example 2.4.1. Show that

$$(x, y) = \frac{1}{g} \sum_{\tau \in G} (\rho_\tau x | \rho_\tau y)$$

is a scalar product in \mathcal{V} and that it satisfies $(\rho_\tau x, \rho_\tau y) = (x, y)$ for all $\tau \in G$ and all $x, y \in \mathcal{V}$.

Exercise 2.18. [Contributed by K.S. Mallesh] Show that the matrices centralized by D_3 have the pattern

$$\begin{bmatrix} \clubsuit & \diamond & \heartsuit & \spadesuit & \times & \star \\ \heartsuit & \clubsuit & \diamond & \star & \spadesuit & \times \\ \diamond & \heartsuit & \clubsuit & \times & \star & \spadesuit \\ \spadesuit & \star & \times & \clubsuit & \heartsuit & \diamond \\ \times & \spadesuit & \star & \diamond & \clubsuit & \heartsuit \\ \star & \times & \spadesuit & \heartsuit & \diamond & \clubsuit \end{bmatrix},$$

where

$$\clubsuit = h_{11} + h_{22} + h_{33} + h_{44} + h_{55} + h_{66}, \quad \diamond = h_{12} + h_{23} + h_{31} + h_{46} + h_{54} + h_{65}$$

$$\heartsuit = h_{13} + h_{21} + h_{32} + h_{45} + h_{56} + h_{64}, \quad \spadesuit = h_{14} + h_{25} + h_{36} + h_{41} + h_{52} + h_{63}$$

$$\times = h_{15} + h_{26} + h_{34} + h_{43} + h_{51} + h_{62}, \quad \star = h_{16} + h_{24} + h_{35} + h_{42} + h_{53} + h_{61}$$

define the transitive orbits relative to a generic 6×6 matrix $H = (h_{ij})$.

Exercise 2.19. For each y in a complex vector space \mathcal{V} , define $y^* : \mathcal{V} \rightarrow \mathbb{C}$ by $y^*(x) = (x, y)$. Show that y^* is a homomorphism of \mathcal{V} and the mapping $y \mapsto y^*$ is an isomorphism from \mathcal{V} into its dual space.

Exercise 2.20. Show that m in (2.14) is determined by the number of conjugacy classes in S_ℓ , when $c \geq \ell$. For $c < \ell$, m is the number of unordered decompositions of a positive integer ℓ as a sum of c non-negative integers, and is given by $m = \sum_{t=1}^c p(\ell, t)$, where, recursively, $p(n, k) = \sum_{t=1}^k p(n-1, t)$. Tables are available, e.g., Takács (1984).

Appendix A: Workshop

Short course on *Symmetry Studies*
TU Eindhoven, Maart 2005
Marlos Viana

In this workshop we will discuss a symmetry study to explore data indexed by the symmetries of a regular triangle (S_3). Note that this is the case in which the structure is in itself a group. All calculations are simple enough and can be done by pencil and paper. If you prefer, however, you may utilize the MAPLE codes shown in Appendix B.

Veel plezier!

- (1) Choosing a group of symmetries: Identify and interpret the symmetries of S_3 and its multiplication table in (2.6);
- (2) Choosing a group action: Let S_3 act on itself by conjugacy, that is,

$$\varphi(\tau, \sigma) = \tau\sigma\tau^{-1}.$$

- (3) Show that φ is a group action (Definition 2.3.1);
- (4) Construct a table similar to those of Example 2.4.2, describing the group action defined above;
- (5) Identify and interpret the orbits;
- (6) Illustrate the proof of Burnside Lemma by counting the number of fixed points and the size of the isotropy groups (stabilizers);
- (7) Use Routine 2.1 in Appendix B to evaluate the linear representations $\rho(\tau)$, $\tau \in S_3$;
- (8) Evaluate the character table of ρ ;
- (9) Identify the character table (2.24) of S_3 ;
- (10) Construct the canonical projections (\mathcal{P}_i) following Section 2.8;
- (11) Determine the dimension of the subspaces in the decomposition $\mathcal{V} = \mathbb{R}^6 = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \mathcal{V}_3$;
- (12) Determine the multiplicities with which the irreducible representations (1, sgn and β) appear in ρ , that is, the multiplicities in the decomposition of \mathcal{V}_i - e.g., Example 2.7.2;
- (13) Verify the properties of a canonical decomposition;
- (14) Define a data vector indexed by S_3 ;
- (15) Identify the invariants associated with each projection;
- (16) Interpret these invariants;
- (17) Decompose the sum of squares associated with the data vector;
- (18) Interpret each component;
- (19) Identify the parametric hypotheses associated with these invariants;
- (20) Identify the non-centrality parameters and degrees of freedom for normally distributed data. This item completes the first part of the lab.
- (21) Now consider another group action of the same group on the same set, namely, the multiplicative action $\varphi(\tau, \sigma) = \sigma\tau$ generated by the Cayley table of S_3 . Now repeat all steps above, compare and interpret the results. This item completes the second part of the lab;
- (22) Now consider the multiplicative action of C_3 (the rotations only) on the same set. This is simply the multiplicative action generated by the Cayley table restricted to rows a, e and f (defining C_3). Now repeat all steps above, compare and interpret the results;

- (23) To conclude this lab, write a summary underlying the effect that choosing different group actions and symmetries has on the analysis of the data indexed by the structure of interest.

2.12. Comments and solutions

The following example illustrates the role of different group actions on the reduction of data indexed by the symmetries (S_3) of a regular triangle. We remark that here the structure of interest is in itself a group.

In the first part of this example we consider the action

$$\left[\begin{array}{l|cccccc} \varphi(\tau, \sigma) = \tau\sigma\tau^{-1} & a & b & c & d & e & f \\ \hline a = 1 & a & b & c & d & e & f \\ b = (12) & a & b & d & c & f & e \\ c = (13) & a & d & c & b & f & e \\ d = (23) & a & c & b & d & f & e \\ e = (123) & a & d & b & c & e & f \\ f = (132) & a & c & d & b & e & f \end{array} \right]$$

of S_3 on itself by conjugacy, so that the resulting orbits are exactly the conjugacy classes of S_3 . The canonical decompositions associated with this action are given by

$$\mathcal{P}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/3 & 1/3 & 1/3 & 0 & 0 \\ 0 & 1/3 & 1/3 & 1/3 & 0 & 0 \\ 0 & 1/3 & 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \end{bmatrix}, \quad \mathcal{P}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & -1/2 \\ 0 & 0 & 0 & 0 & -1/2 & 1/2 \end{bmatrix},$$

and

$$\mathcal{P}_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2/3 & -1/3 & -1/3 & 0 & 0 \\ 0 & -1/3 & 2/3 & -1/3 & 0 & 0 \\ 0 & -1/3 & -1/3 & 2/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

of dimensions 3, 1 and 2, respectively. The invariants $\mathcal{P}x$ in the data vector

$$x' = (a, b, c, d, e, f),$$

indicated by their labels for simpler notation, associated with the $\mathcal{P}_1, \mathcal{P}_2$ and \mathcal{P}_3 are, respectively,

$$\left\{ a, \frac{1}{3}(b + c + d), \frac{1}{2}(e + f) \right\}, \quad \{ \pm(e - f) \},$$

and

$$\left\{ \frac{1}{3} \left[b - \frac{c+d}{2} \right], \frac{1}{3} \left[c - \frac{b+d}{2} \right], \frac{1}{3} \left[d - \frac{b+c}{2} \right] \right\}.$$

The components of the decomposition of $x'x$ are then

$$x'\mathcal{P}_1x = a^2 + \frac{1}{3}(a + b + c)^2 + \frac{1}{2}(e + f)^2, \quad x'\mathcal{P}_2x = \frac{1}{2}(e - f)^2,$$

and

$$x'\mathcal{P}_3x = \frac{2}{3}(b^2 + c^2 + d^2 - bc - bd - cd).$$

The parametric hypotheses afforded by this reduction are

- (1) H : $e = f$ that the two (non-trivial) rotation parameters are the same;
- (2) H : $b = c = d$, that all transposition parameters are the same.

In the second part of this symmetry study we consider the regular action $\varphi(\tau, \sigma) = \sigma\tau$ of S_3 on itself. This action is generated by the Cayley table of S_3 . In contrast, the resulting canonical reductions now are

$$\mathcal{P}_1 = 1/6 \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}, \quad \mathcal{P}_2 = 1/6 \begin{bmatrix} 1 & -1 & -1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 & -1 & -1 \\ -1 & 1 & 1 & 1 & -1 & -1 \\ -1 & 1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 \end{bmatrix}$$

and

$$\mathcal{P}_3 = 1/3 \begin{bmatrix} 2 & 0 & 0 & 0 & -1 & -1 \\ 0 & 2 & -1 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & -1 & 2 & 0 & 0 \\ -1 & 0 & 0 & 0 & 2 & -1 \\ -1 & 0 & 0 & 0 & -1 & 2 \end{bmatrix}.$$

Their dimensions are 1, 1 and 4, respectively. Under the regular action, $x'x$ now decomposes as

$$x'\mathcal{P}_1x = \frac{1}{6}(a + b + c + d + e + f)^2, \quad x'\mathcal{P}_2x = \frac{1}{6}(a + e + f - b - c - d)^2,$$

and

$$x'\mathcal{P}_3x = \frac{1}{3}[(a - e)^2 + (a - f)^2 + (e - f)^2 + (b - c)^2 + (b - d)^2 + (c - d)^2].$$

The parametric hypotheses afforded by this reduction are

- (1) H : $a + e + f = b + c + d$ that the parameter sum of rotations equals the parameter sum of transpositions;
- (2) H : $a = e = f$ and $b = c = d$, of homogeneity of rotation parameters and of homogeneity of transposition parameters.

Observe that these hypotheses are disjoint and that the corresponding quadratic forms are null when each one of them obtains.

Appendix B: Computing Algorithms

Routine 2.1 (Permutation matrices). This procedure generates the permutation matrices for a given permutation.

```
> restart:
> p3 := proc(a1,a2,a3) local f,m:
>   f:=[a1,a2,a3]:
>   m:=(i,j)->1-min(abs(f[i]-j),1):
>   Matrix(3,3,m):
> end:
> p3(3,2,1);
```

Example: If the basis (a,b,c) is mapped to (c,b,a),
the corresponding permutation matrix is:

```
>p3(3,2,1);
```

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Note: To evaluate a $n \times n$ permutation matrix write $[a_1, \dots, a_n]$ and the dimension of matrix m accordingly.

Routine 2.2 (Cayley Tables). This algorithm evaluates the multiplication table for S_3 .

```
> restart:
> with(group):
> t:=[[ ], [[1,2]], [[1,3]], [[2,3]], [[1,2,3 ]], [[1,3,2 ]]];
> ut:=[1,2,3,4,5,6];
> m:=(i,j)->mulperms(op(i,t),op(j,t));
> M:=Matrix(6,6,m):
> CS3:=subs( seq( op(i,t)=op(i,ut), i=1..nops(ut) ), evalm(M) );
```

Routine 2.3 (Invariants for C_{2h} -labeled structure). First note that C_{2h} is isomorphic to $C_2 \otimes C_2$ so that the character table indexed by $\{u, c, o, s\} = \{E, C_2, i, \sigma_h\}$ is 1, 1, 1, 1, 1, -1, 1, -1, 1, 1, -1, -1, 1, -1, -1, 1 respectively. Also note that procedure p8 has a change of basis (defined by the vector ff) so that the representations come in nice blocks.

```
> restart:
> with(LinearAlgebra):
> p8 := proc( a1,a2,a3,a4,a5,a6,a7,a8 )
>   local f,ff,m:
>   f:=[a1,a2,a3,a4,a5,a6,a7,a8]:
>   ff:=[7,8,1,2,3,4,5,6]:
>   m:=(i,j)->1-min(abs(f[i]-ff[j]),1):
>   Matrix(8,8,m):
> end proc:
> c:=p8(1,2,7,8,5,6,3,4); o:=p8(2,1,8,7,6,5,4,3);
> s:=p8(8,7,2,1,4,3,6,5); u:=p8(7,8,1,2,3,4,5,6);
```

```

> Q1:=(u+c+o+s)/4; Q2:=(u-c+o-s)/4;
> Q3:=(u+c-o-s)/4; Q4:=(u-c-o+s)/4;
> x:=<abb,abB,aBb,aBB,Abb,AbB,ABb,ABB>;
> MatrixVectorMultiply(Q1,x);
> MatrixVectorMultiply(Q2,x);
> MatrixVectorMultiply(Q3,x);
> MatrixVectorMultiply(Q4,x);

Routine 2.4 (Regular projections for  $S_3$ ). > restart:
> with(group):

> t:=[[ ], [[1,2]], [[1,3]], [[2,3]], [[1,2,3]], [[1,3,2]]];

> ut:=[1,2,3,4,5,6];
> m:=(i,j)->mulperms(op(i,t),op(j,t));

> M:=Matrix(6,6,m):
> CS3:=subs( seq( op(i,t)=op(i,ut), i=1..nops(ut) ), evalm(M) );

> delta:=(i,j)->floor(2^(-abs(i-j)));

> f:=(i,j,k)->floor(2^(-abs(CS3[k,i]-CS3[1,j]))):

> c1:=<1,1,1,1,1,1>;
> c2:=<2,0,0,0,-1,-1>;
> c3:=<1,-1,-1,-1,1,1>;

> P1:=Matrix(6,6,(i,j)->add(c1[k]*f(i,j,k),k=1..6))/6;
> P2:=Matrix(6,6,(i,j)->add(c2[k]*f(i,j,k),k=1..6))*2/6;
> P3:=Matrix(6,6,(i,j)->add(c3[k]*f(i,j,k),k=1..6))/6;

```

```

Routine 2.5 (Regular projections for  $S_4$ ). > restart:
> with(group):
> t:=[[ ], [[3,4]], [[2,3 ]], [[2,4 ]], [[1,2 ]], [[1,3 ]], [[1,4]],
> [[2,3,4 ]], [[2,4,3 ]], [[1,2,3 ]], [[ 1,2,4]], [[1,3,2 ]],
> [[1,3,4 ]], [[1,4,2 ]], [[1,4,3 ]], [[1,2], [3,4]],
> [[1,3], [2,4 ]], [[1,4], [2,3 ]], [[1,2,3,4 ]],
> [[1,2,4,3 ]], [[1,3,2,4 ]], [[1,3,4,2 ]], [[1,4,3,2 ]],
> [[1,4,2,3 ]]];
> ut:=[1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,
> 21,22,23,24];
> m:=(i,j)->mulperms(op(i,t),op(j,t));
> M:=Matrix(24,24,m):
> CS4:=subs( seq( op(i,t)=op(i,ut), i=1..nops(ut) ), evalm(M) ):
> delta:=(i,j)->floor(2^(-abs(i-j)));
> f:=(i,j,k)->floor(2^(-abs(CS4[k,i]-CS4[1,j]))):
> c1:=<1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1>;
> c2:=<3,1,1,1,1,1,1,0,0,0,0,0,0,0,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1>;
> c3:=<2, 0,0,0,0,0,0, -1,-1,-1,-1,-1,-1,-1,-1,2,2,2, 0,0,0,0,0,0>;
> c4:=<3,-1,-1,-1,-1,-1,-1, 0,0,0,0,0,0,0,-1,-1,-1, 1,1,1,1,1,1,1,1>;
> c5:=<1, -1,-1,-1,-1,-1,-1,1,1,1,1,1,1,1,1,1, -1,-1,-1,-1,-1,-1,-1,-1>;
> P1:=Matrix(24,24,(i,j)->add(c1[k]*f(i,j,k),k=1..24))/24;
> P2:=Matrix(24,24,(i,j)->add(c2[k]*f(i,j,k),k=1..24))*3/24;
> P3:=Matrix(24,24,(i,j)->add(c3[k]*f(i,j,k),k=1..24))*2/24;

```

```
> P4:=Matrix(24,24,(i,j)->add(c4[k]*f(i,j,k),k=1..24))*3/24;
> P5:=Matrix(24,24,(i,j)->add(c5[k]*f(i,j,k),k=1..24))/24;
```

Routine 2.6. This code generates the canonical projections of the regular representation of D_4 .

```
> restart:
> with(group):
> t:=[[ ], [[1,2,3,4 ]], [[1,3], [2,4]], [[1,4,3,2 ]], [[1,4], [2,3 ]],
> mulperms([[1,4], [2,3 ]], [[1,2,3,4]]), mulperms([[1,4], [2,3 ]], [[1,3], [2,4]]),
> mulperms([[1,4], [2,3 ]], [[1,4,3,2]]) ];
> ut:=[1,2,3,4,5,6,7,8];
> m:=(i,j)->mulperms(op(i,t),op(j,t));
> M:=matrix(8,8,m):
> CD4:=subs( seq( op(i,t)=op(i,ut), i=1..nops(ut) ), evalm(M) ):
> delta:=(i,j)->floor(2^(-abs(i-j)));
> f:=(i,j,k)->floor(2^(-abs(CD4[k,i]-CD4[1,j]))):

> c1:=<1,1,1,1,1,1,1,1>;
> c2:=<1,1,1,1,-1,-1,-1,-1>;
> c3:=<1,-1,1,-1,1,-1,1,-1>;
> c4:=<1,-1,1,-1,-1,1,-1,1>;
> c5:=<2,0,-2,0,0,0,0,0>;

> P1:=Matrix(8,8,(i,j)->add(c1[k]*f(i,j,k),k=1..8))/8;
> P2:=Matrix(8,8,(i,j)->add(c2[k]*f(i,j,k),k=1..8))/8;
> P3:=Matrix(8,8,(i,j)->add(c3[k]*f(i,j,k),k=1..8))/8;
> P4:=Matrix(8,8,(i,j)->add(c4[k]*f(i,j,k),k=1..8))/8;
> P5:=Matrix(8,8,(i,j)->add(c5[k]*f(i,j,k),k=1..8))*2/8;
```


Bibliography

- Aigner, M. (1979), *Combinatorial Theory*, Springer-Verlag, New York, NY.
- Andersson, S. (1992), *Normal Statistical Models Given by Group Symmetry*, DMV Seminar Lecture Notes, Günzburg, Germany.
- Arratia, R. and Simon, T. (92), 'The cycle structure of random permutations', *Annals of Probability* **20**, 1567–1591.
- Bacry, H. (1963), *Leçons sur la Théorie des Groupes*, Université D'Aix-Marseille, Marseille, France.
- Bailey, R. A. (1991), 'Strata for randomized experiments', *Journal of the Royal Statistical Society B* (53), 27–78.
- Bishop, D. M. (1973), *Group Theory and Chemistry*, Dover, New York, NY.
- Cameron, P. J. (1994), *Combinatorics: Topics, Techniques and Algorithms*, Cambridge University Press, New York.
- Cameron, P. J. (1999), *Permutation Groups*, Cambridge University Press, New York.
- Campbell, C. (1997), 'The refractive group', *Optometry and Vision Science* **74**, 381–387.
- Cartan, E. (1966), *The Theory of Spinors*, MIT Press, Cambridge, MA.
- Cartier, P. (2001), 'A mad day's work: from Grothendiek to Connes and Kontsevich- the evolution of concepts of space and symmetry', *Bulletin (New Series) of the American Mathematical Society* **38**(4), 389–408.
- Curtis, M. L., ed. (1984), *Matrix Groups*, Springer-Verlag, New York, NY.
- Dawid, A. P. (1988), 'Symmetry models and hypothesis for structured data layouts', *Journal of the Royal Statistical Society B* (50), 1–34.
- de Finetti, B. (1972), *Probability, Induction and Statistics*, Wiley, New York.
- Dean, R. A. (1966), *Elements of Abstract Algebra*, Wiley, New York, NY.
- Diaconis, P. and Freedman, D. (1999), 'Iterated random functions', *SIAM Review* **41**(1), 45–76.
- Diaconis, P. and Shahshahani, M. (1986), Product of random matrices as they arise in the study of random walks on groups, in 'Random Matrices and Their Applications', American Mathematical Society, Providence, RI, pp. 183–195.
- Dixon, J. D. and Mortimer, B. (1996), *Permutation Groups*, Springer, New York, NY.
- Doi, H. (1991), 'Importance of purine and pyrimidine content of local nucleotide sequences (six bases long) for evolution of human immunodeficiency virus type 1', *Evolution* **88**(3), 9282–9286.
- Eaton, M. L. (1983), *Multivariate Statistics- A Vector Space Approach*, Wiley, New York, NY.
- Enquist, M. and Arak, A. (1994), 'Symmetry, beauty and evolution', *Nature* **371**, 169–72.
- Faris, W. (1996), 'Review of Roland Omnès, the interpretation of quantum mechanics', *Notices of the American Mathematical Society* **43**(11), 1328–39.
- Ferris 3rd, F., Freidlin, V., Kassoff, A., Green, S. and Milton, R. (1993), 'Relative letter and position difficulty on visual acuity charts from the early treatment diabetic retinopathy study.', *Am J Ophthalmol* **15**, 735–40.
- Flanders, H. (1989), *Differential Forms with Applications to the Physical Sciences*, Dover (Academic Press, 1963), New York, NY.
- Gao, Y. and Marden, J. (2001), Some rank-based hypothesis test for covariance structure and conditional independence, in M. Viana and D. Richards, eds, 'Contemporary Mathematics- Algebraic methods in statistics and probability', Vol. 287, American Mathematical Society, Providence, RI, pp. 97–110.
- Graf, R. and Schachman, H. (1996), 'Cyclic symmetries. random circular permutation of genes and expressed polypeptide chains: application of the method to the catalytic chains of aspartate transcarbamoylase.', **93**(21), 11591–6.
- Halmos, P. R. (1987), *Finite-Dimensional Vector Spaces*, Springer-Verlag, New York.
- Harris, D. C. and Bertolucci, M. D. (1978), *Symmetry and Spectroscopy- An Introduction to Vibrational and Electronic Spectroscopy*, Oxford University Press, New York, NY.
- Hellige, J. B. (1993), *Hemispheric Asymmetry*, Harvard U. Press, Cambridge, MA.
- Kolchin, V. F. (1971), 'A problem of the allocation of particles in cells and cycles of random permutations', *Theory of Probability and Its Applications* **16**(1), 74–90.
- Lakshminarayanan, V. and Viana, M. (2005), 'Dihedral representations and statistical geometric optics I: Spherocylindrical lenses', *J. Optical Society of America A* **22**(11), 2483–89.
- Lalley, S. (96), 'Cycle structure of riffle shuffles', *Annals of Probability* **24**, 49–73.
- Lam, T. Y. (1998), 'Representations of finite groups: A hundred years, parts i,ii', *Notices of the AMS* **45**(3), 361–372.
- Lee, H. (1998), The covariance structure of concomitants of ordered symmetrically dependent observations, PhD thesis, The University of Illinois at Chicago Division of Biostatistics.

- Lee, H. and Viana, M. (1999), 'The joint covariance structure of ordered symmetrically dependent observations and their concomitants of order statistics', *Statistics and Probability Letters* **43**, 411–414.
- Lestrel, P. E., Takahashi, O. and Kanazawa, E. (2004), 'A quantitative approach for measuring crowding in dental arch: Fourier descriptors', *Am. J. of Orthodontics and Dental Orthopedics*.
- Marshall, A. W. and Olkin, I. (1979), *Theory of Majorization and its Applications*, Academic Press, New York, NY.
- McCullagh, P. (1999), 'Quotient spaces and statistical models', *The Canadian Journal of Statistics* **27**(3), 447–456.
- McCullagh, P. (2000), 'Invariance and factorial models', *J. R. Stat. Soc. Ser. B Stat. Methodol* **62**, 209–256.
- Muirhead, R. J. (1982), *Aspects of Multivariate Statistical Theory*, Wiley, New York.
- Naimark, M. A. and Štern, A. I. (1982), *Theory of Group Representations*, Springer-Verlag, New York, NY.
- Oliveira, N. L., Silveira, A., Kusnoto, B. and Viana, G. (2004), 'Three-dimensional assessment of morphologic changes in the maxilla: A comparison of 2 kinds of palatal expanders', *Am J. of Orthodontics and Dentofacial Orthopedics* pp. 354–362.
- Omnès, R. (1994), *The Interpretation of Quantum Mechanics*, Princeton Press, Princeton, NJ.
- Pauleikhoff, D., Wormald, R., Wright, L., Wessing, A. and Bird, A. (1992), 'Macular disease in an elderly population', *Ger J Ophthalmol* **1**(1), 12–15.
- Pistone, G., Riccomagno, E. and Wynn, P. (2000), *Algebraic Statistics: Computational Commutative Algebra in Statistics*, Chapman & Hall/CRC, Boca Raton, FL.
- Pólya, G. (1954), Patterns of plausible inference, in 'Mathematics and Plausible Reasoning', Vol. 2, Princeton University Press, Princeton, NJ.
- Rao, C. R. (1973), *Linear Statistical Inference and Its Applications*, Wiley, New York.
- Riley, K. F., Hobson, M. P. and Bence, S. J. (2002), *Mathematical Methods for Physics and Engineering*, 2nd edn, Cambridge University Press, New York, NY.
- Rosen, J. (1975), *Symmetry Discovered*, Dover, Mineola, NY.
- Rosen, J. (1995), *Symmetry in Science, An Introduction to the General Theory*, Springer-Verlag, New York.
- Rotman, J. J. (1995), *An Introduction to the Theory of Groups*, 4 edn, Springer-Verlag, New York.
- Ruhla, C. (1989), *The Physics of Chance*, Oxford Press, New York, NY.
- Salmon, T. O., Thibos, L. N. and Bradley, A. (1998), 'Comparison of the eye's wave-front aberration measured psychophysically and with the shack-hartmann wave-front sensor', *J. Optical Society of America A* **15**(9), 2457–64.
- Schrödinger, E. (1967), *What is Life? The Physical Aspects of the Living Cell*, Cambridge University Press, New York, NY.
- Searle, S. R. (1971), *Linear Models*, Wiley, New York, NY.
- Serre, J.-P. (1977), *Linear Representations of Finite Groups*, Springer-Verlag, New York.
- Simon, B. (1996), *Representations of Finite and Compact Groups*, American Mathematical Society, Providence, RI.
- Snedecor, G. W. and Cochran, W. G. (1989), *Statistical Methods*, 8th edn, Iowa State University Press, Ames, IO.
- Stanton, D. and White, D. (1986), *Constructive Combinatorics*, Springer-Verlag, New York.
- Swaddle, J. P. (1999), 'Visual signalling by asymmetry: a review of perceptual processes', *Philos Trans R Soc Lond B Biol Science*.
- Szlyk, J. P., Seiple, W. and Xie, W. (1995), 'Symmetry discrimination in patients with retinitis pigmentosa', *Vision Research* **35**(11), 1633–1640.
- Szlyk, J., Rock, I. and Fisher, C. (1995), 'Level of processing in the perception of symmetrical forms viewed from different angles', *Spatial Vision* **9**(1), 139–150.
- Takács, L. (1984), Combinatorics, in P. R. Krishnaiah and P. K. Sen, eds, 'Handbook of Statistics', Vol. 4, North-Holland, New York, NY, pp. 123–173.
- Tuljapurkar, S. (1986), Demographic applications of random matrix products, in 'Random Matrices and Their Applications', American Mathematical Society, Providence, RI, pp. 319–326.
- Tyler, C. W. (1996), *Human symmetry perception and its computational analysis*, VSP, Utrecht, The Netherlands.
- Viana, M. (2003a), 'Invariance conditions for random curvature models', *Methodology and Computing in Applied Probability* **5**, 439–453.
- Viana, M. (2003b), *Symmetry Studies- An Introduction*, IMPA Institute for Pure and Applied Mathematics (Lecture Notes- 24th Brazilian Mathematics Symposium), Rio de Janeiro, Brazil.
- Viana, M. (2004), Symmetry analysis of refraction data, in 'Proceedings of the 2nd International Workshop in Applied Probability', University of Piraeus, Greece, pp. 336–337.
- Viana, M. and Olkin, I. (2000), 'Symmetrically dependent models arising in visual assessment data', *Biometrics* (56), 1188–1191.
- von Mises, R. (1957), *Probability, Statistics and Truth*, Dover, New York, NY.
- Wagemans, J. (1999), 'Parallel visual processes in symmetry perception: Normality and pathology', *Documenta Ophthalmologica* **95**, 359–370.
- Weyl, H. (1950), *The Theory of Groups and Quantum Mechanics*, Dover, New York, NY.
- Weyl, H. (1952), *Symmetry*, Princeton U. Press, Princeton, NJ.
- Weyl, H. (1953), *The Classical Groups, Their Invariants and Representations*, Princeton University Press, Princeton, NJ.
- Wijsman, R. A. (1990), *Invariant Measures on Groups and Their Use in Statistics*, Vol. 14, Institute of Mathematical Statistics, Hayward, California.

- Wit, E. and McCullagh, P. (2001), The extendibility of statistical models, *in* M. Viana and D. Richards, eds, 'Contemporary Mathematics- Algebraic methods in statistics and probability', Vol. 287, American Mathematical Society, Providence, RI, pp. 327–340.
- Youden, W. J. (1951), *Statistical Methods for Chemists*, Wiley, New York, NY.

Index

- action
 - transitive, 40
- automorphism, 37
- biological sequence, 6
- Bose-Einstein statistic, 14
- Cameron, P.J., 85
- Cartier, P., 21
- Cayley Table
 - S_3 , 35
- character table, 63
- class function, 65
- Cochran, W.G., 21
- composition
 - on the left, 8
 - on the right, 8
- conjugacy class, 66
- Curtis, M.L., 85
- Diaconis, P., 86
- Dixon, J.D., 85
- Doi, H., 22
- Eaton, M., 21
- entropy, 16
- Fermi-Dirac statistic, 15
- group
 - action, 40
 - commutative, 35
 - isotropy, 40
 - order, 35
- group algebra, 49
- Halmos, P., 85
- homomorphism, 37
- inner product
 - invariant, 48
- integer partition, 11
- integer partitions, 11
- isomorphism, 37
- Lam, T.Y., 85
- Lee, H., 22
- lense refraction, 12
- Maxwell-Boltzmann statistic, 14
- McCullagh, P., 22, 81
- Muirhead, R., 21
- Nairmark, M.A., 85
- Olkin, I., 22
- orbit, 40
 - cyclic, 40
 - dihedral, 41
 - permutation, 8
- orbit
 - cyclic, 8
- permutation group, 8
- permutation symmetry, 8
- permutations, 8
- permutations, cyclic, 33
- Pistone, G., 22
- primitive idempotents, 58
- probability laws
 - exchangeable, 12
- quotient space, 10, 42
- Rao, C.R., 21
- representation
 - alternating square, 53
 - character of, 59
 - irreducible, 54
 - irreducible character of, 62
 - irreducible of dimension $(n - 1)$, 70
 - regular, 46, 65
 - symmetric square, 53
 - tensor, 49
- Rotman, J.J., 85
- Ruhla, C., 16
- Searle, S.R., 21
- semidirect product, 38
- semilinear, 62
- Serre, J., 85
- Simon, B., 85
- Snedecor, G.W., 21
- stabilizer, 40
- Stanton, D., 85
- structure, 7
 - factorial, 77
 - fractional, 79
- structured data, 6
- Symmetry Principle, 20
- Takács, L., 88
- Viana, M., 22

Wheatstone bridge, 26
Wijsman, R.A., 22