

# TOWARDS THE CALCULATION OF JACK POLYNOMIALS

by

Leigh Alan Roberts

A thesis  
submitted to the Victoria University of Wellington  
in fulfilment of the  
requirements for the degree of  
Doctor of Philosophy  
in Operations Research and Statistics

Victoria University of Wellington

2001

## ABSTRACT

Jack polynomials are useful in mathematical statistics, but they are awkward to calculate, and their uses have chiefly been theoretical.

In this thesis a determinantal expansion of Jack polynomials in elementary symmetric polynomials is found, complementing a recent result in the literature on expansions as determinants in monomial symmetric functions. These results offer enhanced possibilities for the calculation of these polynomials, and for finding workable approximations to them.

The thesis investigates the structure of the determinants concerned, finding which terms can be expected to dominate, and quantifying the sparsity of the matrices involved.

Expressions are found for the elementary and monomial symmetric polynomials when the variates involved assume the form of arithmetic and geometric progressions. The latter case in particular is expected to facilitate the construction of algorithms suitable for approximating Jack polynomials.

## ACKNOWLEDGEMENTS

It is a pleasure to acknowledge the contribution from several quarters to the writing of this thesis.

Thanks should first pass to an anonymous referee of Roberts (2001), on which Chapter 3 of this thesis is partly based. The referee suggested an alternative approach to simplify a central result. I have acknowledged the assistance in the article cited, and make no further mention of it in the thesis.

Peter Forrester, Luc Lapointe, Yasuko Chikuse and Ian G Macdonald were kind enough to send me copies of some of the articles and booklets which they had authored. Other individuals took the trouble to email me with advice and encouragement. Thanks are due to all of them.

Peter Smith explained to me more about electronic engineering and mobile phones than I really wanted to know, but I appreciated his efforts. It was Peter and Hongsheng Gao who introduced me to zonal polynomials rather longer ago than I care to remember.

Several other friends and colleagues at Victoria University have been of assistance. I see little point in presenting a long list here, but it would be churlish not to acknowledge the substantial debt I owe to Chris Atkin and John Harper on the mathematical side; and Ray Brownrigg and Edith Hodgen on the computing front. I am grateful.

My supervisor David Vere-Jones has been a model of restraint, in allowing me free rein to indulge my interests, with the occasional sojourn back to base, over a lengthy period. But it was David's generosity with his time towards the end of my time, his acuity in guiding me to tie together what was threatening to become a quite unwieldy thesis, and above all his formidable mathematical and statistical expertise in so many unexpected and quite eclectic areas, which eventually made this thesis so much fun to write, all the hard work notwithstanding. To have had David as a

supervisor was a great privilege.

Finally, to my children Imogen and Timothy, and my wife Phillippa, who have borne more than a little ill-humour on my part over what became an unsociable series of years rather than months, and complained less than they might when they saw less of me than they ought, I dedicate this thesis. With much love.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Preliminary . . . . .	1
1.2	Background . . . . .	2
1.3	Aims of the thesis . . . . .	5
1.4	What is in the thesis? . . . . .	9
1.5	Notation . . . . .	15
1.5.1	Ordering of partitions . . . . .	15
1.5.2	d2 chains and d2-1 chains . . . . .	17
1.5.3	Augmented $m_\lambda$ functions; $e_\lambda$ and $p_\lambda$ functions etc. . . . .	17
1.5.4	Bisymmetric matrices, representations . . . . .	18
1.5.5	Note on product of tensors . . . . .	19
1.5.6	Other definitions . . . . .	19

<b>2</b>	<b>Theoretical overview and literature survey</b>	<b>22</b>
2.1	Farrell's listing of the different approaches to zonal polynomials . . .	22
2.1.1	Farrell's approach to zonal polynomials . . . . .	23
2.1.2	The approach through group representation theory . . . . .	23
2.1.3	Saw's approach . . . . .	24
2.1.4	Laplace-Beltrami approach . . . . .	26
2.2	Other approaches to zonal polynomials and Jack polynomials . . . . .	27
2.2.1	Direct definition from orthogonality . . . . .	28
2.2.2	The combinatorial approach of Knop and Sahi . . . . .	29
2.3	Related literature . . . . .	31
2.3.1	Differential equations . . . . .	31
2.3.2	Hypergeometric functions of more than one matrix argument .	32
2.3.3	Jack polynomials in Physics . . . . .	34
2.3.4	Statistics defined on the sphere . . . . .	35
2.3.5	An engineering application . . . . .	38
2.4	Technical overview . . . . .	40
2.4.1	Group $G$ acting transitively on a set $Z$ . . . . .	40

2.4.2	Representations of the group $G$ . . . . .	42
2.5	James' definition of zonal polynomials . . . . .	49
2.5.1	The situation of interest . . . . .	51
2.5.2	Overview of James' reasoning . . . . .	51
2.5.3	Description of the diagram (2.13) . . . . .	52
2.5.4	The basic representation of interest $R_2$ . . . . .	53
2.5.5	The representation $R_7$ leading to $R_1$ . . . . .	55
2.5.6	From $R_{0a}$ to $R_3$ , and from $R_{0b}$ to $R_6$ . . . . .	59
2.5.7	Comparing $R_4$ and $R_5$ . . . . .	61
2.5.8	Comparing $R_5$ and $R_6$ . . . . .	62
2.6	Macdonald's definition of zonal polynomials . . . . .	64
2.6.1	The initial setup . . . . .	64
2.6.2	Development of Macdonald's ideas . . . . .	65
2.7	Vilenkin and Klimyk's work . . . . .	68
2.8	Farrell's work on zonal polynomials and summary of James', Mac- donald's and Farrell's approaches . . . . .	69
2.8.1	Farrell's approach to complex zonal polynomials . . . . .	70

2.8.2	Mini summary of James', Macdonald's and Farrell's arguments	72
2.9	The place of zonal polynomials in statistical theory . . . . .	73
2.9.1	The central place of zonal polynomials in multivariate statistical theory . . . . .	73
2.9.2	Little need for exact calculation . . . . .	74
2.9.3	Two basic problems with the application of zonal polynomials	75
2.9.4	Size of matrices . . . . .	77
3	Determinantal expansions of $J_\rho$ in terms of $e_\lambda$ and $m_\lambda$ functions, and their interrelationship	79
3.1	Preliminary . . . . .	79
3.1.1	Introduction . . . . .	79
3.1.2	Overview . . . . .	80
3.2	Determinantal forms of $J_\rho$ in $e_\lambda$ and $m_\lambda$ functions . . . . .	81
3.2.1	Preliminary technical results . . . . .	81
3.2.2	$J_\rho$ as a determinantal form in $e_\lambda$ functions . . . . .	83
3.2.3	$J_\rho$ as a determinantal form in $m_\lambda$ functions . . . . .	84
3.2.4	Shape of determinantal expansions . . . . .	85
3.3	Symmetry between the operator matrices $\Omega^m$ and $\Omega^e$ . . . . .	87

3.3.1	Preliminary technical material . . . . .	87
3.3.2	The principal theorem . . . . .	87
3.4	The action of $\mathcal{L}$ on the $e_\lambda$ functions . . . . .	89
3.4.1	Technical support for the calculation of $\mathcal{L}e_\lambda$ . . . . .	90
3.5	Consistency of $\Omega^e$ and $\Omega^m$ derived by different means . . . . .	94
3.6	Extended example . . . . .	96
<b>4</b>	<b>The composition of the operator matrices <math>\Omega^e</math> and <math>\Omega^m</math></b>	<b>98</b>
4.1	Derivation of $\Omega^m$ through core chains . . . . .	99
4.2	Sparsity of $\Omega_w$ . . . . .	101
4.3	From the operator matrix $\Omega$ to the coefficient matrix $C$ . . . . .	104
4.4	The composition of the $j$ coefficients . . . . .	105
4.4.1	The composition of the $j_{\rho,\lambda}^m$ coefficients . . . . .	106
4.4.2	The composition of the $j_{\rho,\lambda}^e$ coefficients . . . . .	107
4.4.3	Dominant terms in the determinantal expansions . . . . .	108
4.4.4	Path descriptors in terms of core chains . . . . .	109
4.4.5	Incidence matrices . . . . .	111

4.5	Relative magnitudes of diagonal and off-diagonal terms in $\Omega^m$ . . . . .	111
4.6	Slicing and reflection results on $j_{\rho,\lambda}$ coefficients . . . . .	116
4.6.1	The basic supporting result for the slicing and reflection the- orems . . . . .	117
4.6.2	Slicing theorems . . . . .	118
4.6.3	Reflection results . . . . .	123
4.7	Relative parsimony of expansions in $e_\lambda$ and $m_\lambda$ functions . . . . .	126
<b>5</b>	<b>The operator and coefficient matrices <math>\Omega^m</math> and <math>C^m</math> for indeterminate weight <math>w</math></b>	<b>129</b>
5.1	The generalised operator matrix $g\Omega^m$ . . . . .	130
5.1.1	The generalised coefficient matrix $gC^m$ . . . . .	131
5.2	The $j_{\rho,\lambda}^m$ coefficients for $\lambda$ low in the RLO . . . . .	138
5.2.1	Finding the values of $g_j$ . . . . .	144
5.2.2	Supporting Maple code . . . . .	147
<b>6</b>	<b>Evaluation of the <math>m_\lambda</math> functions</b>	<b>149</b>
6.1	Summary . . . . .	149
6.2	Preliminary . . . . .	150

6.3	The expansion of $am_\lambda$ in $p_\lambda$ functions . . . . .	151
6.3.1	Definitions . . . . .	153
6.3.2	Examples . . . . .	155
6.3.3	Proof of Theorem 6.1 . . . . .	156
6.4	Behaviour of the $am_\lambda$ functions when the underlying variables form an arithmetic progression . . . . .	158
6.5	Behaviour of the $am_\lambda$ functions when the underlying variables form a geometric progression . . . . .	163
6.5.1	Infinite number of variates . . . . .	164
7	<b>Evaluation of the <math>e_\lambda</math> functions when variates form a geometric pro- gression</b>	<b>165</b>
7.1	A tangent: when the variates are symmetric about the origin . . . .	166
7.2	Calculating $e_r$ : one can assume that $r \leq n/2$ . . . . .	166
7.3	The basic recursion for $e_r$ . . . . .	169
7.4	Magnitude of successive terms in the expansion of $e_r$ . . . . .	174
7.5	Approximate evaluation of $e_r$ . . . . .	178
7.5.1	Definitions . . . . .	178
7.5.2	Preliminary work . . . . .	180

7.5.3	Principal results . . . . .	182
7.6	Calculating $e_\lambda$ : the main theorem . . . . .	185
7.7	Approximations to $e_\lambda$ . . . . .	186
8	Conclusion	187
8.1	A wishlist . . . . .	188
8.2	Relative merits of expanding Jack polynomials in the $e_\lambda$ and $m_\lambda$ functions . . . . .	189
8.3	On the assumptions that variates form arithmetic or geometric progressions . . . . .	190
8.4	A beginning, not the end . . . . .	191
	Bibliography	192

# List of Figures

3.1	$\det(\Omega_{(3,2),1}^e)$ and $\det(\Omega_{(5),1}^e)$	86
3.2	$\det(\Omega_{(3,2),1}^m)$ and $\det(\Omega_{(5),1}^m)$	86
4.1	The chain matrix $CM_5$	100
4.2	The operator matrix $\Omega_5^m$	102
4.3	$\Omega_{(4),1}^m$	106
4.4	$\Omega_{(4),1}^e$	107
4.5	d2 path $abbhj$ from $(5)$ to $(1^5)$	110
4.6	d2 path $ae fj$ from $(5)$ to $(1^5)$	110

# List of Tables

- 5.1 The generalised operator matrix  $g\Omega^m$  up to the 3rd block . . . . . 130
- 5.2 The generalised coefficient matrix  $gC^m$  up to the 5th block . . . . . 132

# Chapter 1

## Introduction

### 1.1 Preliminary

The brief title of this thesis gives little indication of its scope. In the first instance, much of the thesis, and in particular most of the discussion on the literature, will treat of zonal polynomials, the special case of Jack polynomials of greatest interest in statistics.

Zonal polynomials have enjoyed a certain currency in mathematical statistics, but their uses have chiefly been theoretical, especially as an aid to the evaluation of integrals arising in multivariate inference. The overall intention of this thesis is to advance towards the stage of being able to use zonal polynomials in practical statistical situations. The principal results concern more efficient means of calculating and approximating zonal polynomials, and those methods apply just as much to Jack polynomials as to zonal polynomials, whence the title.

The title however falls short in several other and perhaps more significant ways. There is an attempt to discuss several putative paths that one could follow to attempt to calculate zonal polynomials, and some false steps taken by the author along this route; we look at how much one really needs to calculate zonal poly-

mials since they are so slightly amenable to calculation, and what attempts have been made to obviate their calculation; we also look at approximations to zonal polynomials, either asymptotically or by assuming special forms of the variates involved; and we look at their generalisations, such as Jack polynomials, Macdonald polynomials, and the increasing number of asymmetric polynomials from which the symmetric polynomials are obtained by symmetrisation. But rather than inserting this lengthy list into an omnibus title, it was decided that one could live with the potential inaccuracy of the abbreviated version.

## 1.2 Background

Zonal polynomials arise essentially from considerations of symmetry of multivariate statistical data, particularly rotational symmetry, and especially when estimating covariance matrices under various statistical hypotheses. The constraints implicit in parametrising these symmetric situations can be imposed in a straightforward fashion in standard Euclidean space, but at the cost of both complicating enormously the evaluation of integrals necessary to obtain density functions, and largely losing the ability to visualise the situation conceptually by obfuscation of the underlying symmetry.

The alternative is to parametrise in order to reflect the symmetry of the underlying situation in the parameter space. Intuition is preserved, and otherwise intractable integrals evaluated. The price to pay is that the Jacobians/differential forms involved are awkward, on which matter see James (1954, p. 41) for a discussion.

The role of James (1954) in alerting statisticians to the advantages in adopting the latter approach, and gathering the mathematical machinery together in an eminently readable fashion to enable the statistical community to avail itself of the opportunity, was pivotal. James gave the invariant measures for Stiefel and Grassman manifolds, and indicated the way in which orthogonal projections lead to factorisation of measures. Invariance and ancillarity are among several areas in statistical theory utilising James' ideas (Hillier & Skeels (1996, p. 161)).

James' paper is all the more remarkable in view of his tender age when it was written, and was an impressive vindication of Cornish's encouragement of James when a student in Adelaide to further his knowledge of algebra, particularly invariants and related topics, when mathematical interest in these areas had dwindled somewhat after their apogee in the late 19th and early 20th centuries. Publication of the paper, essentially the first half of a PhD thesis, was encouraged by several well-known mathematicians and statisticians, including Wilks and Feller. The other half of the thesis applied the theory of the first half, and was published as James (1955a) and James (1955b), which were introduced to the Royal Society by Fisher (Hillier & Skeels (1996, p. 160)).

James had a major influence in subsequent developments, even if he published little directly in the area of zonal polynomials in later years. Saw (1977) and Muirhead (1975), for example, were introduced by James; and Anderson (1965) acknowledges James' part in both suggesting the problem dealt with in his paper and the means of tackling it. Muirhead spent some time in Adelaide with James and Constantine, and Muirhead (1982) remains a hugely influential book on multivariate statistics and zonal polynomials.

Having adopted the approach to parametrisation indicated by James, and having available expressions for the Haar invariant measures under transformation groups appropriate to a given situation, one simultaneously had a natural definition of a uniform distribution on the one hand, and a notion of a distribution unchanged upon group transformation on the other. The familiar notions of translation and scaling in the univariate case had natural generalisations in the multivariate case, and it was surprising how many of the univariate distributions had natural generalisations in a multivariate setting; and how easily some of the univariate relationships generalised happily to a wider framework. Even the fundamental ideas of length and direction of a vector found matrix analogues in spherical statistics: see §2.3.4 on p. 35.

From a statistical viewpoint, the most obvious application of the parametric approach advocated by James is to the linear transformation of data matrices, under which there is invariance of the dispersion matrix under the action of orthogonal matrices. The transformation group is the general linear group  $GL(n)$ , acting upon

real symmetric positive semi-definite matrices with the orthogonal group  $O(n)$  as the isotropy group (at the identity matrix).

Another reason for which the statistician may be more interested in zonal polynomials than in Jack polynomials generally is that there is a large number of integral identities and similar relationships satisfied by zonal polynomials and their complex counterparts Schur functions (Jack polynomials with  $\alpha = 2, 1$  respectively), which have no known analogue for general Jack polynomials. The analogues for general  $\alpha$  may exist; but those researching in this area were in the main statisticians and econometricians who were not aware of Jack polynomials, and who were in any case primarily interested in statistical applications.

Nevertheless, placing the study of zonal polynomials within the context of Jack polynomials is a natural thing to do, especially now when Jack polynomials form a current topic of research in physics and combinatorial mathematics, but zonal polynomials are not much studied for their own sake. Stanley (1989, p. 95) has more to say on the advantages of the additional perspective gained by studying zonal polynomials within the framework of Jack polynomials.

As noted, when  $\alpha = 2$ , Jack polynomials reduce to zonal polynomials. For  $\alpha = 1$  the Jack polynomials correspond to the Schur functions, while in the case  $\alpha = 1/2$  they provide zonal polynomials for the skew-field of the quaternions (Macdonald (1995, pp. 440, 446)). In the limiting cases  $\alpha = 0$  and  $\alpha \rightarrow \infty$ , Jack polynomials reduce to multiples of the elementary and monomial symmetric functions  $e_\lambda$  and  $m_\lambda$ , respectively (Stanley (1989, p. 109))<sup>1</sup>.

A representation of a group is a homomorphic image of the group in a *representation* algebra of operators acting upon a *carrier* vector space, say  $U$ . The algebra containing those operators commuting with every operator in the representation algebra is called the *commutant* algebra.

Given a group  $G$  acting on points taken as cosets of an isotropic subgroup, the central items in the representations of  $G$  are the zonal spherical functions, which are

---

<sup>1</sup>The final section of this chapter contains definitions and notation.

chameleonic in character: they constitute one dimensional subspaces of the carrier space which are invariant under the action of the subgroup, and they also span the commutant (algebra), in a sense discussed in §2.5.5.1 on p. 56.

As noted, the object of central mathematical interest is the general linear group acting on real symmetric positive semi-definite matrices, with isotropy subgroup the orthogonal group. The zonal polynomials are derived from the zonal spherical functions of this representation, which is spelt out in some detail in Chapter 2.

### 1.3 Aims of the thesis

While the zonal polynomials introduced in James (1960) were hardly new, being a special case of a class of functions already known, the name (ostensibly) and their application to statistics were new. James himself refers to earlier, rather inaccessible, papers by Hua from the 1950s in which expressions for the zonal polynomials are given, involving integrals “which seem, at least to the author, to be difficult to evaluate.” (James (1961, §1); see also Macdonald (1995, pp. 405, 413) for discussion on the origin of the name zonal polynomial, and for a comment on the unsuitability thereof). I could obtain neither the earlier papers to which James refers (either in Chinese or unpublished), nor a paper in Russian dating from 1959 cited in Constantine (1963) and James (1964). Nor have I managed to procure a copy of Bhanu-Murti (1960), which apparently provides formulae for zonal polynomials.

It is possible that Hua (1963) contains clues as to the content of Hua’s earlier papers, and some of the later parts of that book treat similar problems to those tackled by James, viz. invariance under congruence transforms of matrices and projections onto subspaces, which topics are central to James’ approach discussed in §2.5. Detective work along these lines is beyond the scope of this thesis, and we take James’ approach to zonal polynomials as our starting point.

When introducing the zonal polynomials, James (1960) did not give the impression that he was inventing the name. It may well be that the name was already in use

in the mathematical sources available to him at Princeton, including Hua's work. Nor does James say anything to change that impression in a recent interview, viz. Hillier & Skeels (1996, pp. 162ff).

In essence, and as already noted, zonal polynomials took advantage of the symmetry inherent in statistical/experimental situations, and found exceptionally fruitful application when the underlying distribution was gaussian. The importance of this fact to multivariate statistics follows from the central position of the normal distribution in multivariate statistical theory, in which it occupies an even more pre-eminent position than it does for univariate theory.

The principal use to which zonal polynomials were put was the expression of density functions of many multivariate statistics (quantities calculated from statistical samples) in terms of infinite expansions of zonal polynomials. In fact such densities were hypergeometric functions of matrix argument, which are most readily defined as expansions in zonal polynomials. Familiarity with these hypergeometric functions in turn codified known multivariate distributions and facilitated the definition of new ones: see Takemura (1984, p. 2), Chikuse & Davis (1986), Srivastava (1968), James (1964, p. 475) and Johnson, Kotz & Balakrishnan (1997), *i.a.*

The use of zonal polynomials allowed a natural parametrisation of complicated problems, which were otherwise totally intractable; and just to write down the density functions of complicated multivariate statistics was an achievement. Those expansions in zonal polynomials represent the density as an analytic function, and so could be used on the one hand to provide transforms, mainly the characteristic function and moment generating function<sup>2</sup>, whence the moments could be obtained; and to obtain so-called continuity theorems on the other, whereby convergence of generating functions ensures convergence of corresponding distributions. Moreover the Laplace Transform of a hypergeometric function produces another hypergeometric function, so the generating functions are also expansions in zonal polynomials.

---

<sup>2</sup>The characteristic function in statistics assumes a meaning different from that in the rest of mathematics: the characteristic function here is the Fourier Transform. The moment generating function is effectively the Laplace Transform.

Nevertheless, despite their intrinsic usefulness for the above theoretical purposes, zonal polynomials have by no means realised their full potential. The difficulties of calculating them, or at least applying their calculated values in statistical practice, are apparently severe; and only a handful of papers make any attempt to evaluate them.

This state of affairs raises several questions.

1. Would zonal polynomials be useful to calculate? What are the possible purposes for which one would want to calculate them?

These points are discussed in §2.9.1 on p. 73.

2. If it is desirable to be able to calculate zonal polynomials, are those jobs for which they are potentially useful currently being done by other means? Or are they not being done?

The main alternative tools are asymptotic limits of distributions, and asymptotic expansions of density functions. There is some discussion in §2.3.4.1 on p. 38.

3. If the jobs which could be done by calculating zonal polynomials are being done through alternative means, how effective are those means? Is it the case that those alternative methods are good enough that one does not really need to bother about zonal polynomials in practice?

We do not have much to say on this question, although it is clear that there would still be substantial interest in calculating zonal polynomials if there were reasonable means available for doing so.

4. If the problems with using zonal polynomials are computational in nature, why have those problems not disappeared with the extraordinary increase in computing power since they were defined in 1960?

In part at least, the answer to this lies in §2.9.3.1 on p. 75.

5. Why should zonal polynomials be so hard to calculate? Is it the case that calculating a small number of zonal polynomials is not so hard, but that replicating many such calculations in an effective manner for large collections of

variate values is infeasible? Is the problem not with the zonal polynomials *per se*, but with their utilisation in the hypergeometric expansions?

These points are discussed in §2.9.3.1 on p. 75.

Or is the problem that the zonal polynomials simply involve matrices that quickly grow to gargantuan dimension in practice?

This seems not to be a problem, at least in statistical applications, as discussed in §2.9.4 on p. 77.

The questions were all the more intriguing in view of the ambivalent nature of people's responses to the mention of zonal polynomials. While some but by no means a majority of statisticians had heard of zonal polynomials, few had tried to acquire a workable understanding of them, and as we have mentioned almost no one had used them for any but theoretical purposes.

It appears that while there is interest in calculating zonal polynomials, no-one has managed to do so in an effective fashion. It is in part a consequence of this that zonal polynomials have gone "off the boil" somewhat, in that far fewer papers are written about them now than was the case say twenty years ago. There are still papers written in the statistics and econometrics literature which give density functions of diagnostic statistics in terms of zonal polynomials, or other expansions using extended invariant polynomials (see §2.3.2 on p. 32), but few papers exploring the nature of the zonal polynomials themselves. The mathematical physics literature contains many papers investigating the nature of the Jack polynomials and other relatives of zonal polynomials, but to this author's knowledge the calculation of these polynomials is not attempted in that literature.

In the light of the above questions, then, I determined to search for better means of calculation of, or at least approximation to, the zonal polynomials; there was the further intention to apply those calculations to statistical problems, in finding tail probabilities, likelihood contours, power functions or similar. More conceptually, the aim was to extend the toolkit available to the statistician. So often the attempt to treat a diagnostic statistic analytically is given up at an early stage in favour of the blunt instrument of simulation, and in part at least it was the intention of this

thesis to make available to the practising statistician more powerful analytical tools to complement the simulation approach.

In the event, reality lagged some way behind the lofty intentions. The thesis represents my attempts to understand the nature of zonal and Jack polynomials and the various areas of mathematics underpinning them on the one hand, and to develop enhanced methods of calculation and approximation on the other. Application of these methods to real-life statistical problems remains a project for the future.

## 1.4 What is in the thesis?

The basic tenor of this thesis, firmly rooted in James (1968), is to apply the Laplace-Beltrami operator to vectors of  $m_\lambda$  or  $e_\lambda$  functions and to use the fact that the Jack polynomial is an eigenfunction of this operator with known eigenvalue to expand the polynomial in  $m_\lambda$  and  $e_\lambda$  functions respectively. The same approach could be used for any other set of basis functions for homogeneous symmetric polynomials, but of the commonly used such basis functions only these two produce triangular coefficient matrices. The thesis proves a determinantal expansion for Jack polynomials in elementary symmetric functions  $e_\lambda$  (Roberts (2001)), tying in with a similar result of Lapointe, Lascoux & Morse (2000) for monomial symmetric functions  $m_\lambda$ , and extending earlier work by the author (Roberts (1998)).

Expansion of Jack polynomials in terms of  $m_\lambda$  functions vis-a-vis that in the  $e_\lambda$  function provides a natural counterbalance, partly because of differing but complementary characteristics of the expansions; and partly because algorithms applying to one expansion may have close analogues in the other. Nevertheless there may be advantages in the expansion in one or the other basis function in different situations, and this thesis comes down on the side of the  $e_\lambda$  functions as being perhaps more useful in general. Further discussion on the relative merits and demerits of expanding in the two basis functions is in §8.2 on p. 189.

The thesis starts by reviewing the literature of zonal polynomials, and to some extent

the more recent literature to do with Jack polynomials and related functions. Any such review has to be associated with a survey of some of the theory underlying zonal polynomials and their origin, leading to this chapter being labelled as “Theoretical overview and literature survey”.

It is inevitable that the review is both too broad and too narrow. On the one hand, zonal polynomials and Jack polynomials enter into the literature of many diverse fields, extending over many disciplines, principally but by no means exclusively engineering, physics, pure mathematics and mathematical statistics. In order to convey an idea of how zonal polynomials and related functions are treated in the literature, one has to choose between a broad but shallow choice of subject area; and a narrow spectrum of material covered in greater depth. We have essentially chosen the latter route.

In §2.3.5 on p. 38 fleeting mention will be made of potential uses of zonal polynomials in electronics and communications, treating the work of Smith and Gao. There is a somewhat more extended coverage of the applications of zonal polynomials and their more recently defined cousins the Jack and Macdonald polynomials, as well as their asymmetric siblings, in relatively recent physics literature, mainly in quantum and nuclear physics. Discussion here will centre on work by Forrester and Baker on the one hand, and Lapointe and Vinet on the other: see §2.3.3 on p. 34.

Nevertheless, the principal focus of this thesis is decidedly in the area of mathematical statistics, in line with the author’s background and interests. While the same point as to the choice to be made between a wider shallower discussion and a narrower deeper treatment applies within the field of statistics itself, the coverage of the mathematical and statistical literature of zonal polynomials is intended to be more complete than the forays into electronics and physics mentioned above.

The definition of zonal polynomials in James (1960) drew on the zonal spherical functions of group representation theory, essentially due to E. Cartan in 1929, who extended the Peter-Weyl result to “spherical functions” transforming a set of points (Hillier & Skeels (1996, p. 164)). The label “zonal” refers to functions which are constant on zones, or orbits, of the isotropy group. See also Hannan (1965a, p. 48).

In another sense, however, and especially in the context of mathematical statistics, zonal polynomials have their origin in James (1954), in which James established the differential geometric framework for averaging over the orthogonal group. This allowed the entry of zonal polynomials onto the stage, informally in James (1955a) and James (1955b), in which the density function of the non-central Wishart distribution is expressed as symmetric homogeneous functions not yet labelled as zonal polynomials; and more formally in James (1960), in which the zonal polynomials are introduced in the context of the density function of the eigenvalues of the sample dispersion matrix. The genuine mathematical introduction of zonal polynomials to the statistical world came a little later in James (1961), in which he gave a fairly comprehensive treatment of them through a group representation approach: this is discussed in §2.5.

As noted earlier, much of the thesis is rooted in the work of James (1968), in which a recursive method for generating zonal polynomials through the Laplace-Beltrami operator was set out. The method seems to have been little used in practice, despite McLaren (1976) writing a computer programme to effect the method. This thesis is largely an attempt to refine and develop this method of calculating zonal polynomials, although applied to Jack polynomials.

This is not to say that there have not been developments in the calculation of zonal polynomials and Jack polynomials since 1968. Denoting the Jack polynomials by  $J_{\lambda}^{(\alpha)}$  or sometimes merely by  $J_{\lambda}$ , there are several methods of calculating  $J_{\lambda}^{(\alpha)}$  for particular partitions  $\lambda$  and parameter values  $\alpha$ ; and there are many formulae and integral identities involving zonal polynomials and Schur functions in particular (see Macdonald (1995) and Stanley (1989), for instance). In addition, Knop & Sahi (1997) have given a method of calculating Jack polynomials directly from the Ferrers diagram which obviates the explicit use of  $e_{\lambda}$  or  $m_{\lambda}$  or any other basis functions (for a discussion see §2.2.2 on p. 29); and almost the same method applies to the asymmetric Jack polynomials. Their method is similar to one of the methods of calculating Schur functions in Stanley (1999, §7.10).

Zeilberger has recently written a programme using *MAPLE*, a symbolic manipulation software package, which finds the algebraic expansion of  $J_{\lambda}^{(\alpha)}$  using the creation

operators of Lapointe and Vinet: see §2.3.3 on p. 34. The creation operators essentially generalise the Rodrigues formula for classical orthogonal polynomials, and the program is available publicly on his web site.

This thesis restricts itself to further development of the material in James (1968), then, not because it is the only possible approach to zonal polynomials and Jack polynomials, which it patently is not (see Chapter 2); but because it seems to offer a sensible way to attempt to find approximations to Jack polynomials for practical use. It is by no means clear that the Knop-Sahi method is practical as a means of calculating Jack polynomials: their paper was initially at least notable for proving a conjecture of Macdonald and Stanley about the nature of the coefficients of the expansion of Jack polynomials in terms of  $m_\lambda$  functions (§2.2.2). And given the algebraic expression of  $J_\lambda^{(\alpha)}$  from Zeilberger's program, the polynomial still needs to be evaluated.

There may well be contexts in which Jack polynomials need to be evaluated exactly. But this situation is not so likely in statistical practice, as discussed in §2.9.2 on p. 74. Nor is it likely that the theoretical physicists will see the need to calculate them, as a general matter anyway: they tend to be more concerned about conceptual and theoretical properties of Jack polynomials and similar functions, and the operators of which they are eigenfunctions.

In a statistical context, zonal polynomials arise especially in slowly converging hypergeometric expansions for density functions, but applied statisticians would generally be more interested in (cumulative) distribution functions which give probabilities. One needs to sum slowly converging hypergeometric expansions for density functions, involving larger and larger weights of partitions indexing the zonal polynomials; one must do this over a grid of variate values; one then needs to integrate numerically to get the probabilities. And the starting point of all of this is the calculation of the individual zonal polynomials, which is non-trivial. That said, one is more interested in workable approximations to and bounds on zonal polynomials, and the repetition of those calculations in an efficient manner, than in precise calculation. See §2.9.3 on p. 75.

More complicated related functions of zonal polynomials, called *extended invariant polynomials* by Chikuse & Davis (1986), are defined in the evaluation of integrals arising when exploring more complicated statistical questions, especially in “non-central” hypothesis testing (when the random variates are not assumed to have zero mean); in deriving the densities of econometric system estimators; finding the (cumulative) distribution function of statistics; and obtaining Edgeworth expansions, *i.a.* These extended invariant polynomials share some but not all of the properties of zonal polynomials. Their calculation raises even more problems than that of zonal polynomials, and we shall mention them only in the literature survey, in §2.3.2 on p. 32.

The first step along the path to obtaining workable approximations to Jack polynomials from James (1968) is to place the  $e_\lambda$  functions in a particular order (the RLO: see §1.5 on p. 15) to give a column vector  $E$ ; and similarly we let the column vector  $M$  comprise the  $m_\lambda$  functions. Let  $\mathcal{L}$  denote a generalised Laplace-Beltrami operator, and define “operator” matrices  $\Omega^e$  and  $\Omega^m$  such that

$$\mathcal{L}E = \Omega^e E \quad \text{and} \quad \mathcal{L}M = \Omega^m M \quad (1.1)$$

Following the literature review, most of the remainder of the thesis turns on the matrices  $\Omega$ . There is a close connection between  $\Omega^e$  and  $\Omega^m$ , in that off the diagonal they are effectively transpose to each other, and their diagonal elements are closely related as well. Part of our treatment mirrors that in Roberts (2001), but we give a slightly longer and more elementary proof of the basic results, and flesh out some of the details.

In Chapter 3 we derive a determinantal expansion for Jack polynomials in terms of  $e_\lambda$  functions to complement that for  $m_\lambda$  functions found by Lapointe et al. (2000), our treatment again being based on Roberts (2001). There are strong similarities between these two expansions, arising from the links between  $\Omega^e$  and  $\Omega^m$ .

Approximations to Jack polynomials can basically proceed in three directions:

1. streamlining the determinantal expansions;
2. approximating the  $e_\lambda$  and  $m_\lambda$  functions; or
3. streamlining the algorithm in Knop & Sahi (1997).

The third question is not tackled in this thesis. Broadly speaking, the first approach is explored in Chapters 4 and 5, while the second is investigated in Chapters 6 and 7.

In Chapter 4 we investigate the nature of  $\Omega^e$  and  $\Omega^m$  a little more closely, with a view to making some progress on the first of the above points. That is, we try to gain some idea of how to expedite the determinantal expansions, without undue regard to the calculation of  $e_\lambda$  and  $m_\lambda$  functions. We also examine how sparse the  $\Omega$  matrices are; and find relationships between diagonal sums and column sums of  $\Omega^m$ , which suffice in principle to obtain bounds on coefficients in the expansion of Jack polynomials in  $m_\lambda$  functions, although that task is undertaken to a greater extent in Roberts (1998) for the specific value  $\alpha = 2$ , and only touched upon in this thesis.

Consider the expansions for the Jack polynomials  $J_\rho^{(\alpha)}$  in terms of the  $e_\lambda$  and  $m_\lambda$  functions:

$$J_\rho^{(\alpha)} = \sum_{\lambda} j_{\rho,\lambda}^e e_\lambda = \sum_{\lambda} j_{\rho,\lambda}^m m_\lambda$$

In Chapter 5 we find the coefficients  $j_{\rho,\lambda}^m$  for partitions  $\rho$  and  $\lambda$  assuming a specific generic form; and also for one or two specific values of  $\lambda$  at the final extremity of the RLO, for any  $\rho$ . The generic operator matrix involved is indexed by partitions  $(w), (w-1, 1), \dots$ , and one assumes for the purpose of this thesis that  $w$  is large enough to ensure that the partitions cropping up are listed in the standard non-increasing manner; there is however some evidence to say that the constraints arising on the  $j_{\rho,\lambda}^m$  coefficients from this matrix are valid whether or not the partitions are listed in the standard ordering. Analogous results could be expected to obtain for the  $j_{\rho,\lambda}^e$  coefficients, but that is not attempted in this thesis.

As regards the second question, viz. that of approximating the  $e_\lambda$  and  $m_\lambda$  functions, it is natural to consider such approximations when the variates assume particular forms. The first such examples to spring to mind are the arithmetic and geometric progressions. In Chapter 6 we investigate the behaviour of the  $m_\lambda$  functions when variates assume the successive forms of the arithmetic and geometric progression; while in Chapter 7 we look at the  $e_\lambda$  functions when the variates stand in geometric progression. It seems that both  $e_\lambda$  and  $m_\lambda$  functions are more easily evaluated

when variates stand in geometric progression than when they form an arithmetic progression.

A related question is of course how useful these expansions are, or rather how well the assumed arithmetic and geometric progression “shapes” of the variates conform to the sets of variates arising in practice. One simple point that can be made in this regard is that when random variates are exponentially distributed, the order statistics have expected values which stand in geometric progression (because the distribution function is an “inverted” exponential  $1 - e^{-\mu x}$  for constant  $\mu > 0$ ). Since the exponential distribution crops up often in applied statistical work, the assumption that the variates of Jack polynomials stand in geometric progression may be reasonable in some practical situations.

## 1.5 Notation

### 1.5.1 Ordering of partitions

Let  $\lambda = (l_1, l_2, \dots, l_r)$  be a partition of  $w$  (see Macdonald (1995, §I.1), Stanley (1989, p. 77), *i. a.*). Unless stated otherwise, partitions are assumed to be in standard form, i.e. listed as positive non-increasing integers called elements or parts with no trailing zeroes. Then  $\lambda$  has weight  $w(\lambda) = \sum_j l_j = w$ , and we write  $\lambda \vdash w(\lambda)$ . The length of  $\lambda$  is given by  $\ell(\lambda) = r$ , and the height of  $\lambda$  is  $h(\lambda) = l_1$ . The conjugate partition to  $\lambda$  is denoted by  $\lambda' = (l'_1, l'_2, \dots)$ , and for partitions  $\lambda, \kappa, \tau, \dots$ , it is understood that  $w = w(\lambda) = w(\kappa) = w(\tau) = \dots$  unless otherwise specified.

The multiplicity of  $i$  in  $\lambda$  is  $m_i(\lambda)$ ; further define

$$u_\lambda = \prod m_i(\lambda)! = m_1! m_2! \dots \quad \text{and} \quad z_\lambda = (1^{m_1} 2^{m_2} \dots) m_1! m_2! \dots$$

in accord with notation in Stanley (1989, p. 77), Macdonald (1995, pp. 24, 110).

For  $\lambda = (l_1, l_2, \dots, l_r) \vdash w$ , define the partition  $2\lambda = (2l_1, 2l_2, \dots, 2l_r) \vdash 2w$ .

The dominance or majorisation partial ordering is denoted by  $\leq$ : thus  $\kappa = (k_1, k_2, \dots) \leq \lambda = (l_1, l_2, \dots) \Leftrightarrow k_1 + k_2 + \dots + k_i \leq l_1 + l_2 + \dots + l_i$  for all  $i$ , provided that  $w(\kappa) = w(\lambda)$ . Macdonald (1995) refers to majorisation as the “natural” partial ordering, while Stanley (1989, p. 77) prefers the label “dominance ordering”. The conventional total ordering of partitions, viz. the reverse lexicographic ordering (RLO), is denoted by  $\overset{R}{\leq}$ : hence  $(4) \overset{R}{\geq} (3, 1)$ . The majorisation partial ordering is consistent with the RLO:  $\kappa \geq \lambda \Rightarrow \kappa \overset{R}{\geq} \lambda$ .

For partitions  $\kappa$  and  $\lambda$  not necessarily of the same weight, we define  $\kappa + \lambda = (k_1, k_2, \dots) + (l_1, l_2, \dots) = (k_1 + l_1, k_2 + l_2, \dots)$ ; and the parts of  $\kappa \cup \lambda$  to contain all the parts of  $\kappa$  together with those of  $\lambda$ . According to Macdonald (1995, p. 5), these operations are dual in the sense that

$$(\kappa \cup \lambda)' = \kappa' + \lambda'$$

The  $m_\lambda$  functions are stacked into a column vector  $M$  in RLO, so that the ordering of the indices of the vector elements from the top is  $(w), (w-1, 1), (w-2, 2), (w-2, 1, 1), \dots$ . There is an analogous stacking of the  $e_\lambda$  functions into a column vector  $E$ , and of the  $J_\lambda^{(\alpha)}$  functions into a column vector  $Y$ .

Matrices with rows and columns indexed by partitions have those rows and columns ordered conformally with  $M$  and  $E$ , viz. in RLO. Thus the top left element of a matrix  $A = (a_{\kappa, \lambda})$  is  $a_{(w), (w)}$  while the element to its right is  $a_{(w), (w-1, 1)}$ , etc. With this ordering of partitions, an upper (lower) triangular matrix has zeroes below (above) the diagonal in the conventional manner. The triangular matrices arising in this thesis will generally be unitriangular in the sense of Macdonald (1995, §I.6), who defines  $A$  to be upper (lower) unitriangular when  $a_{\kappa, \lambda}$  can only be non-zero for  $\kappa \geq \lambda$  ( $\kappa \leq \lambda$ ).

Let  $[\kappa, \lambda]$  denote the interval  $\{\tau : \kappa \geq \tau \geq \lambda\}$ . For a matrix  $A$  with rows and columns indexed by partitions in RLO,  $(A)_{[\kappa, \lambda]}$  represents the submatrix indexed by rows and columns in the interval shown, with other elements omitted.

Finally, the transpose of a matrix  $A$  is denoted by  $A^T$ , and the number of partitions of an integer  $n$  is denoted by  $p(n)$ .

### 1.5.2 d2 chains and d2-1 chains

We define  $\kappa \overset{d2}{>} \lambda$  to mean that

1.  $\kappa > \lambda$ ; and
2. in possibly non-standard orderings of the partitions (and perhaps with an additional zero appended to  $\kappa$ ),  $\kappa$  and  $\lambda$  differ in exactly two elements.

We further define  $\kappa \overset{d2}{>}_1 \lambda$  to mean that  $\kappa \overset{d2}{>} \lambda$  and the discrepancy in elements between  $\kappa$  and  $\lambda$  is unity. The analogy with the raising operator in Macdonald (1979, §1) is immediate; and the  $\overset{d2}{>}_1$  relation is the covering relation for the partially ordered set of partitions with majorisation ordering (e.g. Brylawski (1973)).

Thus  $(9, 8, 5, 4) \overset{d2}{>} (8, 7, 7, 4)$  but  $(9, 8, 5, 4) \not\overset{d2}{>}_1 (8, 7, 7, 4)$ ,

while  $(9, 8, 5, 4) \overset{d2}{>}_1 (8, 8, 5, 5)$ .

A d2-chain of partitions from  $\kappa$  to  $\lambda$  of length  $r$  is a chain

$$\kappa = \kappa_0, \kappa_1, \dots, \kappa_{r-1}, \kappa_r = \lambda$$

in which  $\kappa_i \overset{d2}{>} \kappa_{i+1}$  for all  $i$ .

A d2-1 chain is a d2 chain for which  $\kappa_i \overset{d2}{>}_1 \kappa_{i+1}$  for all  $i$ .

In Chapter 4 it is convenient to speak of “chains” of partitions generating the above diagonal entries of  $\Omega^m$ , and “paths” of partitions traversing the operator matrices. But the distinction is a little precious, and the terminology is not necessarily optimal.

### 1.5.3 Augmented $m_\lambda$ functions; $e_\lambda$ and $p_\lambda$ functions etc.

Let  $r = \ell(\lambda)$ . Let  $S_n$  denote the symmetric group on  $n$  symbols; and  $S_n \ni s : i \rightarrow si$  for  $1 \leq i \leq n$ . Following Roberts (1998, §2.1), we define the *doubly augmented*  $m_\lambda$

function

$$dam_{\lambda} = \sum_{s \in S_n} x_{i_{s1}}^{l_1} x_{i_{s2}}^{l_2} \dots x_{i_{sr}}^{l_r} x_{i_{s(r+1)}}^0 x_{i_{s(r+2)}}^0 \dots x_{i_{sn}}^0 ;$$

then the *augmented*  $m_{\lambda}$  function

$$am_{\lambda} = \frac{dam_{\lambda}}{(n-r)!}$$

as in Takemura (1984, §2.2), and which is denoted as  $\tilde{m}_{\lambda}$  in Macdonald (1995, p. 110). Finally,

$$m_{\lambda} = \frac{am_{\lambda}}{u_{\lambda}}$$

is the conventional *monomial symmetric polynomial*, for which the duplicate monomials in  $am_{\lambda}$  have been removed.

The definitions of  $m_{\lambda}$ , the *elementary symmetric polynomial*  $e_{\lambda}$ , the *power sum symmetric polynomial*  $p_{\lambda}$  and the *Schur function*  $s_{\lambda}$  follow Stanley (1989, p. 77) and Macdonald (1995, §§I.2, I.3).

### 1.5.4 Bisymmetric matrices, representations

For a matrix  $X$ , denote the Kronecker product of  $X$  with itself  $t$  times by  $X^{\otimes t}$ . A polynomial in the elements of  $X$  of degree  $t$  may be written uniquely as  $\text{tr}(A^T X^{\otimes t})$ , where  $A$  is a bisymmetric matrix, for which the rows and columns are indexed by  $(i_1, i_2, \dots, i_t)$  and  $(j_1, j_2, \dots, j_t)$  respectively; and  $\text{tr}$  is short for trace. That  $A$  is bisymmetric means that for  $s \in S_t$ ,

$$(A)_{(i_{s1}, i_{s2}, \dots, i_{st}), (j_{s1}, j_{s2}, \dots, j_{st})} = (A)_{(i_1, i_2, \dots, i_t), (j_1, j_2, \dots, j_t)}$$

The algebra of bisymmetric matrices  $A$  is denoted by  $\mathfrak{A}$ . The only case of interest in this thesis is that for which  $X$  is square, for which see Farrell (1985, § 12.1).

A representation of a group is denoted by

$$\mathfrak{B} \triangleright V \triangleleft \mathfrak{C}$$

where  $\mathfrak{B}$  indicates the representation algebra,  $V$  the carrier vector space and  $\mathfrak{C}$  the commutant algebra. The group  $G$  being represented by the matrices in  $\mathfrak{B}$  is

subsumed in the notation, which roughly follows that in James (1961), and is chosen to facilitate the discussion on that paper in §2.5 on p. 49. The terminology “carrier vector space” comes from Littlewood (1950, ch. 1).

### 1.5.5 Note on product of tensors

Given a vector space  $V$  with basis  $\{e_1, e_2, \dots, e_m\}$ , the tensor product  $T^n V$  is a vector space of dimension  $m^n$ , with basis the elements  $e_i \otimes e_j \otimes \dots$ .

For a matrix  $A$  acting on  $V$ ,  $T^n A$  is that matrix transforming basis tensors such as  $e_i \otimes e_j \dots$  without any symmetry requirements imposed on the tensors;  $S^n A$  transforms the tensors assuming that  $e_1 \otimes e_2 \otimes \dots = e_2 \otimes e_1 \otimes \dots$ ; and the exterior power  $\Lambda^n A$  transforms the tensors assuming that  $e_1 \otimes e_2 \otimes \dots = -e_2 \otimes e_1 \otimes \dots$  and  $e_1 \otimes e_1 \otimes \dots = 0$ . The symmetric and exterior powers of  $A$  provide irreducible representations of  $GL(n)$ . More general arrangements, such as symmetrising over the first 2 indices when there are 3 factors, lead to reducible representations. To find the other irreducible representations, symmetrising over the Young’s symmetrisers is required, see *i.a.* Akin, Buchsbaum & Weyman (1982, p. 276), Farrell (1985, ch. 12).

Our notation follows that of Macdonald (1995), and seems to be the most common. Blokhuis & Seidel (1984) give a good review of tensor products, with slightly different notation.

### 1.5.6 Other definitions

We generally use the Jack polynomial  $J_p^{(\alpha)} = J_p$  without undue regard to normalisation, although  $J_p$  is conventionally defined with normalisation given in (5.6) on p. 137. The only part of this thesis in which normalisation of  $J_p$  assumes any importance is in §5.1.1.2 on p. 137.

Operator matrices  $\Omega^m$  and  $\Omega^e$ , as given in (1.1), are sometimes written as  $\Omega_w^m$  and  $\Omega_w^e$  to emphasise the weight of the partitions used. Their constituent elements may be written as  $\omega_{\kappa,\lambda}^m$  or  $\omega^m(\kappa, \lambda)$ , and  $\omega_{\kappa,\lambda}^e$  or  $\omega^e(\kappa, \lambda)$  respectively. In like manner we may write the coefficients  $j_{\kappa,\lambda}^e$  and  $j_{\kappa,\lambda}^m$  introduced in §3.2 as  $j^e(\kappa, \lambda)$  and  $j^m(\kappa, \lambda)$  respectively; and the eigenvalue  $c_\lambda$  first given in §3.1.2 as  $c(\lambda)$ .

For the operator matrix  $\Omega^m$  we define

$$\eta_\lambda = \sum_{\kappa > \lambda} \omega_{\kappa,\lambda}^m.$$

That is,  $\eta_\lambda$  is the sum of terms above the diagonal in the  $\lambda$ th column. An unrelated symbol  $\eta_t$  is introduced in Definition 7.5 on p. 170. In similar vein, the symbol  $\zeta$  is used with two meanings, viz. in (5.16) on p. 141 and in Definition 7.13 on p. 175. But in neither case are the symbols used in any one chapter with two meanings.

$I$  denotes the identity matrix of an appropriate order. The natural numbers are denoted by  $\mathbb{N}$ , and when 0 is omitted by  $\mathbb{N}_+$ ; and the greatest integral part of a non-negative number  $x$  is denoted by  $[x]$ . The rational, real and complex fields are denoted by  $\mathbb{Q}$ ,  $\mathbb{R}$  and  $\mathbb{C}$  respectively, while the ring of integers is denoted by  $\mathbb{Z}$ .

The *Laplace-Beltrami* operator is defined in §3.1.2 on p. 80.

For a subgroup  $H$  of a group  $G$ , by a left coset we mean  $gH$  for  $g \in G$ . This notation follows that in Macdonald (1968), for instance, and stands in contrast to that in Hall (1959) and Hannan (1965a). We set  $G/H = \{gH : g \in G\}$ .

### 1.5.6.1 Acronyms, abbreviations

RLO stands for *Reverse Lexicographic Ordering*.

We write *iff* for *if and only if*.

In like manner, we write *i.a.* for *inter alia*, viz. *among others*.

The *density function* or simply the *density* is written for the *probability density function*.

The *distribution function* is written for the *cumulative distribution function*.

## Chapter 2

# Theoretical overview and literature survey

### 2.1 Farrell's listing of the different approaches to zonal polynomials

Farrell (1985, p. 282) distinguished four approaches to zonal polynomials:

1. his own "algebraic" approach;
2. through the use of group representations;
3. Saw's and Takemura's approach through "quadratic forms"; and
4. treating the zonal polynomials as spherical functions in the sense of Helgason (1962).

Before extending the list to take cognisance of more recent developments, let us clarify these items briefly. Items 1, 2 and 4 will be treated more fully later.

### 2.1.1 Farrell's approach to zonal polynomials

Farrell's approach to zonal polynomials is direct and elegant, but the zonal polynomials at which he is principally aiming are what he, James (1964) and Takemura (1984, Ch. 5) called complex zonal polynomials (Jack polynomials with  $\alpha = 1$ , better known as Schur functions). The term zonal polynomial in current usage is restricted to Jack polynomials with  $\alpha = 2$ , which are the Jack polynomials of principal interest in mathematical statistics, as already noted.

Farrell's method of attack, utilising the centre of the algebra of bisymmetric matrices, quickly produces proofs of integral formulae which are the direct analogues of those holding in the real case. Also produced is a simple formula for Schur functions in terms of group characters. Unfortunately his methodology which works so elegantly for the complex field  $\mathbb{C}$  fails to extend to the real field  $\mathbb{R}$ .

In his work on the real case  $\alpha = 2$ , he essentially applies the same techniques as in James (1961) and Macdonald (1995), and much of Farrell (1985, Ch. 12) is devoted to giving the readers enough understanding of representation theory and related areas for them to understand James' and Constantine's papers. The basic method is to consider the symmetric group and the general linear group acting on tensor powers of a matrix. A more detailed discussion of Farrell's work is in §2.8 on p. 69.

In a late section of the same long chapter, Farrell (1985, § 12.13) also offers an explanation of Saw's and Takemura's approach to zonal polynomials. We discuss Takemura's development of Saw's ideas below in §2.1.3.

### 2.1.2 The approach through group representation theory

By this rather "catch all" title Farrell presumably means the approach taken in James (1961) and to some extent that in James (1964) and Constantine (1963). The first of these papers is discussed at some length in §2.5 on p. 49.

Farrell seems to regard this methodology as “combinatorial” in nature. More precisely, he labels the computational methods in James (1961) and James (1964) as combinatorial (Farrell (1985, p. 231)), although the combinatorial nature of the second of these papers of James is a little obscure. A closely related approach to that in James (1961), but utilising more high powered algebraic tools, is in Macdonald (1995, Ch. VII). This is also discussed later, in §2.6 on p. 64. In any case much of Farrell’s work, especially in Farrell (1985, ch. 12), could be counted as a “representation theory” approach to zonal polynomials. We comment on Farrell’s work in §2.8 on p. 69.

### 2.1.3 Saw’s approach

Farrell (1985) describes this approach to zonal polynomials as the most elementary available. While this may have been true at the time of his writing, the mantle now has surely passed to Knop & Sahi (1997). Nevertheless, Takemura’s adaptation of Saw’s work is relatively simple, especially to those with some acquaintance with the Wishart distribution, and the link with its motivation in mathematical statistics remains clear. Knop & Sahi (1997) certainly provide an elementary definition of Jack polynomials, which is summarised in §2.2.2; but the motivation for their definition is far from clear.

The original article Saw (1977) is not so easy to understand, and Farrell (1985, §12.13) and Mathai, Provost & Hayakawa (1995, p. 171) both discuss Takemura’s elegant adaptation of Saw’s arguments. We present a slightly rough and ready summary of Takemura’s argument, following Takemura (1984, p. 17) and changing his notation a little.

The zonal polynomial  $C_\kappa$  satisfies the reproductive property

$$\mathcal{E}_W [C_\kappa (X^T X)] = a_\kappa C_\kappa(\Sigma) \quad (2.1)$$

in which the matrix  $X$  has the normal distribution,  $X^T X$  has the Wishart distribution with covariance matrix  $\Sigma$ ,  $\mathcal{E}_W = \mathcal{E}$  denotes the expectation operator with respect to this distribution (James (1964, p. 479), Constantine (1963)), and  $a_\kappa$  is a

constant. Thus one is transferring the expectation operator “into” the zonal polynomial: it is as if the zonal polynomial and expectation operators were commuting.

Takemura takes (2.1) as the basis of his definition of zonal polynomials: see also Macdonald (1995, pp. 428, 439). Define  $E'(A)$  to be the vector  $E$  of  $e_\lambda$  functions from (1.1), but taken in inverse order<sup>1</sup> and evaluated at the eigenvalues of the  $k \times k$  symmetric real matrix  $A$ . Define an operator  $\tau_\nu$  acting on  $E'(A)$  by

$$\tau_\nu E'(A) = \mathcal{E} E'(AW) \quad (2.2)$$

in which  $W$  has the Wishart distribution with covariance matrix the identity and degrees of freedom  $\nu$ . The  $p(w) \times p(w)$  matrix  $\tau_\nu$  is upper triangular, where  $p(w)$  is the number of partitions of  $w$ .

The row eigenvectors of  $\tau_\nu$  do not depend on  $\nu$ . Stack these eigenvectors to form an upper triangular matrix  $\Xi$ . That the eigenvectors can be taken in this triangular form follows from an inductive proof: one first shows that there is an eigenvector with but one final non-zero element; then that there is a second eigenvector with but two final non-zero elements etc., provided that the eigenvalues are distinct. For large enough number of variables  $k$  the eigenvalues will be distinct, as Macdonald (1995, p. 439) points out.

Takemura lists a lemma effectively stating that

$$\Xi \tau_\nu = \Lambda_\nu \Xi \quad (2.3)$$

in which  $\Lambda_\nu$  is a diagonal matrix containing the distinct eigenvalues of  $\tau_\nu$  along the diagonal.

Stacking the zonal polynomials into a vector  $Y = (Y_{(w)}, Y_{(w-1,1)}, \dots, Y_{(1^w)})^T$ , and still following Takemura, define the zonal polynomial by  $Y(A) = \Xi E'(A)$ . The symbol  $Y$  is used rather than the conventional  $Z$  or  $C$  because of indeterminate

<sup>1</sup>The ordering is not quite inverse to the RLO, which would be the lexicographic order. Let  $\lambda_1 = (w), \lambda_2 = (w-1, 1), \lambda_3 = (w-2, 2), \dots, \lambda_t = (1^w)$  be the conventional RLO. The ordering in  $E'$  from the top is  $\lambda'_1, \lambda'_2, \dots, \lambda'_t$  which is not the same as lexicographic order, although they coincide for  $w \leq 5$ .

normalisation. Then, letting both sides of (2.3) act upon  $E'(A)$ , development yields

$$\begin{aligned}\Xi\tau_\nu E'(A) &= \Lambda_\nu \Xi E'(A) \\ \Xi \mathcal{E} E'(AW) &= \Lambda_\nu \Xi E'(A) \\ \mathcal{E} \Xi E'(AW) &= \Lambda_\nu \Xi E'(A) \\ \mathcal{E} Y(AW) &= \Lambda_\nu Y(A)\end{aligned}\tag{2.4}$$

thereby verifying (2.1)

The zonal polynomial thus defined is an eigenfunction of an expectation operator with respect to the Wishart distribution. The action of the operator  $\tau_\nu$  resembles those in (1.1) on p. 13, with triangular matrices of coefficients, but with the stacking of the  $e_\lambda$  functions in the  $E$  vector in “inverse” order.

Nor is the advantage of this approach restricted to ensuring the validity of the reproductive property (2.1) and (2.4). What Farrell calls the splitting property of zonal polynomials can also be proved directly from Takemura’s definition:

$$\mathcal{E}_H Y_\lambda(AHBH^T) = Y_\lambda(A)Y_\lambda(B)/Y_\lambda(I)\tag{2.5}$$

where the expectation is over the invariant distribution over the orthogonal group (i.e. the integration is over the Haar measure) (Takemura (1984, p. 27)). Contrast with (2.33) on p. 72.

Davis (1980, p. 293) suggests adapting Takemura’s approach to the more complicated extended invariant polynomials  $C_\phi^{\kappa,\lambda}$  (see §2.3.2), i.e. stacking them and defining  $C_\phi^{\kappa,\lambda}$  as an eigenvector of an appropriate operator. I know of no work developing this idea.

## 2.1.4 Laplace-Beltrami approach

When Farrell described his fourth approach as treating zonal polynomials as spherical functions in the spirit of Helgason (1962), what he presumably had in mind

was the approach using the Laplace-Beltrami operator in James (1968), which drew on material in Helgason's book. James utilised the fact that the zonal polynomial is an eigenfunction of the Laplace-Beltrami operator, with known eigenvalue. This mirrors the fact that the Jack polynomial is an eigenfunction of a generalised Laplace-Beltrami operator, so that one can adopt the same approach to Jack polynomials as James (1968) did for zonal polynomials. This thesis is largely based on the outworkings of this idea, for which see Chapter 3. In fact the generalised Laplace-Beltrami operator is equivalent to the Calogero-Sutherland operator of physics for symmetric homogeneous polynomials (Roberts (2001)), perhaps explaining why so much work on Jack polynomials has been done by physicists. See §2.3.3 on p. 34.

## 2.2 Other approaches to zonal polynomials and Jack polynomials

Since Farrell drew up his list of ways of defining zonal polynomials (§2.1), several other means of defining Jack polynomials have sprung up.

The most important of these, and the line favoured by several recent authors, is to define Jack polynomials by specifying orthogonality conditions to be satisfied. This method is described in §2.2.1. Implicitly the Jack polynomials are eigenfunctions of the quasi Laplace-Beltrami operator, and the inner product separates eigenfunctions with distinct eigenvalues.

Asymmetric Jack polynomials, as well as many other asymmetric polynomials, are defined as eigenfunctions of more complicated operators, generally comprising both differential and permutation elements. Their symmetric cousins can be obtained by similar means using differential operators, or by symmetrisation over the variates of asymmetric polynomials.

The Rodrigues formulae for classical orthogonal polynomials produce polynomials iteratively by applying a differential operator. In like manner, Jack polynomials can also be defined by successive application of creation operators, for which see §2.3.3.

The parallels between the above means of defining Jack polynomials and the classical orthogonal polynomials are strong, with an underlying differential equation, the generation by successive applications of a differential operator (Rodrigues' formulae), generating functions, and orthogonal eigenfunctions with distinct eigenvalues: see for instance the Bateman papers edited by Erdelyi, Magnus, Oberhettinger & Tricomi (1953), Szegő (1939) and Vilenkin (1968, Intro.). From the physics direction, Baker & Forrester (1997) note the close connection between the Calogero-Sutherland model and the generalised classical polynomials: Jack polynomials have a close connection with the Calogero-Sutherland model, for which see §2.3.3.

### 2.2.1 Direct definition from orthogonality

By this is meant the definition of symmetric polynomials by directly specifying orthogonality of functions of differing index or parameter value, with normalisation conditions imposed.

Following Stanley (1989, p. 77) and Macdonald (1995, pp. 305, 377), define an inner product by

$$\langle p_\kappa, p_\lambda \rangle_\alpha = \delta_{\kappa, \lambda} z_\kappa \alpha^{\ell(\kappa)} \quad (2.6)$$

where  $z_\lambda$  is defined in §1.5.1 on p. 15. With this inner product,  $\langle J_\kappa^{(\alpha)}, J_\lambda^{(\alpha)} \rangle_\alpha = 0$  when  $\kappa \neq \lambda$ .

Underlying inner products of this form is generally an operator of which the orthogonal functions are eigenfunctions, with the eigenvalues dictating the normalisation requirements.

Macdonald (1995, §VI.1) defines a general class of symmetric polynomials using a similar definition to that in (2.6), which includes Hall-Littlewood polynomials and Jack polynomials as special cases. Vilenkin & Klimyk (1995, p. 105) basically follow Macdonald's approach.

For the wider class of symmetric polynomials, Macdonald (1995, §VI.2) also shows

the close relationship between the inner products of the form of (2.6), and relations of the form

$$\sum_{\lambda} j_{\lambda}^{-1} J_{\lambda}(x) J_{\lambda}(y) = \prod_{i,j} (1 - x_i y_j)^{-1/\alpha} \quad (2.7)$$

where  $j_{\lambda} = \langle J_{\lambda}, J_{\lambda} \rangle_{\alpha}$ . See also Macdonald (1995, §I.4), Stanley (1989, p. 79) and Takemura (1984, pp. 37, 58).

## 2.2.2 The combinatorial approach of Knop and Sahi

Farrell's labelling of the approaches in James (1961) and James (1964) as combinatorial (Farrell (1985, pp. 231, 282)) is in retrospect a little confusing, in that other methods more directly combinatorial in nature have since sprung up.

The mantle of most elementary approach to Jack polynomials and zonal polynomials now clearly belongs to Knop & Sahi (1997), who have provided a directly constructionist definition of Jack polynomials from the Ferrers diagram. Practically the same definition works for the asymmetric Jack polynomials, also given in the same paper. Although elementary in nature, this method of generation of Jack polynomials does not necessarily lend itself to computing facility; nor does it easily lead to the basic reproductive and splitting formulae (2.1) and (2.5), which were readily obtained from Takemura's definition. We have already commented on the fact that Takemura's definition closely reflects its motivation, whereas the motivation behind Knop and Sahi's approach is obscure.

The Knop-Sahi approach is to utilise the conventional Ferrers graph (of the conjugate partition  $\lambda'$  for  $J_{\lambda}$ ) and simply consider all ways of mapping the  $n$  variables onto the nodes of the graph. In such an overtly combinatorial setting, the mapping could be called a colouring, with colours the integers  $1, 2, \dots, n$ , or rather the variates  $x_1, x_2, \dots, x_n$ . Roughly speaking one multiplies the variates for each legal colouring and sums all the resulting monomials.

There are three rules set up to administer this process. Firstly, while variates can occur multiply in a single colouring, a variate cannot occur twice in any one column;

secondly, if the same variate occurs in successive columns, the one in the first column cannot be above the second (i.e. in a higher row); and thirdly, when the same variate occurs in successive columns and they are in the same row, the monomial product is to be multiplied by a hook length like factor depending both on  $\alpha$  and the arm and leg lengths of the pair of elements in question. The multiplicative factor or weight for each colouring is either 0 for an illegal colouring, 1 if there are no horizontally adjacent cells containing the same variate; or a polynomial in  $\alpha$  derived as a product of the quasi hook length factors. See Stanley (1989, p. 95) or Macdonald (1995, §I.1) for discussion of hook lengths etc.

As a byproduct of their definition of Jack polynomials, Knop and Sahi prove that in the expansion

$$J_{\rho}^{(\alpha)}(x) = \sum j_{\rho,\lambda}^m(\alpha) m_{\lambda}(x)$$

the functions  $j_{\rho,\lambda}^m/u_{\rho}$  are polynomials in  $\alpha$  with non-negative integral coefficients (recall from §1.5 that  $u_{\rho} = \prod m_i(\rho)!$ ). The result confirmed a conjecture of Macdonald and Stanley (Stanley (1989, p. 110)) as to the nature of these coefficients.

Knop and Sahi's method is reminiscent of results in Stanley (1999, §7.10), in which several combinatorial definitions of Schur functions are given; and is further reminiscent of "arm" and "leg" functions, or more generally hook polynomial and hook length types of functions, in Macdonald (1995, ch. I §1, ch. VI §10), Stanley (1989, p. 95) or Aigner (1979, p. 131), *i.a.*; and in the recent physics literature, Baker & Forrester (1999). Robinson (1961, p. 44) has earlier references. These latter types of functions are roughly defined as the product of functions defined on cells of Ferrers diagrams, typically involving the number of cells to the right of the point (an "arm") and below the point (a "leg"). Sometimes one or other of these arm and leg functions at a point is multiplied by  $\alpha$  before adding them together and taking the product over the Ferrers diagram.

Although above we described the motivation underlying the Knop-Sahi algorithm as obscure, it must be admitted that there are elements at least which are somewhat redolent of previous work on Young Tableaux. The forbidding of the one variate to appear twice in a column is the direct analogue of Young's taking the ordinary permutation group for rows but the permutations corrected for sign in the columns

in the Young's symmetrisers; or of the Young Tableaux being only weakly increasing across rows, but strictly increasing down columns. Moreover the Knop-Sahi process bears a certain resemblance to the insertion of one entry at a time onto a Tableau (corresponding to the "lattice permutations") for the calculation of the characters of the symmetric group. See *i.a.* Littlewood (1950, ch. V), Macdonald (1995, §I.9), Farrell (1985, §12.4).

Analysis which is ostensibly straightforward can arise from complicated underlying mathematics. Knop and Sahi's work is a case in point, with seemingly difficult theory by Opdam (1995) in root systems of semi-simple Lie algebras underlying the beautifully simple deductions drawn from the theory. See also Vilenkin & Klimyk (1995, §3.3.1).

## 2.3 Related literature

In this section we glance at some areas of the literature cognate to zonal and Jack polynomials.

The choice of topics is selective and the coverage not deep, but the aim is to obtain at least some idea of these polynomials "in the round" before we embark on a rather technical overview of Hannan's and James' work, *i.a.*.

### 2.3.1 Differential equations

Muirhead (1970) is an early paper investigating the partial differential equations satisfied by hypergeometric functions of one matrix argument. Following Constantine (1963) in defining hypergeometric functions as expansions in zonal polynomials, Muirhead shows that these functions are solutions of a partial differential equation, with a view to using the properties of the equation to give an asymptotic expansion of the function.

There are close links between the hypergeometric functions of matrix argument, the Laplace transform and Laguerre functions of matrix argument (Herz (1955)). There are defined matrix analogues of several other of the classical orthogonal polynomials, generally as hypergeometric functions of matrix arguments with expansions utilising zonal polynomials and Jack polynomials. The author has not seen the extension of the hypergeometric functions of matrix argument to the  $G$  and  $H$  functions of Meijer, which are defined from the univariate hypergeometric functions. But they would be of interest, since Meijer's functions can express the density function of products of independent beta and gamma random variates (Mathai & Saxena (1973)), which is not possible in general for the hypergeometric function. See also Farrell (1985, §1.1).

### 2.3.2 Hypergeometric functions of more than one matrix argument

Given sample size  $n$  and number of parameters  $k$ , James (1955b) found the need to integrate over  $O(k)$  in his derivation of the density function for the Wishart distribution for general covariance matrix. Then, in his derivation of the density of the eigenvalues of the covariance matrix, James (1960) wished to evaluate an integral over  $O(n)$ , when he produced the formula:

$$\int_{O(n)} C_\lambda(AH^T B H) dH = C_\lambda(A) C_\lambda(B) / C_\lambda(I) \quad (2.8)$$

although not in that notation.

Then in Davis (1979) and Davis (1980) it was shown that

$$\int_{O(n)} C_\kappa(AH^T X H) C_\lambda(BH^T Y H) dH = \sum_{\phi \in \kappa, \lambda} C_\phi^{\kappa, \lambda}(A, B) C_\phi^{\kappa, \lambda}(X, Y) / C_\phi(I) \quad (2.9)$$

where the functions  $C_\phi^{\kappa, \lambda}$  are called the *extended invariant polynomials*. The summation in (2.9) is over those representations  $2\phi$  which are contained in  $2\kappa \otimes 2\lambda$ , the Kronecker product of representations. The underlying representation theory for the evaluation of the integral in (2.9) is given in Davis (1980), and this is set out in somewhat greater detail in Mathai et al. (1995, Appendix). The expansion (2.9)

should be compared with (2.8); and with (2.32) and (2.33) on p. 72: the  $Z_\lambda$  and  $C_\lambda$  are different normalisations of the zonal polynomial.

According to Davis (1980), the original motivation for (2.9) came from an application of a rather striking twist to the Edgeworth expansion in Davis (1976). For dealing with a statistic through its Edgeworth expansion, in the latter paper Davis had the idea of treating a random variate  $x$  as  $x' + z$ , where  $x'$  retains the first 2 moments of  $x$  and can be dealt with through the usual Edgeworth expansion; following which one can take expectations with respect to the pseudo random variate  $z$ , which has zero mean and variance but higher moments which duplicate those of  $x$ . The second step is usually combinatorial in nature, and combining the two steps is equivalent to the original outworking of dealing with  $x$  through an Edgeworth expansion.

Davis was led to integrals such as that in (2.9) when seeking to apply the methodology in Davis (1976) to multivariate analysis of variance. In Davis (1980), the author is at pains to point out that the extended invariant functions are not zonal polynomials, in the sense that there may not be an extended Laplace-Beltrami operator for them; and nor may there be suitable differential equations for hypergeometric functions of 2 matrix arguments as there are for hypergeometric functions of 1 matrix argument. Building on work in Muirhead (1982), Baker & Forrester (1997, p. 214) provide a differential equation satisfied by hypergeometric  ${}_2F_1$  functions of two matrix arguments. But it is not clear that this type of differential equation will be as useful as the simpler analogues for functions of one matrix argument.

The right side of (2.9) is clearly a complicated expansion, even by the standards of conventional zonal polynomials of one matrix argument which seem in any case to be so intractable. The complications are even greater than apparent from (2.9) since the direct product of representations  $2\kappa \otimes 2\lambda$  is not multiplicity free: i.e. there are duplications of some of the  $2\phi$  partitions in  $2\kappa \otimes 2\lambda$ . See Robinson (1961, §3.3) for calculation of the multiplicities involved.

The expression (2.9) has been extended to the product of more than two zonal polynomials in the integrand in Davis (1981), who credits Chikuse with prior claim on the extension, in a cited but unpublished paper of hers.

Neither Davis' utilisation of the Edgeworth expansion in Davis (1976), nor the rather theoretical paper on Cornish-Fisher expansions in Hill & Davis (1968), seem to have been taken up to any extent by the statistical or econometric communities; but the identity (2.8) and its generalisations have created wide interest.

### 2.3.3 Jack polynomials in Physics

The pretentious title to this subsection notwithstanding, we have surveyed only a small sample of papers in the area, mainly to do with Baker and Forrester on the one hand, and Lapointe and Vinet on the other. Most of their papers can be found as preprints on the internet, which usually differ but slightly from the published versions:

<http://mentor.lanl.gov:80/Welcome.html>

One should click on *math* for Jack polynomials or *hep-th* for the Calogero-Sutherland model.

Much of the work currently being done on Jack polynomials is by physicists, and much of that work places those polynomials in the context of broader families of functions, which are eigenfunctions of operators involving differential and permutation elements.

As noted in §3.1.1 on p. 79, the Laplace-Beltrami operator is equivalent, for the homogeneous symmetric functions arising in this thesis at least, to the Calogero-Sutherland operator. There is clearly more to be made of the link between statistics and physics, but such work lies beyond the scope of this thesis.

#### 2.3.3.1 Calogero-Sutherland model

The Calogero-Sutherland model in quantum physics refers to many body problems with inverse square potential, but only in one dimension (e.g. Baker & Forrester (1997)). The fact that it is one-dimensional is of course limiting, but treatment at

that level is at the forefront of what the mathematics is capable of at the moment. The model is of interest partly because it is the simplest solvable system of this type, and can provide a vehicle for testing ideas on fractional statistics and quantum chaos: see for instance Forrester (1995) and Lapointe & Vinet (1996); and for an elementary view of Maxwell-Boltzmann statistics and the like, in the context of occupancy types of combinatorial/probability models, see Moran (1968, pp. 32, 105).

The possible states of nature, or the wave functions, are given by the eigenfunctions of the Schrödinger operator, which are closely related to the Jack polynomials. The non-symmetric Jack polynomials have to do with the same system on a circle with exchangeability of particles, whence the permutation operators inserted into the operators to obtain asymmetric polynomials as solutions, *viv-a-vis* the symmetric functions being obtained by differential operators. See Baker & Forrester (1998), *i. a.*

### 2.3.3.2 Creation operators

Higher order Jack polynomials are built up from lower order polynomials by using the creation operators, which generalise the Rodrigues formula for classical orthogonal polynomials. The simplest treatment to follow is perhaps Lapointe & Vinet (1995); one could see Lapointe & Vinet (1997a) and Lapointe & Vinet (1997b) for more comprehensive treatments.

We have already referred in §1.4 on p. 11 to Zeilberger's use of the creation operators to generate the Jack polynomials.

### 2.3.4 Statistics defined on the sphere

Occasionally called *orientation* statistics, *spherical* statistics or *directional* statistics, or similar labels, this area of statistical theory may at first glance appear an undistinguished candidate for discussion in this thesis. But there are several reasons for

which the choice is a suitable one.

1. In §1.3 we posed a series of questions concerning the use of zonal polynomials in statistical theory and practice. Given the central position of orthogonal matrices in the definition of zonal polynomials (for which see §2.5 on p. 49), and the importance of rotational invariance to the theory of random variates defined on spheres, this area could be expected to be fertile ground for seeing how zonal polynomials are used, or not as the case will largely turn out to be.
2. If the use of zonal polynomials in statistical theory is obviated, natural candidates to be used instead are
  - (a) asymptotic limits of the hypergeometric functions defining the densities; and
  - (b) asymptotic expansions of the hypergeometric functions.

For statistics at least, the asymptotic limit intended is that for letting the sample size become indefinitely large.

This area of theory illustrates both of these aspects. That said, the hypergeometric functions generally enter only as normalising constants (e.g. Khatri & Mardia (1977)), and need to be inverted to provide Taylor expansions or asymptotic expansions for densities. Since the hypergeometric functions appear only as normalising constants, inference within the class of directional statistics to some extent need not involve zonal polynomials.

3. There is the analogue of length and magnitude for matrices, generalising the usual definition for vectors. This would seem an important aspect for multivariate analysis, especially in seeking for generalisations of univariate theory.
4. Central to this area of statistics is the idea of factorisation of measures on differential manifolds. This was the original motivation underlying James (1954). By factorising the Haar measure over the parameter space into that associated with the Stiefel or Grassmannian manifold on the one hand and the remainder on the other, one could integrate out the remainder (the nuisance variables) and leave the integral over the subspace of interest. Muirhead (1982, lem. 9.5.3) gives a good indication of the general idea. Recent developments in this direction include Chikuse (1994), which is described below.

Basic references in spherical statistics include Downs (1972) and Bingham (1974). The fundamental distributions are the (matrix) *Langevin* and *Bingham* distributions, with densities proportional to  $\text{etr}(F^T X)$  and  $\text{etr}(X^T A X)$  respectively, in which *etr* stands for the exponential of the trace. The former can be usefully thought of as arising from the term  $\text{etr}\{(X - M)^T(X - M)\}$  appearing in the multivariate normal density, with the constraint imposed that  $X^T X = 1$ . When the mean matrix  $M$  reduces to a vector  $\mu$ , the Langevin distribution is circularly symmetric about the vector  $\mu$ . The connection between the Bingham density and the multivariate normal is even more apparent to the casual observer, and has the advantage of antipodal symmetry, in that  $X$  and  $-X$  have the same distribution (Bingham (1974)).

An idea as to the flavour of the subject may be gleaned from Chikuse (1994). She decomposes the sample space into orthogonal subspaces, say  $A$  and  $B$ , of dimension  $p$  and  $m - p$  respectively; and then considers a subspaces of dimension  $k$  in  $A$  and  $B$ . First factorising the invariant measure for the whole space into those for these subspaces, she then decomposes the data matrix into orthogonal singular value decompositions, and applies this methodology to finding the distribution of canonical correlation coefficients. She is generalising the methodology of James (1954), as noted in her introduction. She is also honing previous work on this problem, obtaining the results even when  $k \leq p$  and  $k \leq m - p$  are not necessarily true.

The polar and elliptic components are defined with respect to a matrix  $C$  in Downs (1972, p. 666). Putting  $C = I$  gives the conventional breakdown. Polar (also known as directional or angular or orientational) and elliptical (length) components are also called by Downs (1972, p. 668) the modal and concentration matrices respectively. Whatever they are called, the analogues of direction and length for an  $n \times k$  matrix  $X$  are  $X(X^T X)^{-1/2}$  and  $(X^T X)^{1/2}$  respectively, for  $n \geq k$ .

Generalising known univariate results, Chikuse investigates the distribution of the orientation of a random matrix, which reduces to a relatively simple form when the matrix is normally distributed: see Chikuse (1990a) and Chikuse (1990b), *i.a.* In another direction, Bingham, Chang & Richards (1992) describe a means of spherical regression due to Chang, and there are many other developments in this whole area.

The field is an interesting one in that there are in the literature approximations obviating the use of zonal polynomials. Those approximations are generally either as limits of a density for large sample size, or as asymptotic expansions of the same, as noted above.

#### **2.3.4.1 Asymptotic expansions**

Farrell (1985, §13.0) mentions the slow convergence of expansions in zonal polynomials, and recommends resorting to asymptotic expansions of hypergeometric functions.

Muirhead (1982, §9.5) gives an asymptotic limit of the  ${}_0F_0$  hypergeometric function as sample size becomes large, tracing the result back to Anderson (1965), who in turn refers to Girshick. Bingham et al. (1992, p. 321) use this method; and Chikuse (1991b) provides asymptotic limiting distributions for statistics in this general area as sample size becomes infinite.

It would seem preferable however to obtain an asymptotic expansion to a density function, or other function of interest, rather than a limiting form. Such expansions generally do not converge, but can provide good approximations for large sample size (e.g. Erdelyi (1956), Whittaker & Watson (1927, ch. VIII)).

Muirhead (1982, §9.5) cites Muirhead (1978) in this regard. Chikuse provides asymptotic expansions for several distributions in this area, for some of which see Chikuse (1990a), Chikuse (1990b), Chikuse (1991a) and Chikuse (1993); see also Kent (1987). Chikuse (1991a) refers to earlier related results of Watson (1983).

#### **2.3.5 An engineering application**

Smith and Gao have done much recent work in the general area of engineering for mobile radios. More precisely, their work has involved the performance of space

diversity wireless systems. The rough idea is that the myriad of signals reflected back from many reflectors to the source of an electromagnetic signal is more or less normally distributed, being the sum of independent small signals. The overall signals are analysed in the complex domain because there are magnitude and phase associated with the electro-magnetic waves. The quadratic forms arise because the strength of the contributions is the square of the magnitude. Density functions of the quadratic forms are hypergeometric functions of one or two matrix arguments, defined in terms of “complex” zonal polynomials.

Two key recent papers are Gao & Smith (2000) and Smith & Gao (2000). The first of these is concerned with the capacity of the mobile radio systems, roughly the number of users which the system can support, and in that paper Gao and Smith give the joint density function of the distinct elements of

$$X\Gamma X^H, \quad (2.10)$$

in a form which is computable, albeit not too straightforwardly. In (2.10),  $\Gamma$  is a positive definite diagonal matrix;  $X$  has a complex Gaussian distribution; and the superscript  $H$  indicates the Hermitian transpose.

The second of these papers concerns the bit error rate of these systems, and here Smith and Gao provide a computable form of the joint density of the distinct elements of

$$Y^H (X\Gamma X^H)^{-1} Y \quad (2.11)$$

in which  $Y$  also has the complex Gaussian distribution, and  $X$  and  $Y$  are independent.

The density functions of (2.10) and (2.11) are given in Khatri (1966) as hypergeometric functions of matrix argument, but Smith and Gao found that they could not easily apply those expansions in practice. Instead they used the results of Gross & Richards (1989), in which hypergeometric functions of matrix argument(s) are expressed as ratios, with a product of Vandermonde determinants in the denominator. The problem of evaluation did not lie with the denominator but with the numerator, involving a determinant with individual elements containing conventional hypergeometric functions having a single parameter as argument. The Gross-Richards expression is valid for distinct values of the eigenvalues, and Smith and Gao adapted

their results to the case with repeated eigenvalues, which involved rather intricate limiting processes.

Gao, Smith & Clark (1998) treats the special case of (2.11) when  $Y$  reduces to a vector, and Smith & Gao (1999) contains further technical details of the outworkings in this case. The latter paper contains a remarkably simple distribution function for the complex random variate (2.11) when the matrix  $Y$  reduces to a vector.

Gao & Smith (1998b) and Gao & Smith (1998a) are generalisations of the work in Gao et al. (1998).

Smith and Gao's work involves the complex zonal polynomials or Schur functions, and Gross and Richards' results are valid for the complex field. If this approach were to generalise to values of  $\alpha \neq 1$ , it would be a very promising approach to the calculation of Jack polynomials.

## 2.4 Technical overview

It is convenient to start the present section with an approach based on Hannan (1965a) (also published as a separate monograph Hannan (1965b)). Although his article is primarily written for an audience interested in probability, and is naturally coloured by his particular interest in time series, Hannan's viewpoint is rather wider than that of other writers, and parts of his article provide a suitable framework in which one can approach zonal polynomials from a general perspective.

### 2.4.1 Group $G$ acting transitively on a set $Z$

To this end, consider a set  $Z$  of distinct elements or points  $z$ , on which a group  $G$  acts transitively by permuting the elements: thus  $gz_1 = z_2$  means that  $g \in G$  moves  $z_1$  to  $z_2$ . Picking out a fixed but arbitrary point  $z_0$ , the isotropy group  $K$  is

that subgroup of  $G$  which fixes  $z_0$ . Changing the reference point produces conjugate subgroups of  $G$  as isotropy groups. Although in principle a change of reference point will change little of substance, there will often be an obvious choice for  $z_0$ , or one for which resulting expressions are simplified in form.

Examples of groups acting as transitive transformation groups in this way include the following:

1.  $Z$  contains the plots in an analysis of variance applied to an agricultural experiment, with  $G$  the symmetric group acting on the plots. This is the example used in James (1957), which can perhaps lay claim to being the first paper to apply representation theory to reflect symmetries inherent in experimental statistics. The point of James (1954), James (1955a) and James (1955b) was more to facilitate the integration over awkward manifolds, particularly the orthogonal group; it was later that the central role of zonal polynomials for this purpose emerged. The invariant relationship matrices in James (1957) reflect functions in a discrete group which are the equivalent of zonal spherical functions for continuous groups.

The ideas in James (1957) were refined in James (1982), which in fact represents an intermediate step in the development of the theory leading to the introduction of zonal polynomials in James (1960) and James (1961), despite its late appearance (see Hillier & Skeels (1996, pp. 156, 163)), Farrell (1985, §12.0)).

2. As a further example in the analysis of variance,  $Z$  contains characteristics of offspring of pairings of brothers and sisters from different families, with  $G$  comprising the symmetric group acting on the families. This example is the one used in James (1982).
3.  $Z$  contains the points on the surface  $S^2$  of the sphere in  $R^3$ . The reference point  $z_0$  is usually taken as the north pole, and  $G$  as the group of rotations of the sphere, viz.  $O(3)$ . The isotropy group  $K$  contains rotations around the axis joining the north and south poles, and the zones are the orbits of the isotropy group (see §2.4.2.5), viz. the bands of constant latitude.

4.  $Z$  consists of the real symmetric matrices. The group  $G$  is the general linear group  $GL(n) \ni L$ , acting by mapping  $Z \ni Y \rightarrow LYL^T$ . The reference point  $z_0$  is logically chosen as the identity matrix, with isotropy group the orthogonal group  $O(n)$ .

This example provides the mathematical foundation of zonal polynomials in James (1961). See §2.5.

5.  $Z$  is the set of “doublets” in James (1961, §3), a doublet being defined to be a subdivision of the integers  $\{1, 2, \dots, 2f\}$  into mutually exclusive pairs. The group is the symmetric group on  $2f$  elements acting in the conventional manner on the set of doublets. The isotropy group at the element  $(1, 2)(3, 4) \dots (2f - 1, 2f)$  is the hyperoctahedral group (Macdonald (1995, p. 401)).

This example provides a further central element in the definition of zonal polynomials, since the hyperoctahedral group imposes symmetries on a general tensor to make it correspond to a bisymmetric tensor deriving from a symmetric matrix (James (1961, §2); see also §2.5.6.4 on p. 60).

### 2.4.2 Representations of the group $G$

While it is  $Z$  that is of primary interest, to remain at the level of  $G$  acting on  $Z$  renders analysis difficult, since in general  $Z$  has no particular structure. It is natural to remove one stage to consider the points  $z$  as left cosets of the isotropy group and analyse the coset space  $G/K = \{gK : g \in G\}$ .

In fact we choose to work with representations of the group  $G$ . These are operators permuting vectors in a “carrier” vector space, and are homomorphisms  $T(g) = T_g$  of  $G$ : i.e. they mimic the behaviour of the group, in that  $T(g_1)T(g_2) = T(g_1g_2)$ .

Working with general representations of  $G$  can focus more directly on the symmetries inherent in the original problem. In addition, representation theory is a well researched area of mathematics about which much is known, so that one is tapping into a large “pool” of knowledge. In any case one loses nothing by the additional generality, since the situation in which  $G$  operates directly on  $G/K$  is a particular

example contained within the general theory.

For our purposes, representations will in general take the form of either permutation matrices, as in examples 1 and 2 above; or will act by translating functions of  $Z$ , i.e. the carrier space is the space of functions on  $Z$  of a certain form, as in Example 4 on p. 41 above. Symbolically one has in this latter case

$$(T_1(g)\phi)(z) = \phi(g^{-1}z)$$

representing the group by left “translating” a function. Similarly another representation is given by a right translation

$$(T_2(g)\phi)(z) = \phi(zg).$$

Considering the translation of a function is a well trodden path; see *i.a.* Weyl (1946, p. 23), Farrell (1976, p. 235), Vilenkin (1968, p. 12), Helgason (1962). Translation in this sense is sometimes referred to as a shift or shift operation, e.g. Vilenkin (1968, p. 27).

Note that the multiplication in a group algebra is effectively translation of a function: for treating an element in the group algebra as a function on the group, and for  $x, y \in G$  (so that  $x(g) = 0$  for  $g \neq x$  and 1 for  $g = x$ ), then  $(xf)(y) = \sum_{g \in G} x(g)f(g^{-1}y) = f(x^{-1}y)$  (e.g. Macdonald (1995, p. 388)).

Given that we are representing points  $z$  as left cosets  $gK$ , it is natural to consider functions  $f(g)$  invariant on the right: i.e.  $f(g) = f(gk)$ . The zonal polynomials will surface as functions which are invariant from either side, viz. they are bi-invariant, although this bald assertion gives little idea of their significance.

#### 2.4.2.1 The commutant algebra of a representation

The representation algebra of a representation  $T$  is generated by all matrices  $T_g$ ,  $g \in G$ , in the representation. The commutant algebra of  $T$  contains all matrices  $B$  which commute with all matrices in the representation algebra.

There are several other names for the commutant algebra, including commuting algebra (Hannan (1965a), Robinson (1961)), commuting ring (Burrow (1965)), centraliser (James (1961), James (1982)), central (Weyl (1931)), and commutator algebra (*i.a.* Weyl (1946), Farrell (1985)). The use of the last of these names is reasonably widespread, but it could be confused with the commutator in a group context, viz. an element of the form  $x^{-1}y^{-1}xy$ .

The most common label now seems to be the commutant (algebra) (e.g. Ledermann (1977)), to which terminology we shall adhere. It should be noted however that our usage clashes to some extent with older conventions. Turnbull & Aitken (1932, p. 147), for instance, call a matrix  $X$  a commutant of  $A$  and  $B$  when  $AX = XB$ .

The commutant is no less important than the representation algebra itself. This is partly because the commutant is isomorphic to the endomorphism ring of the carrier space, which provides information about the multiplicity of irreducible component representations in the representation  $T$ .

More centrally, however, the commutant reflects closely the symmetries inherent in the situation under investigation. For the representation R1 in (2.13) on p. 50, it will turn out that the zonal spherical function, properly belonging to the carrier space, can usefully be regarded also as a member of the commutant: in fact the zonal spherical functions collectively form an orthogonal basis of the commutant. More precisely, the commutant is spanned by matrices which are zero outside a single diagonal block, and that block is a zonal spherical function times a unit matrix. From the zonal spherical function for R1 is directly derived the zonal polynomial: see §2.5.5.1.1.

The relationship matrices in James (1957) span the commutant, and are the direct analogues of zonal spherical functions, for the finite group defined in that simple analysis of variance setting. That the mathematical introduction of zonal polynomials in James (1961) is perhaps a little hard to follow is partly because the jump from James (1957) to James (1961) depended on theory which was not published until James (1982).

### 2.4.2.2 The endomorphism ring

The isomorphism between endomorphisms of the carrier vector space and the commutant algebra of the representation can be seen as follows.

Recall that a homomorphism of a left module  $U$  into a left module  $W$  is a mapping  $\phi : U \rightarrow W$  such that

$$\phi(au) = a\phi(u) \quad (2.12)$$

where  $a$  acts on  $u \in U$  (resp.  $w \in W$ ) to produce a member of  $U$  (resp.  $W$ ). For a vector space (or linear space), a special case of a module,  $u$  is a vector and  $a$  is a matrix operating on the vector space in the conventional way: in the simplest case  $a$  is simply an element of the coefficient field multiplying the vector. Setting  $U = W$ , the homomorphism is relabelled as an endomorphism. See *i.a.* Ledermann (1977, p. 33), who is more careful in his use of language and refers to the above transforms (2.12) as  $G$ -homomorphisms (when  $a \in G$ ).

Suppose now that  $T$  is a representation of a finite group  $G$  with carrier space  $U$  and let  $a$  be in the representation algebra, i.e.  $a = \sum b_g T_g$  is a representation of  $\sum b_g g$  in the group algebra; then  $\phi$  in (2.12) is in the commutant algebra. In fact, the commutant algebra and the endomorphism ring of the carrier space are isomorphic.

More generally, when  $W$  is an irreducible representation of  $G$ , the dimension of  $\text{Hom}(U, W)$ , called the intertwining number, represents the number of times that  $W$  appears in  $U$ .

### 2.4.2.3 The commutant as reflecting symmetries

In the context of working with the symmetric group acting on  $Z$  (or more strictly with the representation of the symmetric group induced by the trivial representation of the isotropy group  $K$ , so that the permutations merely shuffle the left cosets of  $K$  around), Hannan (1965a, p. 14) points out the equivalence of working with

1. the commutant algebra;
2. point pair invariant functions, defined below; and
3. functions having “rotational symmetry” on  $Z$ . These are functions  $f(z)$  for which  $f(kz) = f(z)$  for all  $k \in K$ , and are called zonal functions, because they are constant on the zones  $Kz$ .

Representing a point  $z \in Z$  as a left coset of an isotropy subgroup in  $G$ , a function on  $Z$  must be a function of  $G$  which is invariant on the right: i.e.  $\phi(gk) = \phi(g)$  for all  $k \in K$ . These functions are called spherical functions. Functions of  $G$  having “rotational symmetry” now are biinvariant, and are called zonal spherical functions:  $\phi(k_1 g k_2) = \phi(g)$  for all  $k_1, k_2 \in K$ .

The terminology can be traced to the case in which  $Z$  contains the points on the surface of the sphere, and zones are bands of constant latitude (Hillier & Skeels (1996, p. 164), Hannan (1965a, §5.1), Vilenkin (1968, p. 30).

We show now the connection between these three classes of objects.

#### 2.4.2.4 Point pair invariant functions

Define an equivalence relation on the direct product  $Z \times Z$  by saying  $(r, s) \sim (u, v)$  if there exists  $g \in G$  such that  $u = gr$ ,  $v = gs$ . That is, the equivalence classes are  $\{(gr, gs) : g \in G\}$ , where  $(r, s)$  is arbitrary but fixed. Such equivalence classes are called *point pair invariants* in Hannan (1965a, p. 13), while James (1961, p. 464) gives them no label. Functions defined on  $Z \times Z$  which are constant on these equivalence classes are called *point pair invariant functions*.

Given the equivalence class generated by  $(r, s)$  as above, define a matrix  $\Gamma(r, s) = (\gamma_{ij})$  by

$$\gamma_{ij} = \begin{cases} 1 & \text{if } (i, j) \sim (r, s) \\ 0 & \text{otherwise} \end{cases}$$

When  $T(g)$  is a permutation of  $Z$ , that the matrices  $\Gamma(r, s)$  belong to the commutant is easily seen as follows:

$$(\Gamma(r, s)T(g))_{z_2, z_0} = \sum (\Gamma(r, s))_{z_2, z_1} (T(g))_{z_1, z_0} = (\Gamma(r, s))_{z_2, gz_0}$$

$$(T(g)\Gamma(r, s))_{z_2, z_0} = \sum (T(g))_{z_2, z_1} (\Gamma(r, s))_{z_1, z_0} = (\Gamma(r, s))_{g^{-1}z_2, z_0} = (\Gamma(r, s))_{z_2, gz_0}$$

In fact, the matrices  $\Gamma(r, s)$  generate the commutant, and correspond directly to the zonal spherical functions for continuous groups. That they generate the commutant can be seen informally as follows.

As a finite set being transformed,  $Z$  can be considered a vector space of dimension  $|Z|$ . For representations of a continuous group,  $Z$  would be considered as a basis of the carrier space.

The linear transformations of  $Z$  are matrices indexed by the elements  $z_i$  of  $Z$ . If  $z_i^*$  denotes a basis for the dual space, the matrix with  $(i, j)$ th entry  $z_i z_j^*$  is the general linear transformation of  $Z$ . Simplifying by indexing this matrix as  $(z_i, z_j)$  and ensuring that  $(z_i, z_j)$  and  $(gz_i, gz_j)$  always have the same value for all  $g \in G$ , informally at least one can see that any operation by a group element and multiplication by this matrix must commute. See also James (1961, Lemma 5).

The simplest example of the matrices  $\Gamma$  is the invariant relationship matrices in James (1957). He defines  $B$ , for instance, in which  $(B)_{ij} = 1$  when plots  $i$  and  $j$  are in the same block, and 0 otherwise; and similarly for the other matrices ( $T$  for when plots are subject to the same treatment,  $G$  for the grand mean and  $I$  the identity). These matrices generate what James calls the invariant relationship algebra, which he recognised as spanning the commutant. The carrier “vector space” in his paper contains linear combinations of the plots.

James (1957) seems not to have found much application in statistical practice. Mann (1960) is one of the few papers taking up James’ ideas, applying them in the context of testing nested hypotheses in a linear model which is a generalisation of the analysis of variance. Even James (1982), a paper essentially bridging the gap between James (1957) and James (1961), cites no papers deriving from its 1957 parent.

### 2.4.2.5 Zones and left invariance of functions

The relation between point pair invariant functions and zonal functions is immediate. We follow Hannan (1965a, p. 13).

Given the point pair invariant function  $\gamma(r, s)$ , set  $f(z) = \gamma(z_0, z)$  for some fixed  $z_0 \in Z$ . Then for  $k \in K$ , the isotropy group for  $z_0$ , one has that  $f(kz) = \gamma(z_0, kz) = \gamma(z_0, z) = f(z)$ , so that  $f(\cdot)$  is a zonal function. Alternatively, given the zonal function  $f(z)$ , so that  $f(kz) = f(z)$ , we set  $\gamma(z_0, z) = f(z)$  and  $\gamma(r, s) = \gamma(gr, gs)$  where  $gr = z_0$ .

### 2.4.2.6 Projection operators onto the isotropy group

Instead of having the transformation group  $G$  act from the left on the left cosets  $gK$ , it is often more convenient to consider the average over the isotropy group within the group algebra of  $G$ :

$$e_2 = \frac{1}{|K|} \sum_{k \in K} k$$

The operator  $e_2$ , projecting the group algebra into the subalgebra spanned by  $K$ , is generally labelled as a variant of  $e$ , and is called  $e_2$  here simply to avoid confusion with  $e$  and  $e_1$  which assume specific meanings later.

Thus the action of  $G$  on the left cosets  $G/K$ :

$$g_1 : gK \rightarrow (g_1g)K = g_1gK$$

is mirrored by the action of  $G$  within the group algebra:

$$g_1 : ge_2 \rightarrow (g_1g)e_2 = g_1ge_2.$$

James (1961) defines  $e_1$  as the projection operator for the isotropy group of the symmetric group acting on the doublets in Example 5, and  $e$  as the projection operator for the orthogonal group in Example 4 on p. 41, this being the isotropy group for the general linear group in this situation (see (2.22)). This latter operator  $e$  is an integral, being the equivalent operation on a continuous group to the average

over a finite group. In similar vein Macdonald (1995, p. 388) defines a simple average over the isotropy group  $K$  when the underlying group of transformations is finite.

## 2.5 James' definition of zonal polynomials

At first glance the approach to zonal polynomials in James (1961) looks convoluted. The reasoning is intricate, revolving around coincidence of commutants for chains of group representations. It seems worthwhile to explore this paper in some detail, partly because the article contains the basic definitional mathematics for the main topic of interest in this thesis; and partly because James' treatment can serve as a vehicle for comment on the other principal approaches to zonal polynomials.

If James (1961) is a little hard to follow, it is partly because the theory supporting the move from James (1957) to James (1961) only appeared much later, in James (1982). And admittedly James (1961) is easier to appreciate now that Macdonald (1995) has given such a scholarly treatment of the essential ideas therein, on which we comment in §2.6.

As a preliminary guide to nomenclature to come, there is a basic and large diagram in James (1961, p. 463) which summarises the chain of reasoning of a major part of his article: this is roughly reproduced below as the central portion of the figure in (2.13), with some additions to the figure and slight changes in notation. There are 6 distinct representations appearing in the central column, labelled as  $R_1, R_2, \dots, R_6$ . The basic situation which James is investigating is on the second line, for instance, and is listed as  $R_2$ . Basic notation is given in §1.5 on p. 15, and descriptions follow.

$$\begin{array}{rclcl}
 R7 & \mathfrak{A} \triangleright \mathfrak{A} \triangleleft \mathfrak{A} & \rightarrow & R1 & \mathfrak{A} \triangleright \mathfrak{A}e \triangleleft e\mathfrak{A}e & \leftarrow R1R & \mathfrak{A}' \triangleright \mathfrak{A}e \\
 & & & R2 & \mathfrak{A}_1 \triangleright V_f \triangleleft \mathfrak{B} & & \\
 R0a & \mathfrak{A} \triangleright P_{2f} \triangleleft S_{2f} & \rightarrow & R3 & \mathfrak{A} \triangleright e_1 P_{2f} \triangleleft e_1 S_{2f} e_1 & & \\
 \hline
 R8 & S_{2f} \triangleright S_{2f} \triangleleft S_{2f} & \rightarrow & R4 & S_{2f} \triangleright S_{2f} e_1 \triangleleft e_1 S_{2f} e_1 & & \\
 & & & R5 & S_{2f} \triangleright \mathfrak{D} \triangleleft \mathfrak{F} & & \\
 R0b & S_{2f} \triangleright P_{2f} \triangleleft \mathfrak{A} & \rightarrow & R6 & S_{2f} \triangleright e P_{2f} \triangleleft e \mathfrak{A} e & & 
 \end{array} \tag{2.13}$$

The three elements within each item denoting a representation are respectively the representation algebra, the carrier space and the commutant algebra. There is sometimes the implication that the first of these acts from the left and the last from the right, as in the regular representations R7 and R8.

The representation algebra for most of the representations above the line in (2.13) is  $\mathfrak{A}$ , consisting of bisymmetric tensors  $A$  of rank  $4f$ , with  $2f$  covariant (lower, “row”) and  $2f$  contravariant (upper, “column”) indices: see *i.a.* Spiegel (1959) for the definitions of covariant and contravariant tensors. These representations are representations of  $GL(n)$ , since  $\mathfrak{A}$  is the enveloping algebra of  $L^{\otimes 2f}$ , for  $L \in GL(n)$ . The exception is R1R, which furnishes a representation of  $O(n)$ .

Below the line are representations of  $S_{2f}$ , the symmetric group on  $2f$  symbols. The representation algebra is the group algebra of the symmetric group, denoted by the same symbol in an abuse of notation.

The more general representations underpinning James’ diagram but not appearing in the original are four in number, and listed on the left side of (2.13). Two of them are listed as R0a and R0b, since they are identical save for the interchange of their representation algebras and their commutants: these representations interrelate  $\mathfrak{A}$  on the one hand, and  $S_{2f}$  on the other. The other basic underlying representations are labelled as R7 and R8: the former of these is the regular representation of  $\mathfrak{A}$ , and the latter is the regular representation of  $S_{2f}$ . The arrows from the left-most column to the middle column indicate that the carrier space has been restricted, using the projection operators  $e$  and  $e_1$  defined in (2.22) and (2.26) on pp. 56 and 61 respectively. Further details of the overall structure of the diagram (2.13) are in

### §2.5.3.

The carrier space for the regular representation R7 is also shown as  $\mathfrak{A}$ , but is capable of a dual and more revealing interpretation. A polynomial in the elements of a matrix  $X$  can be written uniquely as  $\text{tr}(AX^{\otimes 2f})$  where  $A$  is bisymmetric, so that the carrier space of R7 may be considered either as  $\mathfrak{A}$ ; or as consisting of polynomials in the elements of a matrix  $X$ . That is,

$$R7 \quad \mathfrak{A} \triangleright \mathfrak{A} \triangleleft \mathfrak{A} \quad \leftrightarrow \quad \mathfrak{A} \triangleright \text{tr}(\mathfrak{A}X^{\otimes 2f}) \triangleleft \mathfrak{A} \quad (2.14)$$

## 2.5.1 The situation of interest

The basic situation James wishes to analyse is that of Example 4 on p. 41: viz. the carrier space  $V$  consists of functions of real positive definite symmetric matrices  $Y$ , and the group action on  $V$  is given by  $GL(n, R) = G \ni L : \phi(Y) \rightarrow \phi(L^{-1}YL^{-1T})$ . This representation, labelled as R2, is of basic interest in statistics for the following reason.

Consider an  $n \times k$  data matrix, where  $k$  is the number of variates of interest, and  $n$  is the sample size; it will usually be the case that  $n > k$ . Denoting the covariance matrix of  $X$  by  $\Sigma$ , that of  $L^T X$  is  $L^T \Sigma L$ . Disregarding the adjustment required to allow for the non-zero sample mean, the covariance matrix  $\Sigma$  is estimated by  $X^T X/n$ , which is unchanged by data transformations  $X \rightarrow HX$ , where  $H$  is an orthogonal matrix. The adjustment for the sample mean is standard, see *i.a.* the discussion in Mardia, Kent & Bibby (1979, p. 10) on the “centring” matrix.

## 2.5.2 Overview of James’ reasoning

It will turn out that R1, R2 and R3 are equivalent representations; likewise R4, R5 and R6 are equivalent.

Choosing the second interpretation of the carrier space of R7 in (2.14), the carrier

space of R1 contains polynomials of the form  $\text{tr}(AeX^{\otimes 2f})$ , in which the projection operator  $e$  is also a bisymmetric matrix. These are polynomials in the elements of  $X$ , but assume the further specific form of being functions of  $X^T X$ . This shows the link between R1 and R2, the carrier space for the latter being symmetric real matrices.

The zonal spherical functions (ZSFs) for R1 span the commutant  $e\mathfrak{A}e$ . The zonal polynomial ( $ZP$ ) is defined as

$$ZSF(X) = ZP(X^T X)$$

in an obvious if unimaginative notation. The zonal polynomials thus span the commutant  $\mathfrak{B}$  of R2. James (1961) identifies the commutants of the first three representations with those of the second three, and investigates  $\mathfrak{B}$  by finding  $\mathfrak{F}$ . This is the intricate part, since R5 involves the doublets mentioned in Example 5 on p. 42, which caused Farrell to label James' approach as combinatorial.

### 2.5.3 Description of the diagram (2.13)

The restriction from R1 to R1R is indeed a "restriction" of a representation in the conventional terminology: the representation algebra is restricted to that representing a subgroup of the group previously represented, and is the opposite process of the third meaning of induction discussed at the end of §2.5.4 on p. 54.  $\mathfrak{A}'$  is that subalgebra of  $\mathfrak{A}$  generated by orthogonal matrices, as shown in (2.23). The carrier space is unchanged.

In the major part of the diagram (2.13), however, we are not defining restrictions of representations. In projecting the carrier space  $V$  onto a subspace of itself, say  $V_0 = e_2 V$  or  $V e_2$ , we are in contrast taking a component of the original representation: i.e. we are stripping away those parts of the matrices in the previous representation algebra which operate on subspaces orthogonal to  $V_0$ , to leave smaller matrices operating on a carrier space  $V_0$  of reduced dimension. The commutant becomes  $e_2 V e_2$ , which is commutative in the examples above (for R1, R3, R4 and R6). That the commutant is commutative means that the representation with carrier space  $V_0$

is multiplicity free (see Curtis & Reiner (1962, pp. 319, 340) or Ledermann (1977, p. 27), *i.a.*).

The projection operators in the diagram (2.13) are defined below: the projection  $e$  onto the orthogonal group is defined in (2.22); and the projection  $e_1$  onto the hyperoctahedral group is described in §2.5.6.4.

The representation algebra  $\mathfrak{A}_1$  of  $R_2$  is spanned by the matrices  $\Lambda$  arising in §2.5.4. The carrier space  $\mathfrak{D}$  in  $R_5$  is the set of functions on  $D$ , the set of doublets, defined in Example 5 on p. 42. These are described later.

Although not especially emphasised in James (1961), the symmetry in the action of  $GL(n)$  and  $S_{2f}$  is apparent from the diagram (2.13). The dual role of the actions of these two groups is even more apparent from Macdonald (1995, Ch. VII) and Farrell (1985, Ch. 12).

## 2.5.4 The basic representation of interest $R_2$

A representation of  $GL(n, R) = G$  is induced to act on the vector space of polynomials of degree  $f$  in the elements of a general symmetric matrix  $Y$ , say  $\phi(Y)$ , by the congruence transformation  $Y \rightarrow LYL^T$ . The representation is given formally as

$$(R2) \quad \phi \rightarrow L\phi \quad (2.15)$$

and is defined by

$$(L\phi)(Y) = L\phi(Y) = \phi(L^{-1}YL^{-1T}) \quad (2.16)$$

The notation  $L\phi$  for the result of  $L \in G$  acting on the polynomial  $\phi$ , although conventional, is rather misleading, since the matrix multiplying an appropriate vector of polynomials is very much larger than the original matrix  $L$ . Choosing a basis of monomials in the elements of  $Y$ ,  $\phi$  becomes a vector, say  $\varphi$ ; and similarly denote the operating matrix by  $\Lambda$ , so that the representation (2.15) becomes:

$$(R2) \quad \varphi \rightarrow \Lambda\varphi \quad \text{or} \quad L \rightarrow \Lambda$$

The matrices  $\Lambda$ , as  $L$  varies over  $GL(n)$ , span the algebra  $\mathfrak{A}_1$ , a subalgebra of the bisymmetric tensors  $\mathfrak{A}$ . The dimensions of these algebras are compared in (2.17) and (2.18).

Recall that the number of ways of choosing  $m$  elements from  $n$  is  $\binom{m+n-1}{m}$ , when repeated elements are allowed (e.g. Stanley (1986, p. 15)). Further, let  $L^{[j]}$  be the  $j$ th induced matrix of  $L$ . If one considers the transformation  $X = AY$  for  $n$ -vectors  $X$  and  $Y$  and  $n \times n$  matrix  $A$ , then the homogeneous products of  $j$  elements of  $Y$  are transformed into the homogeneous products of  $j$  elements of  $X$  by the  $t \times t$  matrix  $A^{[j]}$ , where  $t = \binom{n+j-1}{j}$ . The terminology and notation is that of Littlewood (1950, p. 178).

The form of  $\Lambda$  is known to be  $(L^{[2]})^{[f]}$  (James (1961), Macdonald (1995, p. 160)). The latter of these sources cited gives the relationship in more modern notation:  $\Lambda = S^f(S^2(L))$  (see *i.a.* Blokhuis & Seidel (1984), who use  $\Sigma_t$  for  $S^t$ , for an explanation of the nature of these matrices). Thus, if  $Y$  and  $L$  be  $n \times n$  matrices, and setting

$$r = \binom{n+1}{2} \quad \text{and} \quad s = \binom{r+f-1}{f} = \frac{n^{2f}}{2^f f!} \left(1 + \frac{f}{n} + \dots\right) \quad (2.17)$$

the vector  $\varphi$  is  $s \times 1$  and the matrix  $\Lambda$  is  $s \times s$ .

It is perhaps worth commenting on the various uses of the word induction here. The first is the representation of  $G$  sending  $L \rightarrow \Lambda$ ; and the second is the matrix  $L$  inducing the matrix  $L^{[j]}$  for the symmetric tensor representation. Another general meaning is the representation on a group induced by a representation on a subgroup (e.g. Ledermann (1977, Ch. 3), Vilenkin (1968, p. 31)). The principal use of the latter example in this thesis is when the representation on the isotropy subgroup is trivial, so that the induced representation permutes the left cosets (representing the points of  $Z$ : see §2.4.1 on p. 40).

### 2.5.5 The representation R7 leading to R1

Let  $X$  be a  $n \times n$  matrix. The dimension of the matrices in the representation algebra of R7 is  $t \times t$ , for

$$t = \binom{n+2f-1}{2f} = \frac{n^{2f}}{(2f)!} \left( 1 + \frac{f(2f-1)}{n} + \dots \right) \quad (2.18)$$

It is easy to show that  $t/s > 2^f \times f! / (2f)!$ , where  $s$  is taken from (2.17); in fact this lower bound is weak unless  $n \gg f$ .

The typical polynomial in the carrier space of R7 is (using the summation convention, and restricting ourselves to  $2f = 2$  indices of each type in order to illustrate)

$$a_{i_1 i_2}^{j_1 j_2} x_{j_1}^{i_1} x_{j_2}^{i_2} \quad (2.19)$$

in which the tensor  $a$  is bisymmetric.

$GL(n) \ni L$  acts on (2.19), from the left say, moving it to the following polynomial in the carrier space

$$(L^{-1})_{i_1}^{k_1} (L^{-1})_{i_2}^{k_2} a_{k_1 k_2}^{j_1 j_2} x_{j_1}^{i_1} x_{j_2}^{i_2}$$

Alternatively,  $L$  may act from the right, moving (2.19) to

$$a_{i_1 i_2}^{j_1 j_2} x_{j_1}^{k_1} x_{j_2}^{k_2} (L)_{k_1}^{i_1} (L)_{k_2}^{i_2}$$

Equivalently, one could operate only on the bisymmetric tensors (retaining the representation algebra but simplifying the carrier space, so that an equivalent representation obtains), representing the polynomial (2.19) as

$$a_{i_1 i_2}^{j_1 j_2} \quad (2.20)$$

From the left  $L$  moves (2.20) to

$$(L^{-1})_{i_1}^{k_1} (L^{-1})_{i_2}^{k_2} a_{k_1 k_2}^{j_1 j_2}$$

and from the right  $L$  moves (2.20) to

$$a_{i_1 i_2}^{k_1 k_2} (L)_{k_1}^{j_1} (L)_{k_2}^{j_2} \quad (2.21)$$

Whether one chooses to have these representations act from the left or the right, the commutant naturally acts from the opposite side. Commutants in general need not be representations of  $G$ , because although they act upon the carrier space, they need not be homomorphisms of  $G$ . But regular representations have regular representations acting from the opposite side as commutants: see for instance Hall (1959, §6.3).

The tensor  $e$  is an operator effecting an average over the orthogonal group from the right (c.f. (2.21)):

$$e_{i_1 i_2}^{j_1 j_2} = \int_{O(n)} (H)_{i_1}^{j_1} (H)_{i_2}^{j_2} dH \quad (2.22)$$

where  $dH$  is the normalised Haar measure on the orthogonal group.

### 2.5.5.1 Zonal spherical functions

Zonal spherical functions arise in the context of representations when there are massive subgroups of class 1: see §2.7 on p. 68. But we shall place the discussion in the context of  $R1$  and  $R1R$ .

Because the commutant  $e\mathfrak{A}e$  is itself commutative (e.g. Farrell (1985, thm. 12.10.4)),  $R1$  is multiplicity free: i.e.  $R1$  is the direct sum of mutually inequivalent irreducible representations. Those irreducible representations are indexed by the even partitions  $2\lambda$  of  $2f$ , from Hannan (1965a, p. 36), James (1961), Macdonald (1995, p. 402), *i.a.*: say

$$R1 = \bigoplus_{\lambda \vdash f} R1_{2\lambda}$$

where  $R1_{2\lambda}$  is the irreducible representation of  $GL(n)$  associated with the partition  $2\lambda \vdash 2f$ .

Each  $R1_{2\lambda}$  however becomes reducible when it is restricted to represent  $O(n)$  rather than  $GL(n)$ :

$$(R1R) \qquad \mathfrak{A}' \supset \mathfrak{A}e$$

where  $\mathfrak{A}'$  is that subalgebra of  $\mathfrak{A}$  generated by the orthogonal matrices only: i.e.

$\mathfrak{A}'$  is spanned by tensors of the form

$$(H)_{i_1}^{j_1} (H)_{i_2}^{j_2} \dots, \quad \text{for } H \in O(n). \quad (2.23)$$

Each  $R1_{2\lambda}$  is restricted in this way to become  $R1R_{2\lambda}$ .

The trivial representation occurs exactly once in  $R1R_{2\lambda}$ , from Frobenius reciprocity (e.g. Burrow (1965)). Thus there is exactly one element of the representation matrices in  $R1R_{2\lambda}$  which is fixed at unity, with the remainder of the row and column in which it occurs identically zero (for an appropriate basis of the carrier space): see i.a. James (1968), Mathai et al. (1995, p. 319).

Moving back to the representation matrices in  $R1_{2\lambda}$ , say from  $T^\lambda(h)$  to  $T^\lambda(g)$ , and assuming both that a suitable basis of the carrier space has been chosen and that the left cosets of  $H$  represent points of  $Z$ , then the first column of functions  $T_{i1}^\lambda(g)$  are functions which are right invariant under  $H$  and therefore *de facto* functions on  $Z$ , called spherical functions or associated spherical functions (Hannan (1965a, p. 16), Vilenkin (1968, p. 30), Hillier & Skeels (1996, p. 164)); and the topmost of these functions is also left invariant under  $H$ , and so is a zonal function: it is called a zonal spherical function, labelled as  $\phi^\lambda(g)$ , and is invariant under the action of  $H$  from either side.

We mention the following points, highlighting the special nature of the zonal spherical functions  $\phi^\lambda(g)$  for  $\lambda \vdash f$ .

1. There is a link between the biinvariant functions and the commutant, noted in §2.4.2.3 on p. 45. For a representation containing the irreducible component  $R1_{2\lambda}$  once only, the corresponding block in the matrix in the commutant assumes the form  $\phi^\lambda(g)I$ . It is in this sense that the zonal spherical functions  $\phi^\lambda(g)$  span the commutant.

Since  $e\mathfrak{A}e$  is a commutative algebra, the identity is the (direct) sum of the primitive idempotents. These primitive idempotents are the zonal spherical functions, or more precisely the blocks  $\phi^\lambda(g)I$  for  $\lambda \vdash f$ .

This form of the commutant generalises to the situation in which the multiplicity exceeds one (Hannan (1965a, p. 16)), but we do not need the extra

generality here because  $R1$  is multiplicity free.

2. If the carrier space contains functions on  $Z$ , i.e. the carrier space contains functions on  $G$  which are right invariant under  $H$ , the zonal spherical function is also in the carrier space: the zonal spherical function spans the 1 dimensional space which is invariant under  $T^\lambda(h)$ .
3. The zonal spherical functions are a basis of biinvariant functions on  $G$ .
4. With the exception of the first element, functions in the first row of  $T^\lambda(g)$  are not constant on left cosets of  $H$ , and do not represent functions on the set  $Z$ .
5. Rows and columns are reversed in the above comments should right cosets correspond to points of  $Z$  rather than left cosets.

The functions  $T_{ij}^\lambda(h)$  in  $R1R$  are mutually orthogonal, with respect to integration over the uniform measure on  $H$ , with the product of functions being that in the group algebra of  $H$ .

Choosing one of the basis functions to be  $\phi^\lambda(h)$ , one can integrate the elements  $T_{ij}^\lambda(h)$  over the orthogonal group, whereupon all integrals vanish save for that with  $(i, j) = (1, 1)$ , viz. the zonal spherical function. The zonal spherical function is identically unity on  $O(n)$ , and one sees the importance of using the zonal spherical functions as basis functions for the purpose of integration over the orthogonal group.

More generally, Vilenkin & Klimyk (1995, §3.3.1) show that zonal spherical functions of a Lie Group  $G$  are eigenfunctions of operators having the form of Laplace-Beltrami operators. These functions are related to different root systems, and need not be hypergeometric functions.

#### 2.5.5.1.1 Zonal polynomials

Macdonald (1995, p. 419) states that a polynomial function  $f(x)$  of the elements of the matrix  $x$  which is invariant under the action from the left of the orthogonal group

$O(n)$  necessarily assumes the form  $f^*(y)$ , a polynomial function of the elements of  $y = x^T x$ . That is

$$f(x) = f(kx) \quad \text{for all } k \in O(n) \text{ and all } x$$

$$\Rightarrow f(x) = f^*(x^T x)$$

The zonal spherical function  $\phi^\lambda(x)$  has been expressed as the zonal polynomial  $\phi^{\lambda^*}(x^T x) = \phi^{\lambda^*}(y)$ .

## 2.5.6 From R0a to R3, and from R0b to R6

### 2.5.6.1 The actions of $\mathfrak{A}$ and $S_{2f}$ on the carrier space $P_{2f}$

Let  $P_{2f}$  denote the space of covariant tensors of rank  $2f$ . Then  $\mathfrak{A} \ni a$  acts on  $x \in P_{2f}$  as follows, where we illustrate with  $f = 1$ :

$$(ax)_{i_1 i_2} = a_{i_1 i_2}^{j_1 j_2} x_{j_1 j_2}$$

The action of the symmetric group  $S_{2f}$  on  $P_{2f}$  is as follows:

$$(sx)_{i_1, i_2} = x_{i_{s1}, i_{s2}}$$

### 2.5.6.2 The actions of $S_{2f}$ and $\mathfrak{A}$ on $P_{2f}$ commute

The representation of  $\mathfrak{A}$  on  $P_{2f}$ , with each  $a \in \mathfrak{A}$  representing itself upon multiplying from the left, has commutant  $S_{2f}$ . This can be seen as follows, still using  $f = 1$  for illustration:

$$\begin{aligned} (asx)_{i_1 i_2} &= a_{i_1 i_2}^{k_1 k_2} (sx)_{k_1 k_2} = a_{i_1 i_2}^{k_1 k_2} x_{k_{s1}, k_{s2}} \\ (sax)_{i_1 i_2} &= (ax)_{i_{s1}, i_{s2}} = a_{i_{s1}, i_{s2}}^{k_1 k_2} x_{k_1 k_2} = a_{i_1 i_2}^{k_{s-1} 1, k_s - 1 2} x_{k_1 k_2} = a_{i_1 i_2}^{k_1 k_2} x_{k_{s1}, k_{s2}} \end{aligned}$$

The point is made more succinctly in Farrell (1980, p. 305). The commutativity of the two actions is also noted in Hannan (1965a, p. 34).

Note that in the coordinate free definition of a bisymmetric matrix in Farrell (1985, Defn. 12.1.8) as a matrix  $A$  satisfying

$$P_\sigma A = A P_\sigma,$$

where  $P_\sigma$  are (orthogonal) permutation matrices operating on the indices, recasting the equation as

$$A = P_\sigma^{-1} A P_\sigma = P_\sigma^T A P_\sigma$$

shows that  $A$  is simply invariant under the rows and columns being subject to the same permutation.

### 2.5.6.3 Representations with $P_{2f}$ as carrier space

It was James' insight to consider

$$(R0a) \quad \mathfrak{A} \supset P_{2f} \triangleleft S_{2f}$$

and

$$(R0b) \quad S_{2f} \supset P_{2f} \triangleleft \mathfrak{A}$$

and to mimic the basic representation of interest  $R2$  by imposing symmetry requirements on  $P_{2f}$  in the carrier spaces in R0a and R0b. Imposing the symmetry arising from the action of the orthogonal group using  $e \in \mathfrak{A}$  as defined above in (2.22) will be used for R0b; and imposing symmetry by permuting the indices with the symmetric group  $S_{2f}$ , leading to the projection  $e_1 \in S_{2f}$ , to be defined below, will be utilised for R0a.

### 2.5.6.4 The hyperoctahedral group and the projection operator $e_1$

The class of polynomials of degree  $f$  in elements of a symmetric matrix  $Y$  contains elements of the form

$$y_{i_1}^{j_1} y_{i_2}^{j_2} y_{i_3}^{j_3} \dots y_{i_f}^{j_f}$$

the enveloping algebra for which class contains tensors of rank  $2f$ , with  $f$  covariant and  $f$  contravariant indices

$$z_{i_1 i_2 i_3 \dots i_f}^{j_1 j_2 j_3 \dots j_f} \quad (2.24)$$

satisfying particular symmetry conditions. These tensors are, firstly, to be bisymmetric with respect to  $S_f$  operating on lower and upper indices simultaneously; and secondly to be of unchanged value when a given vertical pair of indices are interchanged, since  $Y$  is symmetric. The subgroup of  $S_{2f}$  containing these symmetries is the hyperoctahedral group, say  $HG_{2f}$ , or more simply  $HG$ . There are  $f!$  ways of permuting the vertical pairs of indices, to be multiplied by  $2^f$  to allow for the possible interchange of each such vertical pair. For general  $f$ , then,  $HG_{2f}$  is a subgroup of  $S_{2f}$  of order  $2^f f!$ . See Macdonald (1995, p. 401).

Although the form of (2.24) is convenient from the point of view of seeing which symmetries should apply to the indices, in order to consider  $z$  as a member of the carrier space  $P_{2f}$  one wants to consider all the indices as covariant in order to preserve the usual rules of tensor contraction. We rewrite (2.24) in the form

$$z_{i_1, j_1; i_2, j_2; \dots; i_f, j_f} \quad (2.25)$$

from which the connection with the doublets in §2.5.8 will become clear.

In order to impose the requisite symmetry on the tensor in (2.25) above, the operator is a simple average over the elements of the hyperoctahedral group (see the discussion in §2.4.2.6 on p. 48). Following James (1961) we call this operator  $e_1$ :

$$e_1 = \frac{1}{2^f f!} \sum_{s \in HG} s \quad (2.26)$$

The set  $e_1 P_{2f}$  thus consists of elements of the form (2.25) such that for any  $s \in S_f$ ,

$$z_{i_1, j_1; i_2, j_2; \dots; i_f, j_f} = z_{i_{s1}, j_{s1}; i_{s2}, j_{s2}; \dots; i_{sf}, j_{sf}}$$

and for which in addition there is symmetry when interchanging  $i_t$  and  $j_t$  for any  $t$  in (2.25).

## 2.5.7 Comparing R4 and R5

The isomorphism between the representations

$$(R4) \quad S_{2f} \supset S_{2f} e_1 \triangleleft e_1 S_{2f} e_1$$

and

$$(R5) \quad S_{2f} \triangleright \mathfrak{D} \triangleleft \mathfrak{F}$$

follows because  $e_1$  is the projection onto the isotropy group at the doublet (12)(34) ... Thus  $\mathfrak{D}$  is the set of functions on the cosets of the isotropy group; and we choose to work with the average over the isotropy group than with the coset, in line with the comments in §2.4.2.6.

## 2.5.8 Comparing R5 and R6

To show the connection between

$$(R5) \quad S_{2f} \triangleright \mathfrak{D} \triangleleft \mathfrak{F}$$

and

$$(R6) \quad S_{2f} \triangleright eP_{2f} \triangleleft e\mathfrak{A}e$$

is however less straightforward.

Citing the first main invariant theorem in Weyl (1946, ch. 2), James (1961, eqn. (15)) states that the subspace  $eP_{2f}$  is generated by tensors assuming the form of products of Kronecker delta functions. That is, the set  $eP_{2f}$  has as basis the set of functions of the form

$$\delta_{i_{s1}, i_{s2}} \delta_{i_{s3}, i_{s4}} \cdots \delta_{i_{s(2f-1)}, i_{s(2f)}}$$

as  $s$  varies over a transversal of  $HG$  in  $S_{2f}$  (e.g. Macdonald (1968, p. 44)), i.e. one representative is chosen from each coset.

James' claim can be illustrated by an example with  $f = 2$ :

$$\begin{aligned} e(\delta_{j_1, j_3} \delta_{j_2, j_4}) &= \int (H)_{i_1}^{j_1} (H)_{i_2}^{j_2} (H)_{i_3}^{j_3} (H)_{i_4}^{j_4} \delta_{j_1, j_3} \delta_{j_2, j_4} dH \\ &= \int (H)_{i_1}^{j_1} (H)_{i_2}^{j_2} (H^T)_{j_3}^{i_3} (H^T)_{j_4}^{i_4} \delta_{j_1, j_3} \delta_{j_2, j_4} dH \\ &= \int (HH^T)_{i_1}^{i_3} (HH^T)_{i_2}^{i_4} dH = \int \delta_{i_1, i_3} \delta_{i_2, i_4} dH = \delta_{i_1, i_3} \delta_{i_2, i_4} \end{aligned}$$

The connection between R5 and R6 is then immediate. For the tensor  $\delta_{i_1 i_3} \delta_{i_2 i_4} \delta_{i_5 i_6} \dots$ , for instance, there is a direct correspondence with the doublet  $(13)(24)(56) \dots$

The point about  $eP_{2f}$  being generated in the above fashion is more readily understood from Macdonald (1995, p. 419). Let  $v_1, \dots, v_n$  be the columns of an  $n \times n$  matrix such that a polynomial function  $f(v_1, \dots, v_n) = f(kv_1, \dots, kv_n)$  for all  $k \in O(n)$ . Then Macdonald states that  $f$  is a polynomial in the inner products  $(v_i, v_j)$ . Choosing the pairs of vectors  $v_i$  and  $v_j$  corresponds to picking a doublet; and the functions  $f$  correspond to the functions on the doublets, which James (1961) denotes by  $\mathfrak{D}$ , as do we. Macdonald also refers to Weyl (1946, ch. IIA, §9) for the first main theorem on invariants, but the reasoning in his treatment does not depend on Weyl's theorem.

In any case James (1961, eqn. (5)) had already assumed this result of Weyl's earlier in his paper, when he defined the projection operator  $e$  (our equation (2.22)). There an average over the orthogonal group to define R1 was assumed to be of the form  $\phi(x^T x)$  for suitable matrices  $x$ : making this assertion allowed his claim that R1 and R2 were equivalent.

Let  $t_\nu = \delta_{i_{s1} i_{s2}} \delta_{i_{s3} i_{s4}} \dots \in eP_{2f}$  for the doublet  $\nu = (i_{s1} i_{s2}) (i_{s3} i_{s4}) \dots \in D$ , where  $s \in S_{2f}$ ; and  $d_\nu \in \mathbb{D}$  a function assuming the value 1 at  $\nu$ , 0 otherwise. Then the matrices in the commutants contain elements of the form  $\alpha(\mu, \nu) d_\mu d_\nu^*$  or its image  $\alpha(\mu, \nu) t_\mu t_\nu^*$  under the isomorphism between the commutants of R5 and R6, where the asterisk denotes the corresponding member of the dual space, and  $\alpha(\mu, \nu)$  is a point pair invariant function (see §2.4.2.4 on p. 46). James calculates  $\alpha(\mu, \nu)$  as the number of cycles in his diagram "intertwining" the doublets  $\mu$  and  $\nu$ : this number turns out to be the length of  $\lambda$  when  $2\lambda$  defines the equivalence class in question (James (1961, p. 464)).

Reverting to  $f = 3$ , and for the doublets  $\mu = (13)(24)(56)$ ,  $\nu = (14)(23)(56)$ , for instance, the  $(\mu, \nu)$ th element of a matrix say  $\Gamma$  in the commutant of R6 is

$$t_\mu t_\nu^* = \delta_{i_1 i_3} \delta_{i_2 i_4} \delta_{i_5 i_6} \delta^{j_1 j_4} \delta^{j_2 j_3} \delta^{j_5 j_6}$$

Taking the trace of this matrix leads to

$$\text{tr } \Gamma = \delta_{i_1 i_3} \delta_{i_2 i_4} \delta_{i_5 i_6} \delta^{i_1 i_4} \delta^{i_2 i_3} \delta^{i_5 i_6} = \delta_{i_1 i_3} \delta^{i_3 i_2} \delta_{i_2 i_4} \delta^{i_4 i_1} \times \delta_{i_5 i_6} \delta^{i_6 i_5} = n^2$$

in accord with James' claim that this number is  $n^{\alpha(\mu, \nu)}$ : for these values of  $\mu$  and  $\nu$ ,  $\lambda = (2, 1) \vdash 3$  and  $\alpha(\mu, \nu) = \ell(\lambda) = 2$ .

## 2.6 Macdonald's definition of zonal polynomials

The treatment of basically the same material as James' in Macdonald (1995, Ch. I, App. A; Ch. VII) is terse, and has the demerit of demanding rather a strong background in algebra, including the rudiments of category theory. On the other hand, Macdonald's treatment is more or less self contained, in contrast to James' treatment, resting as the latter partially does on rather obscure initial references.

It is not our intention to dissect Macdonald's treatment as closely as we did James'. Rather we shall emphasise the overall approach that Macdonald is taking, placing it in broad terms within the framework of James (1961), and relating specific developments in Macdonald's book to the ubiquitous diagram (2.13) of James.

### 2.6.1 The initial setup

We first show that Macdonald and James (1961) are basically taking a similar approach to the same problem.

Following Macdonald (1995, p. 388) we write  $G$  for a finite group,  $A$  as its group algebra, and  $K$  for a subgroup. Setting  $e = (1/|K|) \sum_{k \in K} k$ , Macdonald claims that  $\text{End}_A(Ae)$  and  $eAe$  are anti-isomorphic. This can be seen as follows.

The elements of  $\text{End}_A(Ae)$  are  $\phi_h$ , say, where  $\phi_h(e) = he$ . One chooses the elements  $h$  from a right transversal, i.e. one representative  $h$  is chosen from each right coset (again see Macdonald (1968, p. 44)).

Then  $\phi(ge) = g\phi(e)$  from the definition of an endomorphism of a module. Using the fact that  $e$  is an idempotent leads to

$$\phi_h(e) = \phi_h(ee) = e\phi_h(e) = ehe$$

$$\phi_h(ge) = g\phi_h(e) = gehe$$

$$\phi_{h_1}\phi_{h_2}(ge) = geh_2eh_1e$$

This is nothing more than the representations R1 and R4 in (2.13) on p. 50, upon substituting  $\mathfrak{A}$  and  $S_{2f}$  respectively for  $G$  in the above.

The result is analogous with that in Macdonald (1995, (3.6), p. 416) stating that the endomorphism ring of the right cosets of  $K$  is isomorphic to the biinvariant functions, in the context of polynomial functions on elements of matrices in  $GL(n)$ , and where  $K$  is the orthogonal group. The operator  $e_K$  in that case is analogous to that in (2.22) on p. 56.

## 2.6.2 Development of Macdonald's ideas

As with James (1961) and Farrell (1985, ch. 12), the basic element on which Macdonald operates is the tensor power of a matrix. From Macdonald (1995, p. 156) we have

$$T^{2m}(V) = T^{2m}V = \bigoplus_{\lambda \vdash 2m} M_\lambda \otimes F_\lambda(V) \quad (2.27)$$

where  $M_\lambda$  is the irreducible representation of  $S_{2m}$  indexed by  $\lambda$  and  $F_\lambda(V)$  is the irreducible (symmetrised tensor) representation of  $GL(n)$  indexed by  $\lambda$ . The notation in (2.27) perhaps needs some clarification. The direct sum is firstly a breakdown of the vector space  $T^{2m}V$  into a direct sum of subspaces, as indicated. The components are however not just vector subspaces, but also act as carrier spaces to matrix operators which represent the indicated groups. And  $T^{2m}V$  is to be thought of not just as a vector space, but as the representation of  $T^{2m}A$  with carrier space  $T^{2m}V$ . More precisely,  $T = T^{2m}V$  is a  $S_{2m} \times G$ -module (Macdonald (1995, p. 421)).

The expansion (2.27) requires only that the coefficient field for  $V$  has characteristic 0, and hence is valid for both  $\mathbb{R}$  and  $\mathbb{C}$ .

The tensor power  $T = T^{2m}V$  defined in Macdonald (1995, p. 417) is the direct analogue of  $P_{2f}$  in James (1961). Denoting the hyperoctahedral group by  $HG$  and the orthogonal group by  $K$ , we have the direct correspondences

$$T^{HG} = e_1 P_{2f} \quad \text{and} \quad T^K = e P_{2f} \quad (2.28)$$

where we write equals for isomorphisms. The superscript in  $T^{HG}$  denotes an averaging over the application of  $HG$  or equivalently the subset invariant under the action of  $HG$ ; and similarly for  $K$ . Thus we are essentially working with R3 and R6, to use the notation in (2.13).

We have noted elsewhere that the complex case is much easier than the real case, and nowhere is this clearer than in Farrell's work. In Macdonald's work this fact is manifested in (2.28), where the complicated action of the wreath product  $HG$  is replaced in the complex case by a simple diagonalisation operator, utilising the complex conjugate of a complex number: see Macdonald (1995, §VII.5).

The Gelfand pairs  $(S_{2m}, HG)$  and  $(G, K)$  provide the representations R4 and R1 respectively, as already noted.

Macdonald (1995, p. 414) takes the basic action defining R7 in (2.13) as

$$p(x) \rightarrow gp(x) = p(xg)$$

Thus a point corresponds to a right coset of the isotropic subgroup, not the left coset as in Hannan (1965a). There is a consequential distinction between (2.16) and the action on symmetric real matrices underlying R2 in (2.13):

$$p(\sigma) \rightarrow gp(\sigma) = p(g^T \sigma g),$$

from Macdonald (1995, p. 417), in which  $x, \sigma \in GL(n)$  and in addition  $\sigma$  is symmetric.

Macdonald (1995, p. 402) shows that  $T^K = \oplus M_{2\lambda}$  from consideration of the induced representation  $1_{HG}^S$ . The character of the induced representation turns out to be the plethysm  $h_n[h_2] = \sum_{\lambda \vdash n} s_{2\lambda}$ , from Read (1968, §7.2) or Littlewood (1950, p. 206), since  $HG$  is the semi-direct (wreath) product of  $S_2$  and  $S_m$ . See Cohn (1981, p.

278) for an explanation of the wreath product, and Read (1968) for the connection between wreath product and plethysms.

Macdonald's proof that  $T^{HG} = \oplus F_{2\lambda}$  is longer, and utilises the fact that regular representations have irreducible components repeated the same number of times as their dimensionality.

He makes the connection between R1 and R3 by showing that also  $e_K P(G) = \oplus F_{2\lambda}$ . Finally, like James, he moves from R1 to R2 by inserting  $Y = XX^T$ , or the equivalent: the mapping  $p(x^T x) \rightarrow \tilde{p}(x)$  in Macdonald (1995, p. 418).

Apart from its innate elegance, Macdonald's approach offers the advantage of assuming at the start that the representations of interest are going to be multiplicity free, i.e. that  $(G, K)$  is a Gelfand pair: this simplifies matters considerably. In other words Macdonald restricts his attention to double cosets of  $K$ , or rather the space of functions on  $G$  which are constant on double cosets, which is to be spanned by the zonal spherical functions. He establishes a simple lemma stating that the pair is Gelfand when  $KxK = Kx^{-1}K$  for all  $x \in G$ . He then applies the lemma to the situations of interest to show that in fact he need only be concerned with the case in which  $(G, K)$  is a Gelfand pair.

This could be contrasted with Vilenkin (1968, p. 29), who defines zonal spherical functions when a representation  $T(g)$  of  $G$  is of class 1 relative to a massive subgroup  $H$ . His approach is adumbrated in §2.7, and the basic equations there are similar to those in Hannan (1965a, pp. 15, 16).

Vilenkin's methodology is well suited to obtaining parallels with classical systems of orthogonal polynomials, and extending into new families of special functions. Also his approach lends itself readily to more complicated scenarios, such as representation theory over non-compact, infinite dimensional Lie groups, etc. Vilenkin & Klimyk (1992, p. 321), for instance, deal with Gelfand pairs for non-compact isotropy groups, for which generalised functions, or "distributions" in the sense of say Antosik, Mikusinski & Sikorski (1973), are required.

But however elegant Vilenkin's outworking, he seems to fail to mention just how often his assumptions are going to be valid, and what sufficient conditions there might be to guarantee the validity of his theoretical outworking. The beauty of Macdonald's approach lies in his applying a simple criterion to show that his simplifying assumptions are valid for the situation of interest.

## 2.7 Vilenkin and Klimyk's work

The work of Vilenkin and Klimyk in the area of Jack polynomials is part of an encyclopaedic publication exploring the interrelationship between special functions and representations of Lie groups. Their collaborative effort appears first as a three-volume set Vilenkin & Klimyk (1991) Vilenkin & Klimyk (1993) and Vilenkin & Klimyk (1992); and a separate volume appeared soon afterwards: Vilenkin & Klimyk (1995). Publication was some years in the pipeline, and in fact Vilenkin died at about the same time that the first volume appeared.

The four volumes bear the strong imprint of Vilenkin (1968), which made its original appearance in Russian in 1965. Vilenkin's and Klimyk's approach facilitates the ordering of a vast array of seemingly disparate results on the various types of classical orthogonal polynomials; and allows extension into new families of invariant polynomials and other special functions (Vilenkin & Klimyk (1991, Intro.), Vilenkin (1968, Preface)).

Their approach to spherical functions is to consider a transformation group  $G$  with isotropy group  $H$  as above, and to define  $H$  to be a massive subgroup of class 1 when there is essentially only one vector  $a$  such that  $ha = a$  for all  $h \in H$ . Setting  $a = e_1$  and  $e_2, e_3, \dots$  as remaining members of an orthonormal basis of the carrier space, Vilenkin (1968, p. 30) defines the zonal spherical function as

$$t_{11}(g) = (T(g)e_1, e_1) \quad (2.29)$$

The inner product may be considered the canonical one, viz.  $(e_i, e_j) = \delta_{ij}$ .

The spherical functions, or the associated spherical functions, are defined as

$$t_{1i}(g) = (T(g)e_i, e_1) \quad (2.30)$$

which are constant on right cosets of  $H$  in  $G$ .

Vilenkin & Klimyk (1991, p. 101 ff) give essentially the same approach; while Hannan (1965a, pp. 15, 16) gives formulae analogous to (2.30), although in other respects his treatment is quite different.

One gets the impression that the work of Vilenkin and the other Russian writers (Gelfand, Naimark etc.) in the area of group representation theory and its applications to special functions was not widely known in the West, at least until the late 1980s or 1990s; and in particular the work was not known amongst the statistical community.

## 2.8 Farrell's work on zonal polynomials and summary of James', Macdonald's and Farrell's approaches

Farrell distinguished between James' and his approaches to zonal polynomials in that they were classified separately in §2.1. His approach to "complex" zonal polynomials is elegant, and distinct from James' approach to real zonal polynomials in James (1961), but his method does not generalise to the real case.

James (1964) is an extraordinary farrago of results which were known to James and Constantine at that time. As Farrell (1985, §12.0) notes, the paper is apparently without error, but contains a large number of striking assertions concerning zonal polynomials of both the real and complex kinds, the provenance of which was unclear. In the remainder of that same chapter Farrell sets out to justify the general tenor of results in James (1964), with an approach which is basically the same as that adopted in James (1961) and Macdonald (1995), although offering greater intuition.

## 2.8.1 Farrell's approach to complex zonal polynomials

Farrell (1980) works with polynomials in the elements of a matrix  $X \in GL(n)$  of the form

$$e(X) = \text{tr} (E_0 X^{\otimes m}) .$$

The matrix  $E_0$  is bisymmetric, and the symbol  $e$  does not have the implication of a projection operator: we simply follow Farrell's notation.

Again denote by  $\mathfrak{A}$  the algebra of bisymmetric matrices, and its centre by  $C(\mathfrak{A})$ . Working in the complex field, Farrell (1980, p. 303) shows the equivalence of

$$e(UXU^*) = e(X) \quad \text{for} \quad U \text{ unitary and } X \in GL(n);$$

$$e(XY) = e(YX) \quad \text{for} \quad X, Y \in GL(n); \quad \text{and}$$

$$E_0 \in C(\mathfrak{A}) .$$

Whereas James' basic approach was to project the bisymmetric matrices onto a subspace invariant under the orthogonal group, Farrell's is to consider  $C(\mathfrak{A})$ . A key lemma in Farrell (1985, §12.6), due to Wedderburn, isolates the nature of the centre of the algebra when subalgebras are mutually commutant. The lemma finds an elegant application to the interrelationship between the symmetric group and the general linear group in R0a and R0b, showing that the centre is spanned by the zonal spherical functions.

### 2.8.1.1 Product of group characters

The irreducible components of R7 in (2.13) on p. 50 are repeated as many times as their number of rows, since R7 is a regular representation. One block of this representation, corresponding to the partition  $\lambda$  say, contains  $f_\lambda$  multiples of an irreducible representation  $F_\lambda$ , say. Then the submatrix corresponding to  $F_\lambda$  in  $C(\mathfrak{A})$  will contain  $A \otimes I_\lambda$ , provided we are working within  $\mathbb{C}$ ; where  $I_\lambda$  denotes the identity matrix of appropriate dimension (eg, see Hannan (1965a, p. 12)); and  $A$  is  $f_\lambda \times f_\lambda$ .

$R1$  is multiplicity free, and the irreducible element of the centre of  $e\mathcal{A}e$  is  $P_{\epsilon_\lambda}$  in Farrell's notation, which corresponds to  $aI_\lambda$ , treated as one diagonal block in an otherwise zero  $n^m \times n^m$  matrix. The  $a$  is a character of  $GL(n)$ ; and each individual diagonal element of the  $I$  matrix gives an equivalent representation to another in which the position of the non-vanishing diagonal element is permuted. Permuting those diagonal elements multiplies the character of  $GL(n)$  by the number of rows of the irreducible representation, viz.  $\chi_\lambda(1)$ .

We thus have the following simple formulation of Schur functions (or "complex" zonal polynomials, or Jack polynomials with  $\alpha = 1$ ) in terms of group characters:

$$\widetilde{C}_\lambda(X) = \chi_\lambda(1)\chi_\lambda(X) \quad (2.31)$$

The simplicity of (2.31) is not replicated when working in the real field  $\mathbb{R}$ , because the commutant assumes a more complicated form.

### 2.8.1.2 Derivation of key formulae for complex case

One advantage of Farrell's approach is a quick derivation of several of the key integral identities which characterise the zonal polynomials, but derived for the complex zonal polynomials. Setting  $U(n)$  to be the unitary group, chief among these identities are the "splitting identity":

$$\int_{U(n)} \widetilde{C}_\lambda(UXU^*Y)dU = \frac{\widetilde{C}_\lambda(X)\widetilde{C}_\lambda(Y)}{\widetilde{C}_\lambda(I)} \quad (2.32)$$

and the reproducing property of the zonal polynomial:

$$\widetilde{C}_\lambda * \widetilde{C}_\lambda = a_\lambda \widetilde{C}_\lambda$$

where the convolution  $*$  denotes the product in the group algebra

$$\widetilde{C}_\lambda * \widetilde{C}_\mu(X) = \int_{U(n)} \widetilde{C}_\lambda(XU^{-1})\widetilde{C}_\mu(U)dU,$$

and  $a_\lambda$  is a constant. There are several other integral identities of this type, eg Farrell (1985, p. 264); and Farrell (1980, p. 304).

There are often equivalent formulae for “real” zonal polynomials. The equivalent to (2.32), for instance, from James (1961, p. 467) or Macdonald (1995, §VII.4), is

$$\int_{O(n)} Z_{\lambda}(AHBH')dH = \frac{Z_{\lambda}(A)Z_{\lambda}(B)}{Z_{\lambda}(I)} \quad (2.33)$$

Nevertheless it is clear in general that working in the complex field is often far simpler than working in the real field. The most obvious example from this thesis is that the Schur functions are far more tractable than zonal polynomials, a fact noted by several authors, including Takemura (1984) and Farrell (1985, §12.0). Moreover, Bingham et al. (1992, p. 316) comment that the complex case is easier than the real for their work in statistics on the sphere (see §2.3.4). Smith and Gao make similar comments, noting in particular that the random variates defined by (2.10) and (2.11) are quite intractable in the real case: see §2.3.5 on p. 38.

## 2.8.2 Mini summary of James’, Macdonald’s and Farrell’s arguments

The nub of the argument in each of James (1961), Farrell (1985, Ch. 12) and Macdonald (1995) is, in the notation of §2.6.2 on p. 65,

$$T_{2\lambda}^{HG \times K} = (T_{2\lambda}^K)^{HG} = (T_{2\lambda}^{HG})^K$$

which gives a one dimensional space in centre of  $T$ , invariant under both  $HG$  and  $K$ , spanned by the zonal spherical function corresponding to  $2\lambda$ .

## 2.9 The place of zonal polynomials in statistical theory

### 2.9.1 The central place of zonal polynomials in multivariate statistical theory

The central position of zonal polynomials in multivariate statistical theory is vicarious, in that it is the density function which is central; and the density function of statistics and random variates arising in multivariate theory tend to be hypergeometric functions of matrix argument(s), which are most readily defined as expansions in zonal polynomials. Statistical theory often proceeds via generating functions, particularly the Laplace transform and the Fourier transform; and when the density functions have hypergeometric function expansions in zonal polynomials, so do these generating functions.

The two arms of statistical inference are estimation of parameters, and hypothesis testing. Looking first to the estimation, to obtain maximum likelihood estimators one needs to integrate the density over the parameter space for fixed values of the variate. For method of moments estimators, one needs the moments of the distribution: given the generating function, the moments are obtained from setting the variates to unity in the zonal polynomials; alternatively one evaluates an integral involving the hypergeometric expansion for the density over the entire variate space to obtain the moments. As for hypothesis testing, the basic aim is to calculate probabilities, for which one has to integrate the density over a subspace of the variate space.

While the main use of zonal polynomials in statistical theory is as a constituent of hypergeometric functions as described above, that is not their only use. Their role in directional statistics is rather indirect, as discussed in §2.3.4 on p. 35; but their position in generalising and improving Stein's variance estimators in higher dimensions was pivotal. Utilising methods which depended on the properties of zonal polynomials, Shorrocks & Zidek (1976) for instance found optimal estimators

of the generalised variance (the determinant of the covariance matrix) for the linear model. See also Zidek (1978) and references therein.

### 2.9.2 Little need for exact calculation

Even if the zonal polynomials are so central to the theory of mathematical statistics, it may not be overly restrictive to seek merely a workable approximation to zonal polynomials rather than precise evaluation. After all, the practitioner is not primarily interested in calculation of a zonal polynomial as a one-off exercise. One is interested in having workable approximations to whole families of them (for all partitions of a given weight, say), and efficient algorithms for repeating those calculations for growing partitions of the one lot of variates on the one hand; and for a grid search for numerical integration over many collections of variate values on the other.

The challenge for the first of these is to use the previously evaluated zonal polynomials for their calculation for partitions of higher weight, in an efficient recursive algorithm. The challenge in the second is to find some way of summarising the properties of a collection of variates, so that one does not have to recalculate completely from scratch upon a change of variates. One possible way of doing this is to fit a geometric progression to a set of variates, easily done by obtaining a straight line regression of the logarithms of (positive) sample variates, when the Jack polynomial can relatively easily be approximated by either the  $e_\lambda$  or  $m_\lambda$  function expansion: we make a preliminary excursion along this path in Chapters 6 and 7. See also the comments in §8.3 on p. 190.

There is clearly scope for accumulating rounding errors, at any of the stages of calculation. But on the other hand it may not be crucial to obtain very precise results. For estimating tail probabilities of a statistic, for instance, an answer of say  $10^{-3}$  for a probability may well be sufficient for one's purpose: an error by a factor of 10, say, still means that a test of a hypothesis can be rejected at the 1% level. An approximation to the value of the density function, or better an exponentially falling upper bound on its value, will facilitate greatly the numerical calculation of

the integral for the probability, or possibly hasten a decision that the probability is nugatory.

### 2.9.3 Two basic problems with the application of zonal polynomials

There are two basic problems with using zonal polynomials in practice.

1. numerical awkwardness
2. interpretation of the successive terms of the hypergeometric expansions in zonal polynomials.

#### 2.9.3.1 Numerical awkwardness

There are several aspects to this.

1. Hypergeometric expansions are slow to converge. Muirhead (1975, p. 284) comments on how hard it is to work with zonal polynomials, citing Sugiyama needing 100 terms of the series for a  ${}_1F_1$  function to converge. To have problems with the  ${}_1F_1$  function augurs ill for the use of the “balanced”  ${}_2F_1$  and  ${}_1F_0$  functions; for in the  ${}_1F_1$  function the  $n!$  term in the denominator of the hypergeometric function expansion has no counterweight in the numerator.
2. Within each “overall” term of the hypergeometric expansion lies a summation over zonal polynomials indexed by partitions of a given weight; and partitions of integers are intrinsically highly non-linear.

One problem with using the RLO, the conventional total ordering of parti-

3. Calculation of each zonal polynomial is difficult. This thesis for instance is concerned with their calculation, or rather their calculation via determinants involving  $e_\lambda$  or  $m_\lambda$  functions. And the calculation of these basis functions may be non-trivial, let alone forming them into determinants.
4. For the one set of variates, one needs efficient ways to reproduce calculations for different partitions of the one weight; and then efficient algorithms to perform calculations for partitions of increasing weight, recursively from the partitions of lower weight. Again, this task is complicated by the highly non-linear nature of partitions.
5. One needs then to repeat the procedure for different sets of variates. Ideally one would like to expedite this by utilising previous calculations.

What is remarkable is that not only are the zonal polynomials hardly ever calculated, there is virtually no literature on their approximations, or on algorithms that can be applied quickly, if approximately, to their estimation. Nor does there seem to be any work done on approximating  $e_\lambda$  and  $m_\lambda$  functions in the literature, or at least in the statistical and combinatorial literatures.

The fact that so little attempt to calculate zonal polynomials is made, even with the extraordinary increase in computing power since their definition in 1960, indicates that the problem lies above all in obtaining sensible algorithms for their calculation. Part of the problem may be the intrinsic and complicated non-linearity of partitions of integers.

### 2.9.3.2 Interpretation of successive terms

The principal reason for which so little work is done on evaluating zonal polynomials may not even be the numerical awkwardness as adumbrated above, so much as the fact that the successive terms in the hypergeometric expansions have no logical interpretation *per se*. In the Edgeworth expansion, for instance, the first terms have an interpretation. The first term corrects for the non-zero third cumulant, the following for the non-zero fourth cumulant; and these cumulants have standard

interpretations in terms of the characteristics of the distributions involved. In some sense one is compensating for increasing levels of non-normality, even if one usually takes the Edgeworth expansion no further, since it does not in general converge.

One would prefer an asymptotic expansion from which, despite the frequent failure to converge, one can often get good approximations to the function in question, and often a bound on the error from the first omitted term in the series. Working with asymptotic expansions in more than one dimension may admittedly be a little more daunting; and there would still be the problem of meaningful statistical interpretation of successive terms.

#### 2.9.4 Size of matrices

Given the few attempts to calculate zonal polynomials numerically, naively one might have thought the problem to be that the order of the matrix argument in the zonal polynomial increased with sample size. Given the rapid increase in  $p(n)$  as  $n \rightarrow \infty$ , even a very modest sample size would then have made the calculation of zonal polynomials prohibitive. But this seems not to be the problem, in statistical applications at least.

The order of the matrices seems not to go up with sample size, at least while one deals with sums of squares/products types of matrices. An example is afforded by the analysis of variance, in which both the “Between” sum of squares and the “Within” sum of squares involve sums of squares and products: both quantities estimate a variance matrix under different scenarios, and the variance matrix has dimension the number of parameters estimated. It is a similar story for the archetypal inference problems producing density functions which are expanded in series of zonal polynomials, since most of them involve estimates of one covariance matrix or another.

To be sure, the number  $k$  of parameters can be large, especially for estimators in

econometric systems of equations:  $p(k)$  is then quite large, behaving as

$$\frac{1}{4k\sqrt{3}} \exp\left(\pi\sqrt{\frac{2k}{3}}\right)$$

for large  $k$  (e.g., Andrews (1998, p. 70)).

On the other hand,  $n \times n$  determinants can be calculated in  $O(n^{2.39})$  time, according to Grötschel & Lovasz (1995, §10), and the determinants arising in Chapter 3 are almost diagonal; they are moreover extremely sparse, according to results in §4.2 on p. 101. The size of the matrices arising in practice should not prove too great a barrier to the application of the determinantal methods discussed in this thesis.

## Chapter 3

# Determinantal expansions of $J_\rho$ in terms of $e_\lambda$ and $m_\lambda$ functions, and their interrelationship

### 3.1 Preliminary

#### 3.1.1 Introduction

In this chapter we find the matrices  $\Omega^m$  and  $\Omega^e$  in (1.1) on p. 13, and exhibit their interrelationship. In chapters 4 and 5 we shall explore the structure of these matrices in greater detail.

Given that the  $m_\lambda$  and  $e_\lambda$  functions span the homogeneous symmetric polynomials, one can find the entries in  $\Omega^m$  and  $\Omega^e$  “from first principles” by expanding  $\mathcal{L}m_\kappa$  in terms of the  $m_\lambda$  functions, and  $\mathcal{L}e_\kappa$  in terms of the  $e_\lambda$  functions. This was the approach taken in Roberts (1998). There the single value  $\alpha = 2$  was used, but in fact the outworking is almost identical for general  $\alpha$ .

Lapointe et al. (2000) have provided the analogue of the  $\Omega^m$  operator matrix for the Calogero-Sutherland operator in mathematical physics, say  $\Omega_{CS}^m$ ; they have also expressed  $J_\rho$  as a determinantal expansion based on  $\Omega_{CS}^m$ . Roberts (2001) shows that for homogeneous symmetric polynomials (but not for general functions), the Calogero-Sutherland and the Laplace-Beltrami operators differ by a constant, i.e. that

$$\Omega_{CS}^m = \Omega^m + kI. \quad (3.1)$$

With  $\Omega^m$  available from Lapointe et al. (2000), Roberts (1998, ch. 2) and (effectively) in Macdonald (1995, p. 327, Ex. 3(c)); and the interrelationship between  $\Omega^m$  and  $\Omega^e$  derived in Roberts (2001) and provided in §3.3 below, one could infer  $\Omega^e$  from  $\Omega^m$ . Aiming at a self contained exposition we reverse the process, deriving  $\Omega^e$  and thereby deducing  $\Omega^m$ .

In the first part of the chapter, we give the basic form of the determinantal expansion of  $J_\lambda$ , following Roberts (2001). This sets the framework for the whole thesis.

We then show the interrelationship between  $\Omega^m$  and  $\Omega^e$ , again following Roberts (2001); and we find the  $\Omega^e$  matrix from first principles, simplifying the proof in Roberts (1998). Finally, we show consistency of the resulting matrix  $\Omega^m$  deduced from  $\Omega^e$  in this way, with the  $\Omega^m$  matrix as given in Lapointe et al. (2000) and Roberts (1998, ch. 2).

### 3.1.2 Overview

Following Stanley (1989, p. 84), and letting  $D_i = \partial/\partial x_i$ , we define the operator:

$$\mathcal{L}^{(\alpha)} = \mathcal{L} = \frac{\alpha}{2} \sum_{i=1}^n x_i^2 D_i^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^n \frac{x_i^2}{x_i - x_j} D_i \quad (3.2)$$

which we shall call the quasi Laplace-Beltrami operator. The operator will usually be denoted by  $\mathcal{L}$ , except in §3.3 where the more precise notation  $\mathcal{L}^{(\alpha)}$  is called for.  $J_\lambda$  is an eigenfunction of  $\mathcal{L}$ :

$$\mathcal{L} J_\lambda = c_\lambda J_\lambda, \quad (3.3)$$

with eigenvalue

$$c_\lambda = \left(n - \frac{\alpha}{2}\right) w - \sum j l_j + \frac{\alpha}{2} \sum l_j^2. \quad (3.4)$$

We shall occasionally write  $c_\lambda^{(\alpha)}$  for  $c_\lambda$ .

The relation  $\mathcal{L}M = \Omega^m M$  breaks down into:

$$\mathcal{L} m_\lambda = c_\lambda m_\lambda + \sum_{\lambda > \kappa} (\Omega_u^m)_{\lambda\kappa} m_\kappa.$$

It will turn out that  $\Omega_u^m$  is upper triangular with zero diagonal, as anticipated in the range given for the summation, and in addition does not depend on  $\alpha$ . In similar vein  $\mathcal{L}E = \Omega^e E$  decomposes into:

$$\mathcal{L} e_\lambda = c_\lambda e_\lambda + \sum_{\kappa > \lambda} (\Omega_l^e)_{\lambda\kappa} e_\kappa.$$

The connection between  $\Omega^e$  and  $\Omega^m$  is given by

$$\Omega_l^e{}^T = -\alpha \Omega_u^m,$$

thereby justifying the suffix  $l$ , since  $\Omega_l^e$  is lower triangular.

Roberts (2001) shows that similar equations hold for the Calogero-Sutherland operator, with the eigenvalue adjusted by the constant  $k$  in (3.1). That is, the expansion of  $J_\lambda$  in terms of  $e_\lambda$  and  $m_\lambda$  functions can be obtained from either the Laplace-Beltrami or the Calogero-Sutherland operator: one merely adjusts the eigenvalue by the constant  $k$  in all the formulae occurring in this chapter.

## 3.2 Determinantal forms of $J_\rho$ in $e_\lambda$ and $m_\lambda$ functions

### 3.2.1 Preliminary technical results

#### Lemma 3.1

Let

$$J_\rho = \sum_{\kappa} j_{\rho\kappa}^m m_\kappa = \sum_{\kappa} j_{\rho\kappa}^e e_\kappa. \quad (3.5)$$

Then  $j_{\rho\kappa}^m = 0$  unless  $\rho \geq \kappa$ ; and  $j_{\rho\kappa}^e = 0$  unless  $\kappa \geq \rho'$ . Moreover  $j_{\rho\rho}^m \neq 0$  and  $j_{\rho\rho'}^e \neq 0$ .

### Proof 3.1

The statements for the  $m_\lambda$  functions are well-known (e.g. Macdonald (1995, pp. 326, 379), Stanley (1989, p. 77)). Given that  $\lambda > \kappa$  iff  $\kappa' > \lambda'$ , the conclusions for the  $e_\lambda$  functions then follow directly by noting (from Takemura (1984, p. 43) or Macdonald (1995, p. 20), *i.a.*) that, for suitable constants  $v_{\kappa\sigma}$ :

$$m_\kappa = e_\kappa + \sum_{\sigma > \kappa'} v_{\kappa\sigma} e_\sigma. \quad \blacksquare$$

### Lemma 3.2

If  $\kappa > \lambda$ , then there is a chain  $\kappa = \kappa_0 \overset{d2}{>}_1 \kappa_1 \overset{d2}{>}_1 \kappa_2 \dots \overset{d2}{>}_1 \kappa_r = \lambda$ .

That is, for each  $q$ , there are  $i$  and  $j$  such that

$$\kappa_q = (k_1, \dots, k_i, \dots, k_j, \dots) \quad \text{and} \quad \kappa_{q+1} = (k_1, \dots, k_i - 1, \dots, k_j + 1, \dots)$$

where the entries other than the  $i$ th and  $j$ th are unaltered.

### Proof 3.2

Let  $i$  be the minimum index such that  $k_i > l_i$  and  $k_i > k_{i+1}$ . Once  $i$  has been fixed, set  $j$  to be the minimum index such that  $k_i > k_j + 1$ . Then, setting  $\kappa^* = (k_1, \dots, k_{i-1}, k_i - 1, k_{i+1}, \dots, k_{j-1}, k_j + 1, k_{j+1}, \dots)$ , one possible choice for  $\kappa_1$  is  $\kappa^*$ . It is straightforward to show that either  $\kappa > \kappa^* > \lambda$  or  $\kappa > \kappa^* = \lambda$ . In the former case the process can be continued. ■

### Theorem 3.3

Given the  $d2$ -1 chain  $\kappa = \kappa_0 \overset{d2}{>}_1 \kappa_1 \overset{d2}{>}_1 \kappa_2 \dots \overset{d2}{>}_1 \kappa_r = \lambda$  of length  $r$ , then

$$c_\kappa - c_\lambda \geq r(\alpha + 1)$$

### Proof 3.3

For any  $\kappa$  and  $\lambda$ , from (3.4) one has that

$$c_\kappa - c_\lambda = \sum j(l_j - k_j) + \frac{\alpha}{2} \sum (k_j - l_j)(k_j + l_j) \quad (3.6)$$

When  $\kappa \stackrel{d2}{>}_1 \lambda$ , the first summation in the right side of (3.6) is  $\geq 1$  and the second is  $\geq 2$ .

### Corollary 3.3

$$\kappa > \lambda \Rightarrow c_\kappa > c_\lambda.$$

The converse to Corollary 3.3 is invalid, as shown in the next example.

### Example 3.4

When  $\kappa \not> \lambda$ , then  $c_\kappa > c_\lambda$ ,  $c_\kappa = c_\lambda$  and  $c_\kappa < c_\lambda$  are all possible.

The instances of  $\kappa$  and  $\lambda$  below are such that  $\kappa \stackrel{R}{>} \lambda$  and  $\kappa \not> \lambda$ ; and the “eigenvalues” shown are in fact  $\tilde{c}_\tau = c_\tau - (n - \frac{\alpha}{2})w$ :

$$\kappa = (5, 5, 1, 1), \lambda = (4, 4, 4) \Rightarrow \tilde{c}_\kappa = -22 + \frac{\alpha}{2} 52 > \tilde{c}_\lambda = -24 + \frac{\alpha}{2} 48$$

$$\kappa = (4, 1, 1), \lambda = (3, 3) \Rightarrow \tilde{c}_\kappa = -9 + \frac{\alpha}{2} 18 = \tilde{c}_\lambda = -9 + \frac{\alpha}{2} 18$$

$$\kappa = (4, 1, 1, 1), \lambda = (3, 3, 1) \Rightarrow \tilde{c}_\kappa = -13 + \frac{\alpha}{2} 19 < \tilde{c}_\lambda = -12 + \frac{\alpha}{2} 19$$

### 3.2.2 $J_\rho$ as a determinantal form in $e_\lambda$ functions

Let  $J_\rho = \sum_\kappa j_{\rho, \kappa}^e e_\kappa = j_\rho^e{}^T E$  in an obvious notation. Setting  $j = j_\rho^e$  in this subsection, one has

$$\mathcal{L} J_\rho = \mathcal{L} j^T E = j^T \mathcal{L} E = j^T \Omega^e E.$$

But from (3.3) one has

$$\mathcal{L} J_\rho = c_\rho J_\rho = c_\rho j^T E$$

so that

$$j^T (\Omega^e - c_\rho I) E = 0.$$

The  $e_\lambda$  functions are linearly independent, so that

$$j^T (\Omega^e - c_\rho I) = 0 . \quad (3.7)$$

Now, from Lemma 3.1,  $j_{\rho,\kappa}^e = 0$  unless  $\kappa \geq \rho'$ . In the matrix  $\Omega^e - c_\rho I$ , we therefore omit all rows and columns indexed by partitions  $\sigma$  such that

$$\rho' \stackrel{R}{>} \sigma ; \quad \text{or} \quad \sigma \stackrel{R}{>} \rho' \quad \text{and} \quad \sigma \not\geq \rho' .$$

The matrix resulting from these deletions is  $\Omega_{\rho,0}^e$ . It is lower triangular, with zero in the last diagonal element, but no further zeroes along the diagonal, by virtue of Corollary 3.3. The vector of coefficients  $j = j_\rho^e$  is likewise reduced, albeit without change of notation; consistent with this, the vector  $E$  is truncated, also without changing notation.

In fact the last column of  $\Omega_{\rho,0}^e$  is zero. One can utilise the final diagonal element to normalise  $J_\lambda$ , or one can insert  $E$  into that vacuous final column. Setting  $z$  to be a column vector of zeroes save for unity in the final position, we have respectively:

$$j^T \Omega_{\rho,2}^e = j^T (\Omega_{\rho,0}^e + (0|z)) = (0, 0, 0, \dots, 0, N) = N z^T$$

where  $j_{\rho,\rho'}^e = N$  and the 0 in  $(0|z)$  is a zero matrix of the appropriate order; and

$$j^T \Omega_{\rho,1}^e = j^T (\Omega_{\rho,0}^e + (0|E)) = J_\rho z^T .$$

With a non-zero final diagonal element, the matrices  $\Omega_{\rho,1}^e$  and  $\Omega_{\rho,2}^e$  are non-singular from Corollary 3.3, and

$$j^T = N z^T (\Omega_{\rho,2}^e)^{-1} = J_\rho z^T (\Omega_{\rho,1}^e)^{-1} .$$

The two inverse matrices have proportional final rows, since the cofactors are identical. Therefore

$$J_\rho = \frac{N}{\det(\Omega_{\rho,2}^e)} \det(\Omega_{\rho,1}^e) .$$

### 3.2.3 $J_\rho$ as a determinantal form in $m_\lambda$ functions

Now letting  $j = j_\rho^m$ , a similar argument to that in §3.2.2 leads to the analogue of (3.7):

$$j^T (\Omega^m - c_\rho I) = 0 . \quad (3.8)$$

From Lemma 3.1 again,  $j_{\rho,\kappa}^m = 0$  unless  $\rho \geq \kappa$ . In the matrix  $\Omega^{lm} - c_\rho I$ , we therefore omit all rows and columns indexed by partitions  $\sigma$  such that

$$\sigma \stackrel{R}{>} \rho; \quad \text{or} \quad \rho \stackrel{R}{>} \sigma \quad \text{and} \quad \rho \not\geq \sigma.$$

The matrix resulting from these deletions is  $\Omega_{\rho,0}^m$ . It is upper triangular, with zero in the first diagonal element, but no further zeroes along the diagonal, by Corollary 3.3. Again the vector of coefficients  $j_\rho^m$  is abbreviated without change of notation, as is  $M$ .

Setting  $y$  to be a column vector of zeroes save for unity in the first position, we have respectively:

$$j^T \Omega_{\rho,2}^m = j^T (\Omega_{\rho,0}^m + (y|0)) = N y^T$$

where  $j_{\rho,\rho}^m = N$ ; and

$$j^T \Omega_{\rho,1}^m = j^T (\Omega_{\rho,0}^m + (M|0)) = J_\rho y^T.$$

With non-zero diagonals, the matrices  $\Omega_{\rho,1}^m$  and  $\Omega_{\rho,2}^m$  are non-singular, and

$$j^T = N y^T (\Omega_{\rho,2}^m)^{-1} = J_\rho y^T (\Omega_{\rho,1}^m)^{-1}.$$

The two inverse matrices have proportional first rows, since the cofactors are identical. We may thus write

$$J_\rho = \frac{N}{\det(\Omega_{\rho,2}^m)} \det(\Omega_{\rho,1}^m).$$

Altogether we have established the following theorem.

### Theorem 3.5

Using the above notation we have

$$J_\rho = \frac{j_{\rho,\rho'}^e}{\det(\Omega_{\rho,2}^e)} \det(\Omega_{\rho,1}^e) = \frac{j_{\rho,\rho}^m}{\det(\Omega_{\rho,2}^m)} \det(\Omega_{\rho,1}^m). \quad (3.9)$$

■

### 3.2.4 Shape of determinantal expansions

Obtaining the expansion of  $J_\rho$  from determinants as in the first equation of (3.9) is illustrated schematically by the diagrams in Figure 3.1.

$$J_{(3,2)} = \begin{vmatrix} d & . & . & . & e_{(5)} & . & . \\ x & d & . & . & e_{(4,1)} & . & . \\ x & x & d & . & e_{(3,2)} & . & . \\ x & x & x & d & e_{(3,1^2)} & . & . \\ x & x & x & x & e_{(2^2,1)} & . & . \\ . & . & . & . & . & 1 & . \\ . & . & . & . & . & . & 1 \end{vmatrix} \quad J_{(5)} = \begin{vmatrix} d & . & . & . & . & . & e_{(5)} \\ x & d & . & . & . & . & e_{(4,1)} \\ x & x & d & . & . & . & e_{(3,2)} \\ x & x & x & d & . & . & e_{(3,1^2)} \\ x & x & x & x & d & . & e_{(2^2,1)} \\ x & x & x & x & x & d & e_{(2,1^3)} \\ x & x & x & x & x & x & e_{(1^5)} \end{vmatrix}$$

Figure 3.1:  $\det(\Omega_{(3,2),1}^e)$  and  $\det(\Omega_{(5),1}^e)$ 

The dots in the diagrams indicate zeroes. For higher  $w(\rho)$ , the triangular matrices containing the elements  $x$  are sparse, as discussed in §4.2. For  $\Omega_{\rho,2}^e$ , the columns of  $e$  functions are replaced by zeroes, save for a normalising constant on the diagonal.

Similarly the expansion in determinantal form of  $J_\rho$  in terms of  $m_\lambda$  functions is illustrated by Figure 3.2.

$$J_{(3,2)} = \begin{vmatrix} 1 & . & . & . & . & . & . \\ . & 1 & . & . & . & . & . \\ . & . & m_{(3,2)} & x & x & x & x \\ . & . & m_{(3,1^2)} & d & x & x & x \\ . & . & m_{(2^2,1)} & . & d & x & x \\ . & . & m_{(2,1^3)} & . & . & d & x \\ . & . & m_{(1^5)} & . & . & . & d \end{vmatrix} \quad J_{(5)} = \begin{vmatrix} m_{(5)} & x & x & x & x & x & x \\ m_{(4,3)} & d & x & x & x & x & x \\ m_{(3,2)} & . & d & x & x & x & x \\ m_{(3,1^2)} & . & . & d & x & x & x \\ m_{(2^2,1)} & . & . & . & d & x & x \\ m_{(2,1^3)} & . & . & . & . & d & x \\ m_{(1^5)} & . & . & . & . & . & d \end{vmatrix}$$

Figure 3.2:  $\det(\Omega_{(3,2),1}^m)$  and  $\det(\Omega_{(5),1}^m)$

### 3.3 Symmetry between the operator matrices $\Omega^m$ and $\Omega^e$

#### 3.3.1 Preliminary technical material

We take as given the inner product  $\langle \cdot, \cdot \rangle_\alpha$  used to define the Jack polynomials, with respect to which they are mutually orthogonal (see Stanley (1989, p. 77), Macdonald (1995, Ch. VI, §§1, 10)): this inner product is given explicitly in (2.6) on p. 28. The notation and treatment here follow that in Macdonald (1995, p. 378).

With respect to  $\langle \cdot, \cdot \rangle_\alpha$ , the set of functions  $\{g_\mu^{(\alpha)}\}$  is defined as those which are dual to the  $m_\lambda$  functions. That is,

$$\langle g_\mu^{(\alpha)}, m_\lambda \rangle_\alpha = \delta_{\mu\lambda}.$$

The operator  $\omega_\alpha$  is defined as

$$\omega_\alpha g_\mu^{(\alpha)} = e_\mu, \quad (3.10)$$

and satisfies the following identity:

$$\omega_\alpha \omega_{1/\alpha} = 1.$$

#### 3.3.2 The principal theorem

The following result is a direct analogue of Roberts (1998, thm. 3.6).

##### Theorem 3.6

*Define operator matrices  $\Omega$  as follows:*

$$\mathcal{L}^{(\alpha)} M = \Omega^{lm}(\alpha) M \quad \mathcal{L}^{(\alpha)} E = \Omega^{le}(\alpha) E \quad (3.11)$$

*and break them up into diagonal and off-diagonal portions*

$$\Omega^{lm}(\alpha) = D^{lm}(\alpha) + \Omega_u^m(\alpha) \quad \Omega^{le}(\alpha) = D^{le}(\alpha) + \Omega_l^e(\alpha),$$

where  $D$  matrices are diagonal; and  $\Omega_l^e(\alpha)$  and  $\Omega_u^m(\alpha)$  have zero diagonals.

Then  $\Omega_u^m(\alpha)$  is upper triangular, and does not depend on  $\alpha$ . Writing  $\Omega_u^m(\alpha)$  as  $\Omega_u^m$ , we have moreover that

$$\Omega_l^e(\alpha) = -\alpha \Omega_u^m{}^T.$$

### Proof 3.6

From Theorem 3.9, the operator matrix  $\Omega^e(\alpha)$  is lower triangular, and the off-diagonal terms contain a factor of  $\alpha$ . The proof is then immediate from Lemma 3.8, noting that the term involving the identity matrix in (3.13) naturally has no impact off the diagonal. ■

### Lemma 3.7

$$-\alpha \mathcal{L}^{(1/\alpha)} = \omega_\alpha \mathcal{L}^{(\alpha)} \omega_{1/\alpha} - (1 + \alpha)k_1 \quad (3.12)$$

where  $k_1 = w(n - 1)$ .

### Proof 3.7

Macdonald (1995, p. 320) defines an operator  $\square^\alpha$ , which he calls the Laplace-Beltrami operator, as

$$\square^\alpha = \mathcal{L}^{(\alpha)} - w(n - 1) = \mathcal{L}^{(\alpha)} - k_1.$$

From Macdonald (1995, p 330, Ex. 3) one has

$$\omega_\alpha \square^\alpha + \alpha \square^{1/\alpha} \omega_\alpha = 0,$$

or equivalently

$$-\alpha \square^{1/\alpha} = \omega_\alpha \square^\alpha \omega_{1/\alpha}.$$

Now the left side of (3.12) becomes

$$-\alpha \mathcal{L}^{(1/\alpha)} = -\alpha (\square^{1/\alpha} + k_1) = \omega_\alpha \square^\alpha \omega_{1/\alpha} - k_1 \alpha = \omega_\alpha \mathcal{L}^{(\alpha)} \omega_{1/\alpha} - k_1(1 + \alpha),$$

agreeing with the right side. ■

### Lemma 3.8

$$\Omega^e(\alpha) = -\alpha \Omega^m(1/\alpha)^T + k_1(1 + \alpha)I \quad (3.13)$$

**Proof 3.8**

From (3.11) we may write

$$\mathcal{L}^{(\alpha)} m_\kappa = \sum_\rho (\Omega^{lm}(\alpha))_{\kappa\rho} m_\rho. \quad (3.14)$$

The fact that the  $\mathcal{L}^{(\alpha)}$  operator is self-adjoint with respect to the inner product  $\langle \cdot, \cdot \rangle_\alpha$  (see e.g. Vilenkin & Klimyk (1995, p. 112), or Stanley (1989, p. 84)) allows us to write

$$\langle \mathcal{L}^{(\alpha)} g_\lambda^{(\alpha)}, m_\kappa \rangle_\alpha = \langle g_\lambda^{(\alpha)}, \mathcal{L}^{(\alpha)} m_\kappa \rangle_\alpha = (\Omega^{lm}(\alpha))_{\kappa\lambda}, \quad (3.15)$$

from which the following is immediate:

$$\mathcal{L}^{(\alpha)} g_\lambda^{(\alpha)} = \sum_\kappa (\Omega^{lm}(\alpha))_{\kappa\lambda} g_\kappa^{(\alpha)}. \quad (3.16)$$

The relation (3.10) and Lemma 3.7 imply

$$-\alpha \mathcal{L}^{(1/\alpha)} e_\lambda = \omega_\alpha \mathcal{L}^{(\alpha)} \omega_{1/\alpha} e_\lambda - k_1(1+\alpha)e_\lambda = \omega_\alpha \mathcal{L}^{(\alpha)} g_\lambda^{(\alpha)} - k_1(1+\alpha)e_\lambda. \quad (3.17)$$

From (3.16) and (3.10), we have

$$\begin{aligned} -\alpha \mathcal{L}^{(1/\alpha)} e_\lambda &= \omega_\alpha \sum_\kappa (\Omega^{lm}(\alpha))_{\kappa\lambda} g_\kappa^{(\alpha)} - k_1(1+\alpha)e_\lambda \\ &= \sum_\kappa (\Omega^{lm}(\alpha))_{\kappa\lambda} e_\kappa - k_1(1+\alpha)e_\lambda. \end{aligned} \quad (3.18)$$

Interchanging  $\alpha$  and  $1/\alpha$ , this expression is equivalent to

$$\mathcal{L}^{(\alpha)} e_\lambda = -\alpha \sum_\kappa (\Omega^{lm}(1/\alpha))_{\kappa\lambda} e_\kappa + k_1(1+\alpha)e_\lambda \equiv \sum_\kappa (\Omega^{le}(\alpha))_{\lambda\kappa} e_\kappa,$$

whence (3.13) follows. ■

### 3.4 The action of $\mathcal{L}$ on the $e_\lambda$ functions

**Theorem 3.9**

Let  $\lambda = (l_1, l_2, \dots, l_r)$ , with  $r = \ell(\lambda)$ . Then

$$\mathcal{L}^{(\alpha)} e_\lambda = c_{\lambda'} e_{\lambda'} - \alpha \sum_{\substack{j,k=1 \\ j>k}}^r \sum_{v=1}^{l_j} (l_k - l_j + 2v) e_{(\dots, l_k+v, \dots, l_j-v, \dots)} \quad (3.19)$$

in which the partition in the final subscript differs from  $\lambda$  only in the  $k$ th and  $j$ th positions, and may not be in the standard weakly decreasing order.

### Proof 3.9

From Theorem 3.12, one has that

$$\mathcal{L} e_\lambda = \sum_{i=1}^r \left( \prod_{\substack{j=1 \\ j \neq i}}^r e_{l_j} \right) \mathcal{L} e_{l_i} + \alpha \sum_{m=1}^n \sum_{\substack{j,k=1 \\ j > k}}^r \left( \prod_{\substack{i=1 \\ i \neq j,k}}^r e_{l_i} \right) x_m^2 \frac{\partial e_{l_j}}{\partial x_m} \frac{\partial e_{l_k}}{\partial x_m} \quad (3.20)$$

while from Theorem 3.17, and for any  $s$ ,

$$\mathcal{L} e_s = c_{(s)'} e_s = c((s)') e_s.$$

Equation (3.20) may now be written as

$$\begin{aligned} \mathcal{L} e_\lambda &= \left( \sum_{i=1}^r c((l_i)') \right) e_\lambda + \alpha \sum_{\substack{j,k=1 \\ j > k}}^r \left( \prod_{\substack{i=1 \\ i \neq j,k}}^r e_{l_i} \right) \sum_{m=1}^n x_m^2 e_{l_{j-1}}(\widehat{x_m}) e_{l_{k-1}}(\widehat{x_m}) \\ &= \left( \sum_{i=1}^r c((l_i)') \right) e_\lambda + \alpha \sum_{\substack{j,k=1 \\ j > k}}^r \left( \prod_{\substack{i=1 \\ i \neq j,k}}^r e_{l_i} \right) \left( l_j e_{(l_k, l_j)} - \sum_{u=0}^{l_j-1} (l_k + l_j - 2u) e_{(l_k + l_j - u, u)} \right) \end{aligned}$$

in which for the last equation we have applied Theorem 3.16. An application of Lemma 3.11 and the substitution of  $v = l_j - u$  complete the proof. ■

### 3.4.1 Technical support for the calculation of $\mathcal{L} e_\lambda$

Lemmas 3.10 and 3.11 have to do with values of the eigenvalues  $c_\lambda$  and  $c_{\lambda'}$ , and are used in the proofs of Theorems 3.9 and 3.17. Theorems 3.12, 3.16 and 3.17 are input to Theorem 3.9.

#### Lemma 3.10

Let  $\lambda = (l_1, l_2, \dots, l_r)$ . Then

$$c_{\lambda'} = \left( n - \alpha - \frac{1}{2} \right) w - \frac{1}{2} \sum_{j=1}^r l_j^2 + \alpha \sum_{j=1}^r j l_j. \quad (3.21)$$

**Proof 3.10**

From (3.4) we have

$$c_{\lambda'} = \left(n - \frac{\alpha}{2}\right) w - \sum j l'_j + \frac{\alpha}{2} \sum (l'_j)^2.$$

Noting that the multiplicity of  $j$  in  $\lambda'$  is  $l_j - l_{j+1}$ , it is tedious but straightforward to verify that

$$\sum_{j=1}^r j l'_j = \frac{1}{2} \left( \sum_{j=1}^r l_j^2 + w \right)$$

and

$$\sum (l'_j)^2 = 2 \sum_{j=1}^r j l_j - w,$$

whence the result follows. ■

**Lemma 3.11**

$$\sum_{j=1}^r c((l_j)') + \alpha \sum_{j=1}^r (j-1) l_j = c_{\lambda'}. \quad (3.22)$$

**Proof 3.11**

From Lemma 3.10 one may write

$$c_{(1^t)} = c_{(t)'} = c((t)') = nt - \frac{t}{2}(t+1)$$

which immediately implies (3.22). ■

**Theorem 3.12**

Let  $f = \prod_{i=1}^r f_i$  be the product of functions  $f_i$  of the indeterminate variates  $x_1, \dots, x_n$ .

Further let  $\mathcal{F}$  be a differential operator of the form

$$\mathcal{F} = \sum_{i=1}^n a_i \frac{\partial}{\partial x_i} + \frac{\alpha}{2} \sum_{i=1}^n x_i^2 \frac{\partial^2}{\partial x_i^2}$$

in which the functions  $a_i$  of the variates  $x_i$  do not involve differential operators.

Then

$$\mathcal{F} f = \sum_{i=1}^r \left( \prod_{\substack{j=1 \\ j \neq i}}^r f_j \right) \mathcal{F} f_i + \alpha \sum_{l=1}^n \sum_{\substack{j,k=1 \\ j>k}}^r \left( \prod_{\substack{i=1 \\ i \neq j,k}}^r f_i \right) x_l^2 \frac{\partial f_j}{\partial x_l} \frac{\partial f_k}{\partial x_l}. \quad (3.23)$$

### Proof 3.12

The first order terms combine in the conventional fashion using the product rule for differentiation, as does the second order operator when both differentiations apply to the one element of the product. This leads to the first term of the right side of (3.23). The second order operators however also lead to cross products, as detailed in the final term in (3.23). ■

### Lemma 3.13

For integral  $t > 0$

$$t \binom{s+2t}{t} = \sum_{v=-t}^{-1} (s-2v) \binom{s+2t}{v+t}.$$

The proof by induction on  $t$  is straightforward. At each stage the value of  $s$  is increased by 2 to use at the next stage. ■

In order to motivate the next lemma, we first consider an example.

### Example 3.14

The monomial  $y_1^2 y_2^2 \dots y_a^2 y_{a+1} y_{a+2} \dots y_{a+b}$  does not appear in  $e_{a+b+1} e_{a-1}$ ; it appears once  $= \binom{b}{0}$  times in  $e_{a+b} e_a$ ; it appears  $b = \binom{b}{1}$  times in  $e_{a+b-1} e_{a+1}$ ; etc. The monomial in question has type  $2^a 1^b$ , following the terminology of Stanley (1999, p. 309). The binomial coefficient  $\binom{b}{c}$  is interpreted as zero if  $c < 0$ . ■

### Lemma 3.15

Suppose that  $2a + b = r + s$ . Then a given monomial of type  $2^a 1^b$  appears  $\binom{b}{t}$  times in  $e_{(r,s)} = e_r e_s$ , where  $t = a + b - r = s - a$ . ■

### Theorem 3.16

Let  $j > k$ . Then

$$\sum_{i=1}^n x_i^2 e_{j-1}(\hat{x}_i) e_{k-1}(\hat{x}_i) = k e_j e_k - \sum_{u=0}^{k-1} (j+k-2u) e_{j+k-u} e_u \quad (3.24)$$

in which the circumflex indicates omission of the variate.

### Proof 3.16

We consider a given monomial of type  $2^{k-t} 1^{j-k+2t}$ , and count the number of times

it appears on the left and right sides of (3.24).

From Lemma 3.15, the number of appearances on the right side of (3.24) is

$$\begin{aligned} & k \binom{j-k+2t}{t} - \sum_{u=0}^{k-1} (j+k-2u) \binom{j-k+2t}{u-k+t} \\ &= k \binom{j-k+2t}{t} - \sum_{u=k-t}^{k-1} (j+k-2u) \binom{j-k+2t}{u-k+t} \end{aligned}$$

while the number on the left side of (3.24) is

$$(k-t) \binom{j-k+2t}{t}.$$

The equality of the number of appearances on either side is a direct consequence of Lemma 3.13, setting  $v = u - k$  and  $s = j - k$ . ■

### Theorem 3.17

$$\mathcal{L} e_r = c_{(r)'} e_r.$$

#### Proof 3.17

Only the final term in (3.2) need be considered. First restricting ourselves to the variates  $x_1$  and  $x_2$ , it is easily shown that

$$\left( \frac{x_1^2}{x_1 - x_2} D_1 + \frac{x_2^2}{x_2 - x_1} D_2 \right) e_r = x_1 x_2 e_{r-2} (\widehat{x_1, x_2}) + (x_1 + x_2) e_{r-1} (\widehat{x_1, x_2}) \quad (3.25)$$

where the extended circumflex again indicates missing variables.

Summing over all possible choices of pairs of variates plainly gives a multiple of  $e_r$ . Counting the number of monomials in the right side of (3.25) yields

$$\binom{n-2}{r-2} \binom{n}{2} + 2 \binom{n-2}{r-1} \binom{n}{2}.$$

Dividing this quantity by  $\binom{n}{r}$ , the number of monomials in  $e_r$ , yields

$$\mathcal{L} e_r = r \left( n - \frac{r+1}{2} \right) e_r.$$

An appeal to Lemma 3.10 completes the proof. ■

### 3.5 Consistency of $\Omega^e$ and $\Omega^m$ derived by different means

We now show the consistency of  $\Omega^e$  as provided in Theorem 3.9; and  $\Omega^m$  as given in Lapointe et al. (2000), Roberts (1998, thm. 2.4), and as a closely related exercise, in Macdonald (1995, p. 327, Ex. 3(c)). We first state the form of  $\Omega^m$ .

#### Theorem 3.18

Let  $\lambda \vdash w$ ,  $\ell(\lambda) = r$ , and write  $\lambda$  in the conventional non-increasing order with a trailing zero:  $\lambda = (l_1, l_2, \dots, l_r, 0)$ . Then

$$\mathcal{L} m_\lambda = c_\lambda m_\lambda + \sum_{\substack{l_k, l_j \\ l_k > l_j \geq 0}} \Delta_{l_k, l_j} \sum_{i=1}^{[\Delta_{l_k, l_j}/2]} K_{(\ell_k, \ell_j), i} m_{(\ell_k - i, \ell_j + i) \cup \lambda_2} \quad (3.26)$$

where

$$K_{(\ell_k, \ell_j), i} = \begin{cases} \begin{pmatrix} t_1 \\ 1 \end{pmatrix} \begin{pmatrix} t_2 \\ 1 \end{pmatrix} & \text{when } \ell_k - i > \ell_j + i \\ \begin{pmatrix} t \\ 2 \end{pmatrix} & \text{when } \ell_k - i = \ell_j + i \end{cases} \quad (3.27)$$

The first summation in (3.26) is over all distinct subpartitions  $\lambda_1 = (\ell_k, \ell_j)$ ,  $\ell_k > \ell_j \geq 0$ , such that  $\lambda = \lambda_1 \cup \lambda_2$ ;  $\Delta_{l_k, l_j} = \ell_k - \ell_j$ ;  $t_1$  is the multiplicity of  $\ell_k - i$  in  $(\ell_k - i, \ell_j + i) \cup \lambda_2$ , while  $t_2$  is the multiplicity of  $\ell_j + i$  in  $(\ell_k - i, \ell_j + i) \cup \lambda_2$ ; unless  $\ell_k - i = \ell_j + i$ , in which case  $t$  is the multiplicity of  $\ell_k - i$ .

#### Proof 3.18

The proof in Lapointe et al. (2000) draws on their earlier papers. A self contained proof is in Roberts (1998, p. 19). ■

#### Theorem 3.19

*As regards off-diagonal terms, the expansions of  $\mathcal{L}e_\lambda$  in Theorem 3.9, and that of  $\mathcal{L}m_\lambda$  in Theorem 3.18, are consistent with Theorem 3.6.*

#### Proof 3.19

Choose  $\lambda = (l_k, l_j, \dots)$  and  $\kappa = (l_k + v, l_j - v, \dots)$  where  $\lambda$  and  $\kappa$  differ in exactly

2 elements, and the listings given are not necessarily in standard weakly decreasing order.

Consider  $(\Omega^e)_{\lambda,\kappa}$ , for which  $\lambda$  is the source partition, and within the  $\lambda$ th row one moves left from the  $\lambda$ th column to the  $\kappa$ th column in  $(\Omega^e)$ . Disregarding the factor of  $\alpha$ ,

$$(\Omega^e)_{\lambda,\kappa} = l_k + v - (l_j - v)$$

according to Theorem 3.9.

Now consider  $(\Omega^m)_{\kappa,\lambda}$ . The “source” partition is to be considered  $\kappa$ , not  $\lambda$ , in Theorem 3.18; and one moves within the  $\kappa$ th row right from the  $\kappa$ th column to the  $\lambda$ th column in  $\Omega^m$ . Should the multiplicities of  $l_k$  and  $l_j$  in  $\lambda$  be unity, and still disregarding  $\alpha$ ,

$$(\Omega^m)_{\kappa,\lambda} = \Delta_{l_k+v, l_j-v} = l_k + v - (l_j - v) = l_k - l_j + 2v = (\Omega^e)_{\lambda,\kappa}$$

The factor  $K$  allows for multiple occurrences of  $l_k - i$  and  $l_j + i$  in  $\lambda$ , since these are counted singly in the first summation of (3.26) and multiply in Theorem 3.9. Should these values be equal, the multiple is the number of times the pair of values can be chosen. ■

The argument in the next theorem is analogous to that in Roberts (2001, Thm. 8).

### Theorem 3.20

Given that  $\mathcal{L}m_\rho = \sum_{\rho \geq \kappa} a_{\rho,\kappa} m_\kappa$ , then  $a_{\rho,\rho} = c_\rho$ , where  $c_\rho$  is the eigenvalue of  $\mathcal{L}$  given in (3.4).

### Proof 3.20

We use the notation from Lemma 3.1:

$$J_\rho = \sum_{\rho \geq \kappa} j_{\rho\kappa}^m m_\kappa$$

whence

$$\begin{aligned} \mathcal{L}J_\rho &= \sum_{\rho \geq \kappa} j_{\rho\kappa}^m \mathcal{L}m_\kappa \\ &= \sum_{\rho \geq \kappa} j_{\rho\kappa}^m \sum_{\kappa \geq \tau} a_{\kappa,\tau} m_\tau \end{aligned}$$

$$= j_{\rho\rho}^m a_{\rho\rho} m_\rho + \text{terms of lower order in the RLO} \quad (3.28)$$

But also from (3.3) one has

$$\begin{aligned} \mathcal{L}J_\rho &= c_\rho J_\rho = c_\rho \sum_{\rho \geq \kappa} j_{\rho\kappa}^m m_\kappa \\ &= c_\rho j_{\rho\rho}^m m_\rho + \text{terms of lower order in the RLO} \end{aligned} \quad (3.29)$$

Also from Lemma 3.1 one knows that  $j_{\rho\rho}^m \neq 0$ , whence the result from a comparison of (3.28) and (3.29). ■

## 3.6 Extended example

Before discussing methods for generating the  $\Omega^m$  matrix in Chapter 4, we illustrate the above material with an extended example.

For partitions of weight 4:

$$\begin{aligned} \Omega^m &= \begin{pmatrix} c_{(4)} & 4 & 4 & & \\ & c_{(3,1)} & 2 & 6 & \\ & & c_{(2,2)} & 2 & \\ & & & c_{(2,1,1)} & 12 \\ & & & & c_{(1,1,1,1)} \end{pmatrix} \\ \Omega_{(4),0}^m &= \begin{pmatrix} 0 & 4 & 4 & & \\ & c_{(3,1)} - c_{(4)} & 2 & 6 & \\ & & c_{(2,2)} - c_{(4)} & 2 & \\ & & & c_{(2,1,1)} - c_{(4)} & 12 \\ & & & & c_{(1,1,1,1)} - c_{(4)} \end{pmatrix} \\ J_{(4)} &= \begin{vmatrix} m_{(4)} & 4 & 4 & & \\ m_{(3,1)} & -1 - 3\alpha & 2 & 6 & \\ m_{(2,2)} & & -2 - 4\alpha & 2 & \\ m_{(2,1,1)} & & & -3 - 5\alpha & 12 \\ m_{(1,1,1,1)} & & & & -6 - 6\alpha \end{vmatrix} \end{aligned}$$

The above example conforms with the example given in Lapointe et al. (2000).

$$\Omega_{(2,2),0}^m = \begin{pmatrix} 0 & 2 & & \\ & c_{(2,1,1)} - c_{(2,2)} & 12 & \\ & & c_{(1,1,1,1)} - c_{(2,2)} & \end{pmatrix}$$

$$J_{(2,2)} = \begin{vmatrix} m_{(2,2)} & 2 & & \\ m_{(2,1,1)} & -1 - \alpha & 12 & \\ m_{(1,1,1,1)} & & -4 - 2\alpha & \end{vmatrix}$$

Turning to the  $e_\lambda$  basis functions,

$$\Omega^e = \begin{pmatrix} c_{(1,1,1,1)} & & & & \\ -4\alpha & c_{(2,1,1)} & & & \\ -4\alpha & -2\alpha & c_{(2,2)} & & \\ & -6\alpha & -2\alpha & c_{(3,1)} & \\ & & & -12\alpha & c_{(4)} \end{pmatrix}$$

$$J_{(4)} = \begin{vmatrix} -6 - 6\alpha & & & & e_{(4)} \\ -4\alpha & -3 - 5\alpha & & & e_{(3,1)} \\ -4\alpha & -2\alpha & -2 - 4\alpha & & e_{(2,2)} \\ & -6\alpha & -2\alpha & -1 - 3\alpha & e_{(2,1,1)} \\ & & & -12\alpha & e_{(1,1,1,1)} \end{vmatrix}$$

$$\Omega_{(3,1),0}^e = \begin{pmatrix} c_{(1,1,1,1)} - c_{(3,1)} & & & & \\ -4\alpha & c_{(2,1,1)} - c_{(3,1)} & & & \\ -4\alpha & -2\alpha & c_{(2,2)} - c_{(3,1)} & & \\ & -6\alpha & -2\alpha & 0 & \end{pmatrix}$$

$$J_{(3,1)} = \begin{vmatrix} -5 - 3\alpha & & & e_{(4)} \\ -4\alpha & -2 - 2\alpha & & e_{(3,1)} \\ -4\alpha & -2\alpha & -1 - \alpha & e_{(2,2)} \\ & -6\alpha & -2\alpha & e_{(2,1,1)} \end{vmatrix}$$

## Chapter 4

# The composition of the operator matrices $\Omega^e$ and $\Omega^m$

The  $\Omega^m$  and  $\Omega^e$  matrices can be used directly to obtain the Jack polynomials from the determinants derived in Chapter 3. Alternatively one may prefer to work with the relations

$$J_\rho = \sum_{\lambda} j_{\rho\lambda}^m m_\lambda = \sum_{\lambda} j_{\rho\lambda}^e e_\lambda, \quad (4.1)$$

reproduced from (3.5) in Lemma 3.1.

Mathematically there is of course no distinction between these two approaches, but highlighting the  $j$  coefficients in (4.1) throws those coefficients into relief, and pushes the calculation of  $e_\lambda$  and  $m_\lambda$  functions into the background. This is convenient because it may turn out that only a few of the terms in this expansion dominate; or because when weights of partitions increase, there are relationships connecting the coefficients  $j$  calculated for the different weights, as we shall see in §4.6.

We work chiefly with  $\Omega^m$ , since the off-diagonal elements do not contain  $\alpha$ . We first define the chains by which the off-diagonal elements of  $\Omega^m$  are defined; and then the paths through the elements of  $\Omega^m$  which give the coefficients of the  $e_\lambda$  and  $m_\lambda$  functions in the determinantal expansions. The distinction between paths and chains is convenient, but terminology here is awkward: both paths and chains as

used here are chains in the sense of partially ordered sets.

Consideration of all possible chains enables one to count the number of off-diagonal elements in  $\Omega^m$ , thereby showing that  $\Omega^m$  is highly sparse when partition weights increase.

We then define d2 paths, showing that the number of such paths from  $\rho$  to  $\lambda$ , and the lengths of those paths, are important clues in deciding when one should omit part of the calculations of  $j_{\rho,\lambda}$  coefficients without losing much accuracy. This sort of decision can be placed in the context of powers of incidence matrices.

## 4.1 Derivation of $\Omega^m$ through core chains

A convenient way in which to derive the above-diagonal elements of  $\Omega^m$  is to draw up a chart or matrix of “chains”, commencing from all distinct elements  $l$  of the partitions  $\lambda$ , and generating the chain  $(l), (l-1, 1), (l-2, 2), \dots$

### Definition 4.1

A “core chain” or “chain” is a sequence of length  $r = \lfloor \frac{l}{2} \rfloor$  of partitions, say  $\kappa_0, \kappa_1, \dots, \kappa_r$ , such that  $\kappa_0 = (l) \cup \lambda$ ,  $\kappa_1 = (l-1, 1) \cup \lambda$ ,  $\kappa_2 = (l-2, 2) \cup \lambda, \dots$ . Symbolically, the chain is generated as

$$\kappa_0 = (l) \cup \lambda \rightarrow (l-1, 1) \cup \lambda \rightarrow \dots \rightarrow \begin{cases} (\frac{l}{2}, \frac{l}{2}) \cup \lambda & \text{when } l \text{ is even} \\ (\frac{l+1}{2}, \frac{l-1}{2}) \cup \lambda & \text{when } l \text{ is odd} \end{cases}$$

It is convenient to allow the degenerate case in which  $l = 1$ . ■

Consider the chain matrix  $CM_5$  in Figure 4.1, which lists the d2-1 chains used to obtain the operating matrix when  $w = 5$ . The numbers constituting the first and third elements in each entry above the diagonal arise from Theorem 3.18 on p. 94. The first element is the difference  $l_i - l_j$ ; the letter in between the two numbers identifies the chain by which that entry is generated; while the third element is  $K$  as

$$CM_5 = \begin{pmatrix} \lambda & (5) & (4,1) & (3,2) & (3,1^2) & (2^2,1) & (2,1^3) & (1^5) \\ (5) & a & 5a1 & 5a1 & & & & \\ (4,1) & & a,b,c & 3a1 & 4b2 & 4b1 & & \\ (3,2) & & & d,e & 2e1 & 3d2 & & \\ (3,1^2) & & & & b,f,g & 2b1 & 3f3 & \\ (2^2,1) & & & & & h,i & 2h3 & \\ (2,1^3) & & & & & & j,k & 2j10 \\ (1^5) & & & & & & & l \end{pmatrix}$$

Figure 4.1: The chain matrix  $CM_5$

given in (3.27) on p. 94. The numbers in the entry  $3d2$  in position  $((3,2), (2,2,1))$ , for instance, arise from  $3 = 3 - 0$  and  $2 = \binom{2}{1}\binom{1}{1}$ .

The core chain  $a$ , viz.  $(5,0) \rightarrow (4,1) \rightarrow (3,2)$ , contains the partitions occurring in the expansion of  $\mathcal{L} m_{(5)}$ ; it also contains incidentally the sequence  $(4,1) \rightarrow (3,2)$  occurring in  $\mathcal{L} m_{(4,1)}$ . In the operator matrix  $\Omega_5^m$  there correspond three elements above the diagonal, viz.  $\omega_{(5),(4,1)}^m$  and  $\omega_{(5),(3,2)}^m$  in the first row and  $\omega_{(4,1),(3,2)}^m$  in the second, as indicated in Figure 4.1. Similarly chain  $b$  consists of  $(4,1), (3,1,1), (2,2,1)$ , generated as  $(4) \cup (1) \rightarrow (3,1) \cup (1) \rightarrow (2,2) \cup (1)$ , and gives rise to 3 elements above the diagonal in the operator matrix.

Chains producing just one entry (paths  $c, g, i, k, l$  in this case) spring from the subpartition  $(1,0)$ , and produce no off-diagonal terms in  $\Omega^m$ . Those chains shown with just two elements, such as  $d$  and  $e$ , each give rise to just one element in the  $CM_5$  matrix above the diagonal. In the general case chains with length  $r$  give rise to  $r(r+1)/2$  elements above the diagonal in  $CM_n$ . Any particular element above the diagonal arises from just one chain, and it is easy to see that this is always the case, for any weight  $w$  of the partitions considered.

#### Lemma 4.2

The number of chains in the  $CM_w$  matrix is  $\sum_0^{w-1} p(j)$ , where  $p(0) = 1$ . If we omit

chains containing but one element (those chains for which the source subpartition is  $(1, 0)$ ), the number of chains reduces to  $\sum_0^{w-2} p(j) = \sum_2^w p(w - j)$ .

#### Proof 4.2

The number of chains is the total number of (non-zero) elements in the entire set of partitions of  $w$ , counting only distinct elements within each partition. The number of partitions of  $w$  containing an element 1 is  $p(w - 1)$ , since these partitions are just a partition of  $w - 1$  with a 1 adjoined. Similarly the number of partitions of  $w$  containing an element 2 is  $p(w - 2)$ , etc. ■

#### Theorem 4.3

The number of non-zero elements above the diagonal in  $CM_w$  is

$$\sum_{j=1}^w p(w - j) \frac{[j/2] ([j/2] + 1)}{2} = \sum_{j=2}^w p(w - j) \frac{[j/2] ([j/2] + 1)}{2}.$$

#### Proof 4.3

For an element  $j$  in  $\lambda$ , from Definition 4.1 one has that the number of partitions in the corresponding chain in  $CM_w$  is  $r + 1 = [j/2] + 1 = q$ , say. The number of pairs of elements able to be chosen from these is  $\binom{q}{2}$ : and for each such pair the first can represent the source partition on the diagonal, and the second the target partition above the diagonal in  $CM_w$ . As noted in Proof 4.2,  $p(w - j)$  is the number of chains of length  $j$ . ■

Multiplying the numbers above the diagonal in  $CM_5$  in Figure 4.1, and inserting the Laplace-Beltrami eigenvalues  $c_\lambda$  in the diagonal  $(\lambda, \lambda)$ th position, leads to the operator matrix  $\Omega_5^m$ , shown in Figure 4.2.

## 4.2 Sparsity of $\Omega_w$

The matrix  $\Omega^m$  is highly sparse, in that the proportion of non-zero elements above the diagonal tends quickly to zero as  $w \rightarrow \infty$ .

$\Omega_5^m =$ 

	(5)	(4, 1)	(3, 2)	(3, 1, 1)	(2, 2, 1)	(2, 1, 1, 1)	(1, 1, 1, 1, 1)
(5)	$c_{(5)}$	5	5	0	0	0	0
(4, 1)	0	$c_{(4,1)}$	3	8	4	0	0
(3, 2)	0	0	$c_{(3,2)}$	2	6	0	0
(3, 1 <sup>2</sup> )	0	0	0	$c_{(3,1,1)}$	2	9	0
(2 <sup>2</sup> , 1)	0	0	0	0	$c_{(2,2,1)}$	6	0
(2, 1 <sup>3</sup> )	0	0	0	0	0	$c_{(2,1,1,1)}$	20
(1 <sup>5</sup> )	0	0	0	0	0	0	$c_{(1,1,1,1,1)}$

Figure 4.2: The operator matrix  $\Omega_5^m$

Lemma 4.4

Let  $0 < A < B$ , and let  $t_{(j)} = t(t-1) \dots (t-j+1)$  denote the falling factorial, with  $t_{(0)} = 1$ . Then

$$\int_A^B x^s e^{K\sqrt{x}} dx = 2 \sum_{j=0}^\infty \frac{(-1)^j t_{(j)}}{K^{j+1}} \left( B^{(t-j)/2} e^{K\sqrt{B}} - A^{(t-j)/2} e^{K\sqrt{A}} \right)$$

in which  $t = 2s + 1$ , and  $K > 0$  is a constant.

Proof 4.4

The proof consists of substituting  $y = \sqrt{x}$  and integrating by parts. The series terminates, since  $t$  is a non-negative integer in applications of the lemma in this paper. ■

Lemma 4.5

As  $w \rightarrow \infty$ , and for  $w_0 > K^{-2} + 1$ ,

$$\sum_{j=w_0}^{w-2} (w-j) \left( \frac{w-j}{2} + 1 \right) \frac{e^{K\sqrt{j}}}{\sqrt{j}} = e^{K\sqrt{w}} \left( \frac{8w}{K^3} + O(\sqrt{w}) \right) \tag{4.2}$$

**Proof 4.5**

For  $x > K^{-2}$  and  $r \geq 0$ , the function

$$x^{r-1/2} e^{K\sqrt{x}}$$

is non-decreasing in  $x$ . For  $w_0 - 1 > K^{-2}$ , then,

$$\int_{w_0-1}^{w-2} x^{r-1/2} e^{K\sqrt{x}} dx \leq \sum_{j=w_0}^{w-2} j^{r-1/2} e^{K\sqrt{j}} \leq \int_{w_0}^{w-1} x^{r-1/2} e^{K\sqrt{x}} dx \quad (4.3)$$

Set  $H_r = \sum_{j=w_0}^{w-2} j^{r-1/2} e^{K\sqrt{j}}$ . Then, as  $w \rightarrow \infty$ , from (4.3) and Lemma 4.4 we have that

$$\begin{aligned} H_0 &= e^{K\sqrt{w}} \left( \frac{2}{K} + o(1) \right), \\ H_1 &= e^{K\sqrt{w}} \left( \frac{2w}{K} - \frac{4\sqrt{w}}{K^2} + \frac{4}{K^3} + o(1) \right) \quad \text{and} \\ H_2 &= e^{K\sqrt{w}} \left( \frac{2w^2}{K} - \frac{8w^{3/2}}{K^2} + \frac{24w}{K^3} + O(\sqrt{w}) \right) \end{aligned}$$

One can rewrite the left hand side of (4.2) as

$$\left( \frac{w^2}{2} + w \right) H_0 - (w+1)H_1 + \frac{1}{2} H_2,$$

which simplifies to the right hand side. ■

**Theorem 4.6**

Set  $K = \pi\sqrt{2/3}$ . As  $w \rightarrow \infty$ , the fraction  $f_w$  of the elements above the diagonal of the  $\Omega_w$  matrix which are non-zero satisfies

$$f_w = O\left(\frac{384w^3}{K^3} e^{-K\sqrt{w}}\right).$$

**Proof 4.6**

From Theorem 4.3, the fraction of non-zero elements in the top half of  $\Omega_w$  is

$$f_w = \frac{\sum_{j=2}^w \frac{[j/2]([j/2] + 1)}{2} p(w-j)}{\frac{p(w)(p(w) - 1)}{2}}$$

$$\leq \frac{\sum_{j=2}^w \frac{j}{2} \left( \frac{j}{2} + 1 \right) p(w-j)}{p(w)(p(w)-1)} = \frac{R}{p(w)(p(w)-1)}, \quad (4.4)$$

say. Concentrating on the numerator  $R$  of the last expression,

$$R = \sum_{j=0}^{w-2} \frac{w-j}{2} \left( \frac{w-j}{2} + 1 \right) p(j) = \left( \sum_{j=0}^{w_0-1} + \sum_{j=w_0}^{w-2} \right) \frac{w-j}{2} \left( \frac{w-j}{2} + 1 \right) p(j)$$

for any  $w_0$  intermediate between 0 and  $w-2$ . Choose  $w_0$  such that for  $j \geq w_0$ ,  $p(j) \leq e^{K\sqrt{j}}/\sqrt{j}$ . This is possible, since as  $w \rightarrow \infty$ ,  $p(w) \sim e^{K\sqrt{w}}/(4w\sqrt{3})$  (e.g. Apostol (1976, p. 316 ff)). Applying Lemma 4.5 yields

$$\begin{aligned} R &\leq \sum_{j=0}^{w_0-1} \frac{w-j}{2} \left( \frac{w-j}{2} + 1 \right) p(j) + \sum_{j=w_0}^{w-2} \frac{w-j}{2} \left( \frac{w-j}{2} + 1 \right) \frac{e^{K\sqrt{j}}}{\sqrt{j}} \\ &= e^{K\sqrt{w}} \left( \frac{8w}{K^3} + O(\sqrt{w}) \right) \end{aligned}$$

Returning to (4.4), one sees that, as  $w \rightarrow \infty$ ,

$$f_w = O \left( \frac{8we^{K\sqrt{w}}/K^3}{\{e^{K\sqrt{w}}/(4w\sqrt{3})\}^2} \right) = O \left( \frac{384w^3}{K^3} e^{-K\sqrt{w}} \right) \quad (4.5)$$

#### Corollary 4.7

The fraction  $f_w$  of non-zero elements in the upper half of the  $\Omega_w$  matrix satisfies  $f_w \rightarrow 0$  as  $w \rightarrow \infty$ . ■

### 4.3 From the operator matrix $\Omega$ to the coefficient matrix $C$

One can rewrite the relationships in (4.1) or (3.5) more succinctly as:

$$Y = C^m M = C^e E$$

in which  $C^m = (j_{\kappa, \lambda}^m)$  and  $C^e = (j_{\kappa, \lambda}^e)$  will be referred to as the (Jack polynomial) coefficient matrices with respect to the  $m_\lambda$  and  $e_\lambda$  functions respectively; and  $Y^T = (J_{(w)}, J_{(w-1,1)}, \dots)$ .

From (3.8) one has

$$j_{\rho}^m{}^T (\Omega^m - c_{\rho} I) = 0 . \quad (4.6)$$

The constraint in (4.6) arising from the  $\lambda$ th entry in the row vector is

$$\sum_{\substack{\kappa \\ \rho \geq \kappa > \lambda}} j_{\rho, \kappa}^m \omega_{\kappa, \lambda}^m = (c_{\rho} - c_{\lambda}) j_{\rho, \lambda}^m \quad (4.7)$$

The system of equations given in (4.6) and (4.7) provides a means of evaluating the  $j_{\rho, \lambda}^m$  coefficients by back substitution. The outworking is similar for the  $e_{\lambda}$  function, for which (4.7) becomes

$$\sum_{\substack{\kappa \\ \lambda > \kappa \geq \rho'}} j_{\rho, \kappa}^e \omega_{\kappa, \lambda}^e = (c_{\rho} - c_{\lambda'}) j_{\rho, \lambda}^e \quad (4.8)$$

One benefit from setting out the calculations as in (4.7) and (4.8) is that finding bounds on the  $j$  coefficients is facilitated. Should, for instance, one have an upper bound on the  $j_{\rho, \kappa}^m$  coefficients for  $\kappa > \lambda$ , and an upper bound on the elements  $\omega_{\kappa, \lambda}^m$  in the  $\lambda$ th column, one can get an upper bound for the coefficient  $j_{\rho, \lambda}^m$  from (4.7). Roberts (1998, §5.3) has attempted some analysis in this direction, since the column sums above the diagonal in  $\Omega^m$  are easy to calculate: see §4.5. Also the results in Chapter 5 concerning the  $j^m$  coefficients as  $w \rightarrow \infty$  are based on (4.7).

## 4.4 The composition of the $j$ coefficients

The matrices  $\Omega_{\rho, 1}$  arising in Chapter 3, by whose determinants one can calculate Jack polynomials, are almost diagonal: there is a single diagonal either immediately above or below the principal diagonal, coupled to a triangular matrix which becomes quite sparse as  $w \rightarrow \infty$ . The terms arising in those expansions assume a specific form, illustrated schematically below.

### 4.4.1 The composition of the $j_{\rho,\lambda}^m$ coefficients

For the expansion of  $J_\rho$  in  $m_\lambda$  functions, according to Theorem 3.5 on p. 85 we first delete the redundant initial rows and columns in  $\Omega_{\rho,1}^m$ . One then finds all possible paths from the top left corner to the bottom right corner, where the first part of the path is a d2 path, and the second part moves straight down the main diagonal. The process is illustrated in the diagram for the calculation of  $J_{(4)}$ . Labelling the elements of the matrix by  $\omega_{\kappa,\lambda}$  with the first row and column indexed by  $(4)$  and the last by  $(1^4)$ , the d2 path proceeds from  $\omega_{(4),(4)}$  to  $\omega_{(2,1,1),(2,1,1)}$ , where  $\omega_{(4),(4)}$  in fact contains the element  $m_{(4)}$ ; while the diagonal part of the path proceeds from  $\omega_{(2,1,1),(2,1,1)}$  to  $\omega_{(1^4),(1^4)}$ .

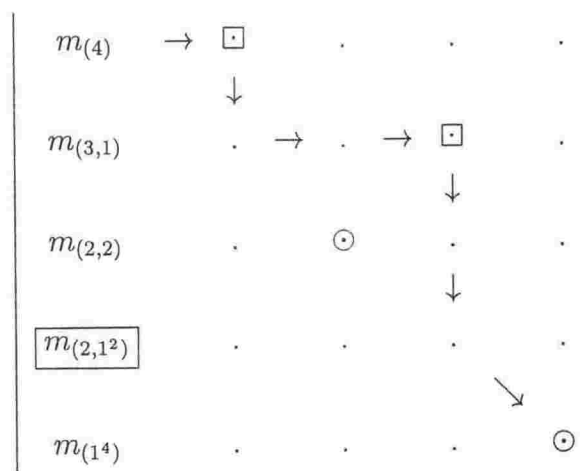


Figure 4.3:  $\Omega_{(4),1}^m$

The above-diagonal terms corresponding to this path are marked by the squares, and the only other off-diagonal representative is the function  $m_{(2,1^2)}$ . With the remaining terms denoted by circles, the corresponding term in the expansion is

$$\omega_{(4),(3,1)}^m \omega_{(3,1),(2,1,1)}^m m_{(2,1,1)} d_{(3,1)} d_{(1^4)}$$

where  $d_\kappa = c_\kappa - c_{(4)}$ .

The coefficient of  $m_{(2,1^2)}$  is found by summing over all possible such paths: a d2 path from the top left element to the diagonal element  $\omega_{(2,1,1),(2,1,1)}$ , followed by a path down the remainder of the diagonal to the bottom right element.

For weights exceeding 5 there will be in general further redundant rows and columns, indexed by  $\kappa$  such that  $\rho \stackrel{R}{>} \kappa$  but  $\rho \not\asymp \kappa$ . The d2 paths simply bypass those elements, such that the diagonal elements  $\omega_{\kappa,\kappa} = c_\kappa - c_\rho$  enter into the determinant, but then cancel, as shown in Theorem 3.5. The problem with this is that the diagonal element will vanish if  $c_\kappa = c_\rho$ , which happens occasionally: see Example 3.4 on p. 83. So one should jettison redundant rows and columns before starting calculations, and not just for purposes of expediting the calculation.

Moving the first column of  $m_\lambda$  functions to the right of the matrix in Figure 4.3 yields a matrix of the same shape as that used in Lapointe et al. (2000). Our approach through d2 paths of  $\Omega$  is equivalent to their calculation of Jack polynomials as determinants of “Hessenberg” matrices.

4.4.2 The composition of the  $j^e_{\rho,\lambda}$  coefficients

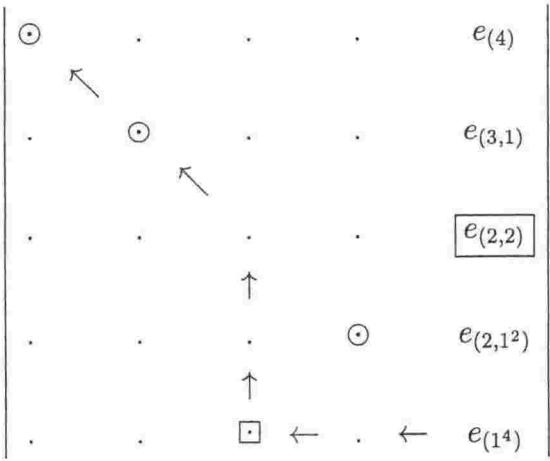


Figure 4.4:  $\Omega_{(4),1}^e$

The story is analogous for the  $e_\lambda$  functions. After deleting redundant final rows and columns in  $\Omega_{(4),1}^e$ , the d2 path proceeds from the bottom right corner upwards below the diagonal to the preassigned diagonal element, and along the diagonal element thereafter. For the example shown in Figure 4.4, the term in the expansion of  $J_{(4)}$  is

$$\omega_{(1^4),(2,2)}^e e_{(2,2)} d'_{(4)} d'_{(3,1)} d'_{(2,1,1)}$$

where  $d'_\kappa = c_{\kappa'} - c_{(4)}$ . Again, the coefficient of  $e_{(2,2)}$  is given as the sum of the terms arising from such combinations of d2 and diagonal paths.

Letting the d2 path have length  $s$ , the contribution to the coefficient of  $e_\kappa$  contains a factor of  $(-\alpha)^s$ , since the diagonal terms are positive. For the expansion in  $m_\lambda$  functions, on the other hand, the diagonal terms are negative: given the sign changes implicit in the definition of the determinant, all contributions to the Jack polynomial are of the one sign for these latter expansions.

### 4.4.3 Dominant terms in the determinantal expansions

The more important terms in the determinantal expansions can be expected to arise as follows.

Consider the expansion in terms of the  $m_\lambda$  functions. Assuming the  $m_\lambda$  functions are relatively insensitive to change in the index partition  $\lambda$  (and we shall see in Chs. 6 and 7 that this is not always the case), the dominant terms can be expected to be those with the d2 chains (starting at the top left corner) terminating about the middle of the matrix. This enables all the lower terms (towards the bottom right corner) on the diagonal to enter into the expression, and those diagonal terms are large. Within the d2 chain itself, the chain should keep close to the diagonal at the upper left corner of the matrix, in order to minimise the number of upper diagonal terms used, since those diagonal terms have small magnitude. Lower down the d2 chain, one should move away more from the diagonal, since those diagonal terms are somewhat larger.

Likewise, restricting ourselves to the coefficient of  $m_\lambda$  in  $J_\rho$ , the dominant terms are expected to be those with d2 chains (over  $[\rho, \lambda]$ ) which hug the diagonal at the top end, and move away from it at the bottom end (towards the  $(\lambda, \lambda)$ th element).

The coefficients  $j_{\rho, \lambda}^m$  contain relatively few terms for  $\rho$  and  $\lambda$  close in the RLO; but they contain large numbers of terms for  $\rho$  high and  $\lambda$  low in the RLO, because the interval  $[\rho, \lambda]$  then supports many d2 chains. Otherwise expressed, the  $j_{\rho, \lambda}^m$  coefficient

in the latter case arises towards the end of the recursive backward substitution in (4.7), also rendering the coefficient subject to substantial rounding errors.

Analogous comments apply to the expansion of Jack polynomials in  $e_\lambda$  functions, save that the diagonal elements of greatest magnitude are now at the top of the operator matrix rather than the bottom. The situation concerning dominant terms is less transparent, since the sign changes in the elements make it harder to gauge the importance of individual terms in the overall picture.

The diagonal terms can be expected to be much larger in magnitude than the off-diagonal terms, at least for the bottom end of the diagonal for the  $m_\lambda$  functions and the top end for the  $e_\lambda$  functions. This can be seen informally from the results in §4.5, in which most of the results suggest that column sums  $\eta_\lambda$  of  $\Omega^m$  are of comparable magnitude to the diagonal elements  $c_\lambda - c_\rho$ : see in particular Corollary 4.16.

Our discussion has centred on the  $m_\lambda$  functions, since in (4.7) it is the columns of  $\Omega^m$  which are involved with the back substitution. Similarly from (4.8) it is the columns of  $\Omega^e$  which are involved, i.e. the rows of  $\Omega^m$ . The row sums of  $\Omega^m$  are less tractable than the column sums, and the off-diagonal elements are greater towards the bottom right corner of  $\Omega^m$  than at the top left, because the value of  $K$  in Theorem 3.18 is large (the multiplicities of the elements are large for  $\lambda$  low in the RLO). But overall one again expects column sums and diagonal elements for  $\Omega^e$  to be broadly comparable in magnitude, so that individual off-diagonal terms will generally be less than diagonal elements.

#### 4.4.4 Path descriptors in terms of core chains

Following on from Figure 4.1, Figures 4.5 and 4.6 illustrate two particular d2 paths from (5) to (1<sup>5</sup>).

Descriptors such as *abbhj* and *aefj* for those examples need not identify uniquely the

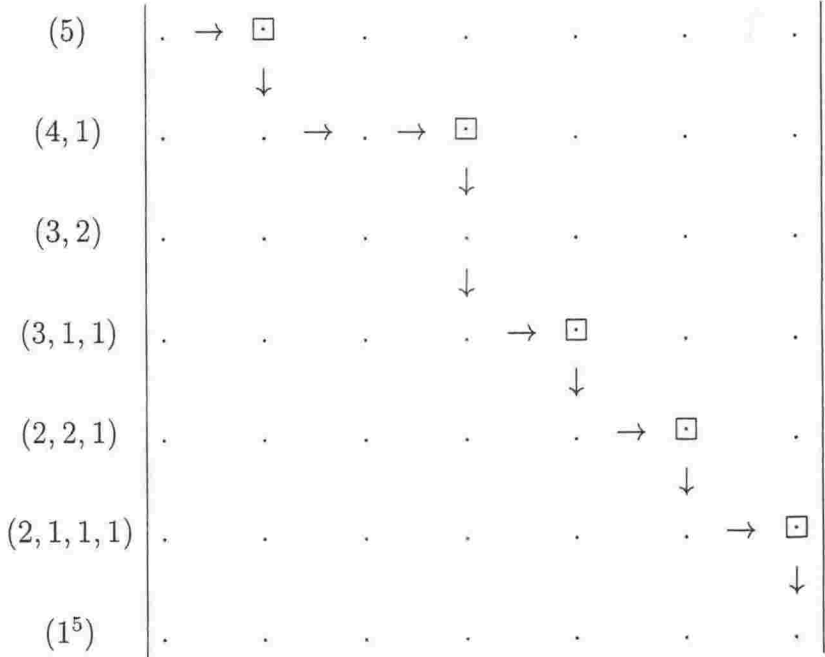


Figure 4.5: d2 path *abbhj* from (5) to (1<sup>5</sup>)

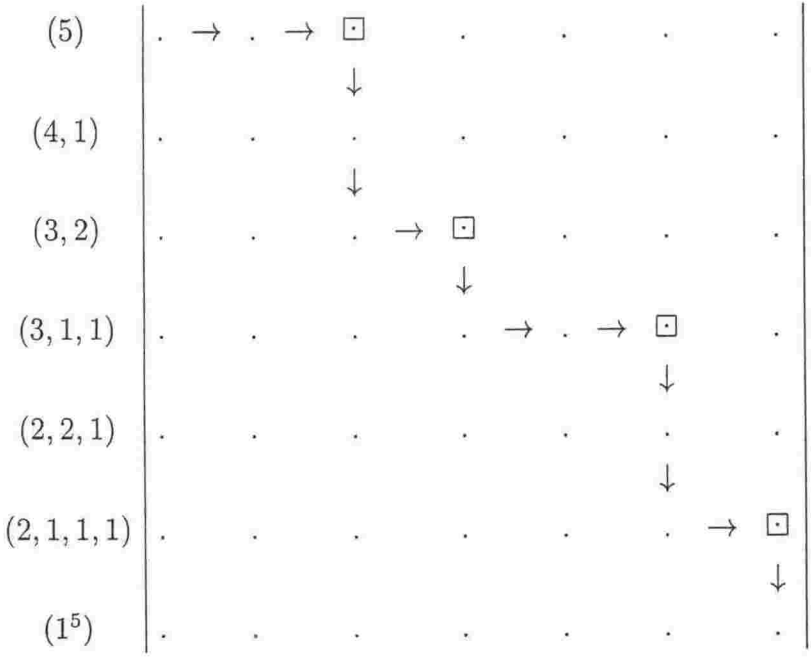


Figure 4.6: d2 path *ae fj* from (5) to (1<sup>5</sup>)

underlying d2 path. For partition weight 6, for instance, consider the core chain

$$(6) \rightarrow (5, 1) \rightarrow (4, 2) \rightarrow (3, 3)$$

which is labelled as, say,  $a$ . Then chains from  $\rho = (6)$  to  $\lambda < (3, 3)$  may contain the subchains  $(6) \rightarrow (5, 1) \rightarrow (3, 3) \rightarrow \dots$  and  $(6) \rightarrow (4, 2) \rightarrow (3, 3) \rightarrow \dots$ , both of which are indexed by  $aa \dots$

Apart from this ambiguity from multiple passes through the one chain, the labelling of d2 paths by the core chains through which successive off-diagonal elements pass identifies the path uniquely.

#### 4.4.5 Incidence matrices

A simple means of investigating the structure of paths passing from partitions  $\kappa$  to  $\lambda$  in  $\Omega^m$  is to define an incidence matrix  $Q$  say, with unity inserted in the place of every non-zero element above the diagonal in  $\Omega_w^m$ . Then the number of paths from  $\kappa$  to  $\lambda$  of length  $s$  is the  $(\kappa, \lambda)$ th entry in  $Q^s$ . See Stanley (1986, p. 113) for further details, although note that the  $\overset{d2}{>}$  ordering is not a partial ordering, since it is not transitive.

For counting the number of d2-1 paths from  $\kappa$  to  $\lambda$  along the lines of Lemma 3.2, one could similarly define an incidence matrix with unit  $(\kappa, \lambda)$ th entry when  $\kappa \overset{d2}{>}_1 \lambda$ . Note that  $\overset{d2}{>}_1$  is not a partial ordering either, although it provides the covering relation for the majorisation partial ordering  $>$  (e.g. Brylawski (1973)). The maximum length of a path from  $\kappa$  to  $\lambda$  is given in Greene & Kleitman (1986).

### 4.5 Relative magnitudes of diagonal and off-diagonal terms in $\Omega^m$

**Lemma 4.8**

The sum of the elements in  $\Omega_w^m$  (strictly) above the  $(\lambda, \lambda)$ th diagonal position is given by

$$\eta_\lambda = \sum_{\substack{\kappa \\ \kappa > \lambda}} \omega_{\kappa, \lambda}^m = \sum_{\substack{j, k=1 \\ j < k}}^{\infty} l_j l_k + \sum_{j=1}^{\infty} (j-1) l_j \quad (4.9)$$

$$= w(w-1) - \sum_{j=1}^{\infty} l_j^2 + \sum_{j=1}^{\infty} j l_j - \sum_{\substack{j, k=1 \\ j < k}}^{\infty} l_j l_k \quad (4.10)$$

$$= \frac{w}{2}(w-2) - \frac{1}{2} \sum_{j=1}^{\infty} l_j^2 + \sum_{j=1}^{\infty} j l_j \quad (4.11)$$

$$= \frac{w}{2}(w-1) - \frac{1}{2} \sum_{j=1}^{\infty} j^2 m_j + \sum_{\substack{j, k=1 \\ j < k}}^{\infty} j m_j m_k + \frac{1}{2} \sum_{j=1}^{\infty} j m_j^2 \quad (4.12)$$

**Proof 4.8**

Consider the first two entries  $l_1, l_2$  of  $\lambda$ . This pair gives rise to elements  $\omega_{\kappa, \lambda}$ , with row indices the partitions  $\kappa = (l_1 + c, l_2 - c, \dots)$ , for  $c = 1, \dots, l_2$ , the difference in each row pair being  $l_1 - l_2 + 2c$ . Summing these latter terms gives  $l_2(l_1 - l_2) + 2 \sum_{c=1}^{l_2} c = l_2(l_1 + 1)$ . Repeating this for all pairs  $(l_1, l_i)$  for  $i > 1$  yields the sum  $\sum_{i>1} l_i(l_1 + 1) = (w - l_1)(l_1 + 1)$ ; and further repeating the process for all pairs yields the sum for the column above the diagonal of

$$(w - l_1)(l_1 + 1) + (w - l_1 - l_2)(l_2 + 1) + (w - l_1 - l_2 - l_3)(l_3 + 1) + \dots$$

$$\begin{aligned} &= \sum_{j=1}^{\infty} w(l_j + 1) - \sum_{j=1}^{\infty} l_j(l_j + 1) - \sum_{\substack{j, k=1 \\ j < k}}^{\infty} l_j(l_k + 1) \\ &= w(w + r) - \sum_{j=1}^{\infty} l_j^2 - w - \sum_{\substack{j, k=1 \\ j < k}}^{\infty} l_j l_k - \sum_{j=1}^{\infty} l_j(r - j) \end{aligned}$$

where  $r = \ell(\lambda)$ . This expression easily reduces to (4.10).

The equivalence of (4.9), (4.10), (4.11) and (4.12) is straightforward to show. The infinite upper limit to the summations is of course notional: the upper limit to  $j$  is  $\ell(\lambda)$  when  $l_j$  is involved, and  $h(\lambda)$  when  $m_j$  is involved. ■

**Theorem 4.9**

$$\eta_\lambda + \eta_{\lambda'} = w(w - 1). \quad (4.13)$$

**Proof 4.9**

The multiplicities for the conjugate partition  $\lambda'$  are given by  $m'_j = l_j - l_{j+1}$ . Substitution of  $m'_j$  for  $m_j$  in (4.12), noting that

$$\sum_{\substack{j,k=1 \\ j < k}}^{\infty} j m'_j m'_k = \sum_{j=1}^{\infty} j m'_j l_{j+1}, \quad (4.14)$$

and some simplification yield the result.

**Corollary 4.9**

$$\eta_{\lambda'} = \frac{w^2}{2} + \frac{1}{2} \sum_{j=1}^{\infty} l_j^2 - \sum_{j=1}^{\infty} j l_j \quad (4.15)$$

$$\eta_{\lambda'} - \eta_\lambda = w + \sum_{j=1}^{\infty} l_j^2 - 2 \sum_{j=1}^{\infty} j l_j \quad (4.16)$$

These are immediate from (4.11) and (4.13). Alternatively (4.15) is easily proved directly from (4.12) and (4.14). ■

**Lemma 4.10**

If  $\lambda = (w - t) \cup \kappa$ , with  $w(\kappa) = t$  and  $h(\kappa) \leq w - t$ , then

$$\begin{aligned} \eta_\lambda - \eta_\kappa &= t(w - t + 1) \\ c_\lambda - c_\kappa &= \left(n - \frac{\alpha}{2}\right)(w - t) - w + \frac{\alpha}{2}(w - t)^2 \end{aligned} \quad (4.17)$$

The equation (4.17) will be applied in §5.1. ■

**Lemma 4.11**

For any partitions  $\kappa$  and  $\lambda$  such that  $w(\kappa) = w(\lambda)$ ,

$$\begin{aligned} c_\kappa^{(\alpha)} - c_\lambda^{(\alpha)} &= \sum j(l_j - k_j) + \frac{\alpha}{2} \sum (k_j^2 - l_j^2) && \text{from (3.4)} \\ c_{\kappa'}^{(\alpha)} - c_{\lambda'}^{(\alpha)} &= -\alpha \sum j(l_j - k_j) - \frac{1}{2} \sum (k_j^2 - l_j^2) && \text{from (3.21)} \end{aligned}$$

$$\begin{aligned}
 c_{\kappa}^{(\alpha)} - c_{\lambda}^{(\alpha)} &= -\alpha \left( c_{\kappa'}^{(1/\alpha)} - c_{\lambda'}^{(1/\alpha)} \right) \\
 \eta_{\kappa} - \eta_{\lambda} &= -\sum j(l_j - k_j) - \frac{1}{2} \sum (k_j^2 - l_j^2) = -(\eta_{\kappa'} - \eta_{\lambda'}) \quad \text{from (4.11), (4.13)} \\
 \eta_{\kappa} - \eta_{\lambda} &= c_{\kappa'}^{(1)} - c_{\lambda'}^{(1)} = -\left( c_{\kappa}^{(1)} - c_{\lambda}^{(1)} \right)
 \end{aligned}$$

**Proof 4.11**

The provenance of the first, second and fourth equations is as indicated. The remaining equations are consequences of the other three. ■

**Lemma 4.12**

Let  $\kappa > \lambda$ . Then

$$c_{\kappa} > c_{\lambda} \quad \text{and} \quad \eta_{\kappa} < \eta_{\lambda}$$

**Proof 4.12**

Corollary 3.3 on p. 83 gives the first result for any  $\alpha > 0$ , while the second follows from the final equation of Lemma 4.11. ■

**Lemma 4.13**

$$\frac{c_{\lambda} - c_{\lambda'}}{\alpha + 1} = \frac{w}{2} (w - 1) - \eta_{\lambda} = -\frac{w}{2} (w - 1) + \eta_{\lambda'}$$

**Proof 4.13**

From (3.4) and Lemma 3.10 one has

$$c_{\lambda} - c_{\lambda'} = (\alpha + 1) \left[ \frac{w}{2} - \sum j l_j + \frac{1}{2} \sum l_j^2 \right] \quad (4.18)$$

The result follows from equation (4.11) and Theorem 4.9. ■

**Theorem 4.14**

If  $\eta_{\lambda} \neq \eta_{\lambda'}$ , then

$$\frac{c_{\lambda} - c_{\lambda'}}{\eta_{\lambda'} - \eta_{\lambda}} = \frac{\alpha + 1}{2}$$

**Proof 4.14**

The proof is immediate from (4.16) and (4.18). The result is also a corollary of Theorem 4.15.

**Corollary 4.14**

$$\frac{c_{(w)} - c_{(1^w)}}{\eta_{(1^w)}} = \frac{\alpha + 1}{2}$$

■

**Theorem 4.15**

If  $\eta_\kappa \neq \eta_\lambda$ , then

$$\frac{c_\kappa - c_\lambda}{\eta_\lambda - \eta_\kappa} + \frac{c_{\lambda'} - c_{\kappa'}}{\eta_{\kappa'} - \eta_{\lambda'}} = \alpha + 1$$

**Proof 4.15**

The proof is immediate from Lemma 4.11.

■

**Theorem 4.16**

Let  $\kappa > \lambda$ . Then

$$\begin{aligned} 1 &< \frac{c_\kappa - c_\lambda}{\eta_\lambda - \eta_\kappa} < \alpha && \text{when } \alpha > 1; \\ \alpha &< \frac{c_\kappa - c_\lambda}{\eta_\lambda - \eta_\kappa} < 1 && \text{when } \alpha < 1; \quad \text{and} \\ \frac{c_\kappa - c_\lambda}{\eta_\lambda - \eta_\kappa} &= 1 && \text{when } \alpha = 1. \end{aligned}$$

**Proof 4.16**

Recalling that  $\kappa > \lambda \Leftrightarrow \lambda' > \kappa'$ , the theorem follows immediately from Lemmas 4.11 and 4.12.

**Corollary 4.16**

With  $\rho > \lambda$ , set

$$\gamma = \frac{c_{(w)} - c_\lambda}{c_\rho - c_\lambda}$$

Then

$$\frac{\gamma}{\alpha} < \frac{\eta_\lambda}{c_\rho - c_\lambda} < \gamma \quad \text{when } \alpha > 1; \tag{4.19}$$

$$\gamma < \frac{\eta_\lambda}{c_\rho - c_\lambda} < \frac{\gamma}{\alpha} \quad \text{when } \alpha < 1; \quad \text{and}$$

$$\frac{\eta_\lambda}{c_\rho - c_\lambda} = \gamma \quad \text{when } \alpha = 1. \quad \blacksquare$$

The significance of Corollary 4.16 is as follows. From (4.7) on p. 105, and letting  $L$  be an upper bound on  $j_{\rho,\kappa}^m$  for  $\kappa > \lambda$ , one has that

$$j_{\rho,\lambda}^m \leq L \frac{\eta_\lambda}{c_\rho - c_\lambda} \quad (4.20)$$

This type of result can form the basis for obtaining bounds on  $j_{\rho,\lambda}^m$  coefficients as  $\lambda$  moves down the RLO. The basic result (4.20) can be refined, in that the upper bound on the  $j_{\rho,\kappa}^m$  coefficients need only apply to those partitions  $\kappa$  such that  $\omega_{\kappa,\lambda}^m > 0$ . Work along these lines is in Roberts (1998, §5.3).

One caveat needs to be borne in mind as regards the Corollary 4.16. One does not in fact want  $\eta_\lambda$ , because that column sum includes elements in the redundant first rows of the operator matrix (the rows indexed by  $\kappa \stackrel{R}{>} \rho$  when calculating  $J_\rho$ ). The constraint in (4.20) is therefore weaker than one would wish.

## 4.6 Slicing and reflection results on $j_{\rho,\lambda}$ coefficients

Evaluating  $j$  coefficients through d2 paths in the fashion of §4.4 provides elementary proofs of “slicing” and “reflection” results, extending results of Takemura (1984, p. 59) to general  $\alpha > 0$ . The terminology is non-standard. Our development broadly follows that of Roberts (1998, §4.3), who refers to reflecting results as “dual” results; but this seems poor terminology, partly because the word dual assumes so many meanings in mathematics, and partly because its usage there is inconsistent with the use of the word dual in Stanley (1989, §3), and in Macdonald (1995). Diagonal elements of the coefficient matrix  $C$  in §4.3 are normalised to unity.

For expansions in the  $m_\lambda$  functions, the “slicing” theorems essentially allow us to slice off either the common first row or common first column from the standard Ferrers diagrams for the root and target partitions  $\rho$  and  $\lambda$ , and leave the  $j_{\rho,\lambda}^m$

coefficient unchanged. For the  $e_\lambda$  functions, the procedure is to slice off the first row for the root partition  $\rho$  and the first column of the target partition  $\lambda$ ; or to slice off the first column of  $\rho$  and the first row of  $\lambda$ , in both cases leaving the  $j_{\rho,\lambda}^e$  coefficient unchanged.

The reflection result has to do with reflecting the partition  $\lambda$ , or rather its Ferrers diagram, in a rectangular  $p \times q$  grid of cells. Denoting the reflecting partition of  $\lambda$  by  $\bar{\lambda}$ ,

$$(q^p) = \lambda + \bar{\lambda}$$

where  $\bar{\lambda}$  is in non-decreasing order, and contains initial zeroes if  $h(\lambda) = q$ ; clearly  $w(\lambda) + w(\bar{\lambda}) = pq$ . Let  $w(\lambda) = w$  as previously, and  $w(\bar{\lambda}) = \bar{w}$ .

While the slicing and reflection results certainly assist in the calculation of  $J_\rho$ , they are not necessarily of much help in calculating the “difficult”  $j$  coefficients. The hard  $j_{\rho,\lambda}^m$  coefficients to calculate are those with  $\rho$  high and  $\lambda$  low in the RLO, since  $[\rho, \lambda]$  then supports many d2 paths; and the Ferrers diagrams do not have a common initial row or column.

The situation is not fundamentally different for  $e_\lambda$  functions. Coefficients  $j_{\rho,\lambda}^e$  are hard to calculate when  $\rho$  and  $\lambda$  are high in the RLO, which is when the slicing results do not apply. As for the reflection results, the  $j_{\rho,\lambda}$  coefficients which are hard to calculate tend to have reflecting partitions  $\bar{\rho}$  and  $\bar{\lambda}$  of far greater weight than  $\rho$  and  $\lambda$ .

#### 4.6.1 The basic supporting result for the slicing and reflection theorems

The following lemma underpins all of the slicing and reflection results. Throughout this section we normalise so that  $j_{\rho,\rho'}^e = j_{\rho,\rho}^m = 1$ . Also recall the definitions of  $\Omega_{\rho,0}^e$  and  $\Omega_{\rho,0}^m$  from §§3.2.2 and 3.2.3 on pp. 83 and 84 respectively.

**Lemma 4.17**

$$\begin{aligned} (\Omega_{\rho_1,0}^m)_{[\rho_1,\lambda_1]} &= (\Omega_{\rho_2,0}^m)_{[\rho_2,\lambda_2]} \Rightarrow j_{\rho_1,\lambda_1}^m = j_{\rho_2,\lambda_2}^m \\ (\Omega_{\rho_1,0}^e)_{[\lambda_1,\rho_1']} &= (\Omega_{\rho_2,0}^e)_{[\lambda_2,\rho_2']} \Rightarrow j_{\rho_1,\lambda_1}^e = j_{\rho_2,\lambda_2}^e \end{aligned}$$

**Proof 4.17**

The coefficients  $j_{\rho,\lambda}^m$  are formed from d2 paths over  $[\rho, \lambda]$ , according to the development in §4.4.

The further terms arising in  $j_{\rho,\lambda}^m$  are the diagonal terms  $(\Omega_{\rho,0}^m)_{\kappa,\kappa}$  for which  $\lambda > \kappa$ . None of those terms is zero, from Corollary 3.3, since all indices in  $\Omega_{\rho,0}^m$  are majorised by  $\rho$ . And all those terms cancel, as they appear in both numerator and denominator in Theorem 3.5.

The proof of the second result is analogous. ■

## 4.6.2 Slicing theorems

We divide this section into technical lemmas underpinning the slicing theorems and the theorems themselves. The supporting results are conveniently subdivided into results to do with off-diagonal and diagonal elements, and eigenvalues.

### 4.6.2.1 Supporting technical results

#### 4.6.2.1.1 Off-diagonal elements

**Lemma 4.18**

Let  $\kappa > \lambda$ , with  $p \geq \ell(\lambda)$ . Then

$$(\Omega_w^m)_{\kappa,\lambda} = (\Omega_{w+p}^m)_{\kappa+(1^p),\lambda+(1^p)} \quad (4.21)$$

$$(\Omega_w^e)_{\lambda,\kappa} = (\Omega_{w+p}^e)_{\lambda+(1^p),\kappa+(1^p)} \quad (4.22)$$

Still with  $\kappa > \lambda$ , let  $q \geq h(\kappa)$ . Then

$$(\Omega_w^m)_{\kappa, \lambda} = (\Omega_{w+q}^m)_{\kappa \cup (q), \lambda \cup (q)} \quad (4.23)$$

$$(\Omega_w^e)_{\lambda, \kappa} = (\Omega_{w+q}^e)_{\lambda \cup (q), \kappa \cup (q)} \quad (4.24)$$

#### Proof 4.18

$$\kappa \stackrel{d2}{>} \lambda \quad \text{iff} \quad \kappa + (1^p) \stackrel{d2}{>} \lambda + (1^p)$$

and

$$\kappa \stackrel{d2}{>} \lambda \quad \text{iff} \quad \kappa \cup (q) \stackrel{d2}{>} \lambda \cup (q)$$

so that (4.21) and (4.23) are a consequence of Theorem 3.18 on p. 94. The other equations (4.22) and (4.24) follow from Theorem 3.6 on p. 87. ■

#### 4.6.2.1.2 Diagonal elements

##### Lemma 4.19

If  $p \geq \ell(\lambda)$ , then

$$(\Omega_{\rho,0}^m)_{\lambda, \lambda} = (\Omega_{\rho+(1^p),0}^m)_{\lambda+(1^p), \lambda+(1^p)} \quad (4.25)$$

$$(\Omega_{\rho,0}^e)_{\lambda, \lambda} = (\Omega_{\rho \cup (p),0}^e)_{\lambda+(1^p), \lambda+(1^p)} \quad (4.26)$$

If  $q \geq h(\lambda)$ , then

$$(\Omega_{\rho,0}^m)_{\lambda, \lambda} = (\Omega_{\rho \cup (q),0}^m)_{\lambda \cup (q), \lambda \cup (q)} \quad (4.27)$$

$$(\Omega_{\rho,0}^e)_{\lambda, \lambda} = (\Omega_{\rho+(1^q),0}^e)_{\lambda \cup (q), \lambda \cup (q)} \quad (4.28)$$

#### Proof 4.19

$$\rho \geq \lambda \quad \text{iff} \quad \rho + (1^p) \geq \lambda + (1^p),$$

while the  $(\lambda, \lambda)$ th element of  $\Omega_\rho^m$  is  $c_\lambda - c_\rho$ , so that the proof of (4.25) is immediate from Lemma 4.20. Given Lemma 4.21, the proof of (4.27) is similarly evident from

$$\rho \geq \lambda \quad \text{iff} \quad \rho \cup (q) \geq \lambda \cup (q).$$

For (4.26) one has that

$$\lambda \geq \rho' \quad \text{iff} \quad \lambda + (1^p) \geq \rho' + (1^p)$$

while the diagonal elements are  $c_{\lambda'} - c_\rho$  and  $c_{\lambda' \cup (p)} - c_{\rho \cup (p)}$ , whence the result follows from Lemma 4.20. Similarly

$$\lambda \geq \rho' \quad \text{iff} \quad \lambda \cup (q) \geq \rho' \cup (q),$$

with diagonal elements  $c_{\lambda'} - c_\rho$  and  $c_{\lambda' + (1^q)} - c_{\rho + (1^q)}$ , and (4.28) is immediate from Lemma 4.21. ■

#### 4.6.2.1.3 Eigenvalues

The point of the next two lemmas is that the difference in eigenvalues depends on  $\lambda$  only through its weight  $w$ . Recall the definition of  $c_\lambda$  in (3.4) on p. 81.

##### Lemma 4.20

Let  $p \geq \ell(\lambda)$ . Then

$$c_{\lambda + (1^p)} - c_\lambda = np - \frac{p(p+1)}{2} + \alpha w$$
■

##### Lemma 4.21

Let  $q \geq h(\lambda)$ . Then

$$c_{\lambda \cup (q)} - c_\lambda = nq + \alpha \frac{q(q-1)}{2} - (q + w)$$
■

#### 4.6.2.2 The slicing theorems

There are four basic theorems to be stated here. For completeness we mention a known fifth result, Theorem 4.26, which is a direct corollary of both Theorems 4.22 and 4.25.

##### Theorem 4.22

Let  $\rho > \lambda$ , and let  $p \geq \ell(\lambda)$ . Then

$$j_{\rho, \lambda}^m = j_{\rho + (1^p), \lambda + (1^p)}^m \tag{4.29}$$

**Proof 4.22**

Noting that  $\rho > \kappa > \lambda$  iff  $\rho + (1^p) > \kappa + (1^p) > \lambda + (1^p)$ , we consider  $(\Omega_{\rho,0}^m)_{[\rho,\lambda]}$  and  $(\Omega_{\rho+(1^p),0}^m)_{[\rho+(1^p),\lambda+(1^p)]}$ .

These matrices are identical off the diagonal, from (4.21), while all the diagonal terms coincide from (4.25). The result (4.29) then follows from Lemma 4.17. ■

**Theorem 4.23**

Let  $\lambda > \rho'$ ; and let  $p \geq \ell(\lambda)$  and  $p \geq h(\rho)$ . Then

$$j_{\rho,\lambda}^e = j_{\rho \cup (p), \lambda + (1^p)}^e \quad (4.30)$$

**Proof 4.23**

Seeing that

$$\lambda > \kappa > \rho' \quad \text{iff} \quad \lambda + (1^p) > \kappa + (1^p) > \rho' + (1^p),$$

one contrasts  $(\Omega_{\rho,0}^e)_{[\lambda,\rho']}$  and  $(\Omega_{\rho \cup (p),0}^e)_{[\lambda+(1^p),\rho'+(1^p)]}$ . Equation (4.22) again yields that the off-diagonal terms are equal.

The diagonal terms of the matrices are  $c_{\kappa'} - c_\rho$  and  $c_{\kappa' \cup (p)} - c_{\rho \cup (p)}$ , which are equal by Lemma 4.21. The result follows from Lemma 4.17. ■

**Theorem 4.24**

Let  $\rho > \lambda$ , and let  $q \geq h(\lambda)$ . Then

$$j_{\rho,\lambda}^m = j_{\rho \cup (q), \lambda \cup (q)}^m \quad (4.31)$$

**Proof 4.24**

Noting that  $\rho > \kappa > \lambda$  iff  $\rho \cup (q) > \kappa \cup (q) > \lambda \cup (q)$ , we consider  $(\Omega_{\rho,0}^m)_{[\rho,\lambda]}$  and  $(\Omega_{\rho \cup (q),0}^m)_{[\rho \cup (q), \lambda \cup (q)]}$ .

These matrices are identical off the diagonal, from (4.23). All the diagonal terms assume the form  $c_\kappa - c_\rho$  and  $c_{\kappa \cup (q)} - c_{\rho \cup (q)}$  respectively, and are identical from Lemma 4.21. The result follows from Lemma 4.17. ■

**Theorem 4.25**

Let  $\lambda > \rho'$ ; and let  $q \geq h(\lambda)$  and  $q \geq \ell(\rho)$ . Then

$$j_{\rho, \lambda}^e = j_{\rho+(1^q), \lambda \cup (q)}^e \quad (4.32)$$

**Proof 4.25**

$$\lambda > \kappa > \rho' \quad \text{iff} \quad \lambda \cup (q) > \kappa \cup (q) > \rho' \cup (q)$$

One contrasts  $(\Omega_{\rho, 0}^e)_{[\lambda, \rho']}$  and  $(\Omega_{\rho+(1^q), 0}^e)_{[\lambda \cup (q), \rho' \cup (q)]}$ . Applying (4.24) yields that the off-diagonal terms of these matrices are equal.

The diagonal terms are  $c_{\kappa'} - c_{\rho}$  and  $c_{\kappa'+(1^q)} - c_{\rho+(1^q)}$ , which are equal by Lemma 4.20. The result follows from Lemma 4.17. ■

The following theorem is well known: see i.a. Stanley (1989, p. 90), James (1964, p. 493).

**Theorem 4.26**

When there are exactly  $p$  variates, and when  $\ell(\rho) = p$ ,

$$J_{\rho+(1^p)} = x_1 x_2 \dots x_p J_{\rho} = e_p J_{\rho}$$

**Proof 4.26**

$$J_{\rho+(1^p)} = \sum_{\rho \geq \lambda} j_{\rho+(1^p), \lambda+(1^p)}^m m_{\lambda+(1^p)} = x_1 x_2 \dots x_p \sum_{\rho \geq \lambda} j_{\rho, \lambda}^m m_{\lambda}$$

where we have used theorem 4.22 on the one hand, and the fact that

$$m_{\lambda+(1^p)} = x_1 x_2 \dots x_p m_{\lambda}$$

on the other. ■

Theorem 4.26 could also be proved by expanding in  $e_{\lambda}$  functions and applying Theorem 4.25. There seems no obvious analogous expression for  $J_{\rho \cup (q)}$  in terms of  $J_{\rho}$ .

### 4.6.3 Reflection results

This section is divided into technical lemmas underpinning the reflection theorems, the theorems themselves, and an example. The preliminary supporting results are again subdivided into results to do with off-diagonal and diagonal elements, and eigenvalues.

#### 4.6.3.1 Supporting technical results

##### 4.6.3.1.1 Off-diagonal elements

#### Lemma 4.27

Define the reflecting partitions  $\bar{\kappa}$  and  $\bar{\lambda}$ , and the reflecting weight  $\bar{w}$ , as in §4.6.

Let  $\kappa > \lambda$ ,  $p \geq \ell(\lambda)$  and  $q \geq h(\kappa)$ . Then

$$(\Omega_w^m)_{\kappa, \lambda} = (\Omega_{\bar{w}}^m)_{\bar{\kappa}, \bar{\lambda}}$$

#### Proof 4.27

$$\kappa \stackrel{d2}{>} \lambda \quad \text{iff} \quad \bar{\kappa} \stackrel{d2}{>} \bar{\lambda}$$

Visualising the  $p \times q$  grid in which the Ferrers diagrams are being reflected shows that in the calculations specified in Theorem 3.18, the difference in partition elements is mirrored whether looking from the left of the diagram or the right, so that the  $\Delta$  in that formula is the same for both cases; likewise the multiplicities in  $K$  in that formula are identical.

#### Corollary 4.27

From Theorem 3.6 one has

$$(\Omega_w^e)_{\lambda, \kappa} = (\Omega_{\bar{w}}^e)_{\bar{\lambda}, \bar{\kappa}}$$

■

#### 4.6.3.1.2 Diagonal elements

##### Lemma 4.28

$$(\Omega_{\rho,0}^m)_{\lambda,\lambda} = (\Omega_{\bar{\rho},0}^m)_{\bar{\lambda},\bar{\lambda}}$$

##### Proof 4.28

Diagonal elements are  $c_\lambda - c_\rho$  and  $c_{\bar{\lambda}} - c_{\bar{\rho}}$ . These are equal, from Theorem 4.32. ■

The following result is needed for Lemma 4.30.

##### Lemma 4.29

Let  $p = q$ . Then  $\bar{\lambda}' = \bar{\lambda}'$ . ■

##### Lemma 4.30

When  $p = q$ ,

$$(\Omega_{\rho,0}^e)_{\lambda,\lambda} = (\Omega_{\bar{\rho},0}^e)_{\bar{\lambda},\bar{\lambda}}$$

##### Proof 4.30

Diagonal elements are  $c_{\lambda'} - c_\rho$  and  $c_{\bar{\lambda}'} - c_{\bar{\rho}}$ . But  $c_{\bar{\lambda}'} = c_{\bar{\lambda}'}$  from Lemma 4.29, and one can appeal to Theorem 4.32 to see that  $c_{\lambda'} - c_\rho = c_{\bar{\lambda}'} - c_{\bar{\rho}}$ . ■

#### 4.6.3.1.3 Eigenvalues

##### Lemma 4.31

Suppose that  $\kappa \stackrel{d2}{>} \lambda$ . Then  $\bar{\kappa} \stackrel{d2}{>} \bar{\lambda}$ ; and

$$c_\kappa - c_\lambda = c_{\bar{\kappa}} - c_{\bar{\lambda}}$$

##### Proof 4.31

For any partitions  $\kappa$  and  $\lambda$  of the same weight, recall from (3.6) that

$$c_\kappa - c_\lambda = \sum j(l_j - k_j) + \frac{\alpha}{2} \sum (k_j - l_j)(k_j + l_j)$$

Now suppose that  $\kappa \stackrel{d2}{>} \lambda$ ; and further suppose that  $\kappa$  and  $\lambda$  differ in just the  $j_1$ th and  $j_2$ th elements. Then

$$\begin{aligned} c_\kappa - c_\lambda &= -j_1 + j_2 + \frac{\alpha}{2} (k_{j_1} + l_{j_1} - k_{j_2} - l_{j_2}) \\ &= -(p+1-j_2) + (p+1-j_1) + \frac{\alpha}{2} [q - k_{j_2} + q - l_{j_2} - (q - k_{j_1}) - (q - l_{j_1})] \\ &= c_{\bar{\kappa}} - c_{\bar{\lambda}} \end{aligned}$$

#### Theorem 4.32

Suppose that  $\kappa > \lambda$ . Then  $\bar{\kappa} > \bar{\lambda}$ ; and

$$c_\kappa - c_\lambda = c_{\bar{\kappa}} - c_{\bar{\lambda}}$$

#### Proof 4.32

Apply Lemma 3.2 on p. 82 and Lemma 4.31. ■

#### 4.6.3.2 The reflection theorems

#### Theorem 4.33

$$j_{\rho, \lambda}^m = j_{\bar{\rho}, \bar{\lambda}}^m$$

#### Proof 4.33

From Lemmas 4.27 and 4.28 one has

$$(\Omega_{\rho, 0}^m)_{[\rho, \lambda]} = (\Omega_{\bar{\rho}, 0}^m)_{[\bar{\rho}, \bar{\lambda}]}$$

The result now follows from Lemma 4.17. ■

#### Theorem 4.34

If  $p = q$ , then

$$j_{\rho, \lambda}^e = j_{\bar{\rho}, \bar{\lambda}}^e$$

**Proof 4.34**

From Corollary 4.27 and Lemma 4.30, one has

$$(\Omega_{\rho,0}^e)_{[\lambda,\rho']} = (\Omega_{\bar{\rho},0}^e)_{[\bar{\lambda},\bar{\rho}']} = (\Omega_{\bar{\rho},0}^e)_{[\bar{\lambda},\bar{\rho}']}$$

where we have in addition used Lemma 4.29. The result now follows from Theorem 4.17. ■

**4.6.3.3 Example**

**Example 4.35**

Let  $\rho = (3, 2) = (3, 2, 0)$  and  $\lambda = (2, 2, 1)$ . Set  $h = 4$ ,  $r = 3$ , whereby  $\bar{\rho} = (4 - 3, 4 - 2, 4 - 0) = (4, 2, 1)$ , and  $\bar{\lambda} = (4 - 2, 4 - 2, 4 - 1) = (3, 2, 2)$ . Thus  $j_{(3,2),(2,2,1)}^m = j_{(4,2,1),(3,2,2)}^m$  and  $j_{(3,2),(2,2,1)}^e = j_{(4,2,1),(3,2,2)}^e$ . ■

## 4.7 Relative parsimony of expansions in $e_\lambda$ and $m_\lambda$ functions

It is clear intuitively that expansion in  $m_\lambda$  functions may be more parsimonious in general than expansion in  $e_\lambda$  functions, because the  $m_\lambda$  expansion tends to employ the bottom right corner of the  $\Omega^m$  matrix and delete the top left corner, while the reverse applies for the  $e_\lambda$  expansion. But the top left corner contains more matrix elements and provides the springboard for more d2 paths than does the bottom right corner.

We are unable to prove generally that expansions in  $m_\lambda$  functions are more parsimonious than those utilising  $e_\lambda$  functions, i.e. that their expansions contain fewer terms. But the circumstantial evidence is reasonably convincing. We shall illustrate with  $\Omega^m$ , so that for expansion in  $e_\lambda$  functions one moves in a d2 path above the diagonal from the bottom right corner.

Firstly we note that the first row of  $\Omega_w^m$  contains  $\lfloor \frac{w}{2} \rfloor$  elements, disregarding the diagonal element; while the final column contains but one element above the diagonal. The expansion of  $J_{(w-1,1)}$  in  $e_\lambda$  functions certainly contains more terms than the expansion in terms of  $m_\lambda$  functions, since removing the final column does nothing to reduce the number of d2 paths, while deleting the first row removes several paths.

Reasoning further along this same track, the number of elements to the right of the diagonal for rows indexed by partitions of length not exceeding 2 is easily shown to be  $3w^2/8 + O(w)$ ; but the number of elements above the diagonal for their conjugate partitions, viz.  $(2, 2, 2, \dots), \dots, (1^w)$ , or equivalently those partitions with height not exceeding 2, is just  $2w + O(1)$ . See Lemma 4.36.

**Lemma 4.36**

Let  $nrow_\lambda$  denote the number of terms in the  $\lambda$ th row of  $\Omega_w^m$ , excluding the diagonal element; and let  $ncol_\lambda$  denote the number of terms in the  $\lambda$ th column of  $\Omega_w^m$ , excluding the diagonal element. Then

$$\sum_{\ell(\lambda) \leq 2} nrow_\lambda = \begin{cases} \frac{w(3w-2)}{8} & \text{when } w \text{ is even} \\ \frac{(w-1)(3w-1)}{8} & \text{when } w \text{ is odd} \end{cases}$$

$$\sum_{h(\lambda) \leq 2} ncol_\lambda = \begin{cases} 2w-3 & \text{when } w \text{ is even} \\ 2w-4 & \text{when } w \text{ is odd} \end{cases}$$

**Proof 4.36**

First considering  $nrow_\kappa$ , and applying reasoning similar to that in Thm 4.3 on p. 101, the number of terms  $\omega_{\kappa,\lambda}$  with  $\ell(\kappa) \leq 2$  and  $\ell(\lambda) \leq 2$  is

$$\frac{\lfloor \frac{w}{2} \rfloor (\lfloor \frac{w}{2} \rfloor + 1)}{2}$$

To this must be added the number of terms with  $\ell(\kappa) = 2$  and  $\ell(\lambda) = 3$ , which can arise from the pairings  $\kappa = (w-t, 0)$  and  $(t, 0)$ . The number of these terms arising is respectively  $\lfloor \frac{w-t}{2} \rfloor$  and  $\lfloor \frac{t}{2} \rfloor$ . So

$$\sum_{\ell(\kappa) \leq 2} nrow_\kappa = \frac{\lfloor \frac{w}{2} \rfloor (\lfloor \frac{w}{2} \rfloor + 1)}{2} + \sum_{t=1}^{\lfloor \frac{w}{2} \rfloor} \left( \left\lfloor \frac{w-t}{2} \right\rfloor + \left\lfloor \frac{t}{2} \right\rfloor \right)$$

in which the final term of the summation is to be divided by 2 when  $w$  is even.

This may be rewritten more simply as

$$\sum_{\ell(\kappa) \leq 2} nrow_{\kappa} = \frac{\left\lfloor \frac{w}{2} \right\rfloor \left( \left\lfloor \frac{w}{2} \right\rfloor + 1 \right)}{2} + \sum_{t=1}^{w-1} \left\lfloor \frac{t}{2} \right\rfloor$$

For even  $w$ ,

$$\sum_{\ell(\kappa) \leq 2} nrow_{\kappa} = \frac{\frac{w}{2} \left( \frac{w}{2} + 1 \right)}{2} + 2 \sum_{t=1}^{\frac{w}{2}-1} t = \frac{w(w+2)}{8} + \left( \frac{w}{2} - 1 \right) \frac{w}{2} = \frac{w(3w-2)}{8};$$

while for odd  $w$ ,

$$\sum_{\ell(\kappa) \leq 2} nrow_{\kappa} = \frac{\frac{w-1}{2} \left( \frac{w-1}{2} + 1 \right)}{2} + 2 \sum_{t=1}^{\frac{w-3}{2}} t + \frac{w-1}{2} = \frac{(w-1)(3w-1)}{8}.$$

The formulae for  $ncol_{\lambda}$  follow more simply.

A partition  $\lambda$  of the form  $(2^a 1^b)$ , for  $a > 1, b > 1$ , has four elements above it in the  $\Omega$  matrix: two from the  $(2,2)$  pairing, to give  $(3,1)$  and  $(4,0)$ ; one from  $(2,1)$  to give  $(3,0)$ ; and one from  $(1,1)$  to give  $(2,0)$ . The partition  $(1^w)$  has but one element above its diagonal position, while  $(2, 1^{w-2})$  has two elements above it.

When  $w$  is even,  $(2^{w/2})$  has 2 elements above it, giving a total of

$$4 \left( \frac{w}{2} - 2 \right) + 1 + 2 + 2 = 2w - 3$$

elements above the diagonal elements  $\lambda$  such that  $h(\lambda) \leq 2$ .

When  $w$  is odd and  $w > 1$ , the partition  $(2^{(w-1)/2}, 1)$  has 3 elements above the diagonal, giving a total of

$$4 \left( \frac{w-1}{2} - 2 \right) + 1 + 2 + 3 = 2w - 4$$

elements above the diagonal elements  $\lambda$  such that  $h(\lambda) \leq 2$ . ■

## Chapter 5

# The operator and coefficient matrices $\Omega^m$ and $C^m$ for indeterminate weight $w$

The purpose of this chapter is to place the treatment of the operator and coefficient matrices in the last two chapters in a context valid for general weight  $w$ , however large.

We find the  $j_{\rho,\lambda}^m$  coefficients for partitions  $\rho$  and  $\lambda$  of a specific simple form, and also for a handful of values of  $\lambda$  low in the RLO for any  $\rho$ . To prove these results it is not necessary that  $w \rightarrow \infty$ , merely that  $w$  be sufficiently large that the partitions considered are listed in the standard non-increasing manner. There is circumstantial evidence that the results are in fact valid for any positive values of  $w$ .

The operator matrices  $\Omega^m$  and  $\Omega^e$  will be labelled as  $g\Omega^m$  and  $g\Omega^e$  in this chapter in part as a reminder that we are dealing with a generalised situation in which  $w$  is indefinite. In like manner the coefficient matrices  $C^m$  and  $C^e$  will be denoted by  $gC^m$  and  $gC^e$ . The generalised matrices  $g\Omega^m$ ,  $g\Omega^e$ ,  $gC^m$  and  $gC^e$  have infinitely many rows and columns, whereas  $\Omega^m$ ,  $\Omega^e$ ,  $C^m$  and  $C^e$  are  $p(w) \times p(w)$ .

## 5.1 The generalised operator matrix $g\Omega^m$

The rows and columns in Table 5.1 are indexed by the partitions  $(w)$ ,  $(w-1, 1)$ ,  $(w-2, 2)$ ,  $(w-2, 1, 1)$ ,  $\dots$ , with entries calculated from Theorem 3.18 on the assumption that  $w$  is arbitrarily large.

	$(w)$	$(w-1, 1)$	$(w-2, 2)$	$(w-2, 1^2)$	$(w-3, 3)$	$(w-3, 2, 1)$	$(w-3, 1^3)$
$(w)$	$c_{(w)}$	$1 \times w$	$1 \times w$	0	$1 \times w$	0	0
$(w-1, 1)$	0	$c_{(w-1,1)}$	$1(w-2)$	$2(w-1)$	$1(w-2)$	$1(w-1)$	0
$(w-2, 2)$	0	0	$c_{(w-2,2)}$	$1 \times 2$	$1(w-4)$	$1(w-2)$	0
$(w-2, 1, 1)$	0	0	0	$c_{(w-2,1^2)}$	0	$1(w-3)$	$3(w-2)$
$(w-3, 3)$	0	0	0	0	$c_{(w-3,3)}$	$1 \times 3$	0
$(w-3, 2, 1)$	0	0	0	0	0	$c_{(w-3,2,1)}$	$3 \times 2$
$(w-3, 1^3)$	0	0	0	0	0	0	$c_{(w-3,1^3)}$

Table 5.1: The generalised operator matrix  $g\Omega^m$  up to the 3rd block

At first sight, for a given value of  $w$  the generalised operator matrix  $g\Omega^m$  is only a valid construct as far as the  $[w/2]$ th block: up to that block the  $g\Omega^m$  matrix mimics the standard  $\Omega_w^m$  matrix, while beyond that block some partitions are in non-standard form, with the leading element not necessarily the largest. But it is likely that every column of the generalised operator matrix  $g\Omega$  provides a valid constraint on the  $j_{\rho,\lambda}^m$  coefficients, provided that the eigenvalues  $c_\lambda$  are calculated as if the first element were in fact the largest: Roberts (1998, Ch. 7) illustrates this for the case of  $\alpha = 2$  through an extended example.

Were this conjecture indeed to be true, i.e. if every column of  $g\Omega^m$  gave rise to a valid constraint on the  $j^m$  coefficients along the lines of (4.7) on p. 105, one could consider taking linear combinations of columns in  $g\Omega^m$  to produce new constraints on the  $j^m$  coefficients. The resulting inequalities and bounds on the coefficients would be valid for any value of  $w$ .

### Example 5.1

For  $w = 6$ , for instance, up to the 3rd block, viz. the block headed by partitions with leading element  $w - 3$ , the partitions are in standard form, and the top left corner of the  $g\Omega^m$  matrix reproduces the top left corner of the standard  $\Omega_6^m$  matrix.

For  $w = 4$ , on the other hand, and setting  $\rho = (4)$ , the column in (4.7) on p. 105 indexed by  $(w - 3, 2, 1) = (1, 2, 1)$  produces the constraint

$$3j_{\rho,(3,1)}^m + 2j_{\rho,(2,2)}^m + j_{\rho,(2,1,1)}^m + 3j_{\rho,(1,3)}^m = (4 + 5\alpha)j_{\rho,(1,2,1)}^m \quad (5.1)$$

which corresponds to the  $(w - 3, 2, 1)$ th column of Table 5.1.

This constraint is valid, provided that one identifies partitions  $(3, 1)$  and  $(1, 3)$  on the one hand, and  $(2, 1, 1)$  and  $(1, 2, 1)$  on the other. For the constraint arising from the  $(w - 2, 1, 1)$ th column in Table 5.1 is

$$6j_{\rho,(3,1)}^m + 2j_{\rho,(2,2)}^m = (3 + 5\alpha)j_{\rho,(2,1,1)}^m,$$

which is identical to (5.1). ■

The  $s$ th diagonal block in  $g\Omega^m$  contains the matrix  $\Omega_s^m$ , for  $s = 1, 2, \dots$ , save that the diagonal elements are displaced by a constant, from (4.17) on p. 113. The constancy of the off-diagonal elements is a consequence of Theorem 3.18 on p. 94.

#### 5.1.1 The generalised coefficient matrix $gC^m$

Rows and columns in Table 5.2 are still indexed by the partitions  $(w)$ ,  $(w - 1, 1)$ ,  $(w - 2, 2)$ ,  $(w - 2, 1, 1)$ ,  $\dots$ , but the notation is compacted. Thus partition 11 in the 4th group is  $(w - 4, 2, 1, 1)$ , with  $(2, 1, 1)$  being the 4th partition of 4 in the RLO; and partition 17 in the 5th group is  $(w - 5, 2, 2, 1)$ , since  $(2, 2, 1)$  is the 5th partition of 5 in the RLO. This matrix will be called the generalised Jack polynomial coefficient matrix, and denoted by  $gC^m$ . The diagonal elements are set to unity.

Our main result, Theorem 5.9, gives the entries marked  $X$  in  $gC^m$ . Within each block (bordered by lines), elements to the left of the  $X$  elements are zero: for those

pn		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
block		0	1	2	2	3	3	3	4	4	4	4	4	5	5	5	5	5	5	5
1	0	1	X	X		X	W		X					X						
2	1		1	X		X			X					X						
3	2			1		X			X					X						
4	2				1		X			X					X	W				
5	3					1			X					X						
6	3						1		0	X				0	X					
7	3							1	0	0	0	X		0	0	0	X			
8	4								1					<u>X</u>						
9	4									1					<u>X</u>					
10	4										1					<u>X</u>				
11	4											1					<u>X</u>			
12	4												1						<u>X</u>	
13	5													1						
14	5														1					
15	5															1				
16	5																1			
17	5																	1		
18	5																		1	
19	5																			1

Table 5.2: The generalised coefficient matrix  $gC^m$  up to the 5th block

elements the partition indexing the row does not majorise the partition indexing the column. Those zeroes are inserted explicitly in the (3,4)th and (3,5)th blocks, using an obvious notation, in order to illustrate.

We commence with the definition of the row and column partitions which define the elements  $X$ .

### Definition 5.2

A  $(w, s, b, \nu)$ -standard partition, denoted by  $\lambda_{s,b,\nu}$ , has the form  $\lambda_{s,b,\nu} = (w - s, s - b) \cup \nu$ , where  $0 \leq b < s$ ,  $w(\nu) = b$ ,  $h(\nu) \leq s - b$ , and  $s \leq w/2$ . ■

For fixed  $b$ , and setting  $\lambda_{r,b,\nu}$  to represent the row partition and  $\lambda_{t,b,\nu}$  the column partition, the  $(\lambda_{r,b,\nu}, \lambda_{t,b,\nu})$ th elements in the  $gC^m$  matrix are precisely those marked with an  $X$  in Fig. 5.2, for  $t > r$ . As  $\nu$  varies, one moves from one  $X$  symbol to another within the  $(r, t)$ th block. The particular form of  $\nu$  is immaterial for the principal theorem in the next section.

The pattern of  $X$ s within a block stabilises as we move to the right along the same row block: the pattern arising from the  $r$ th row block is stable by the  $2(r - 1)$ th column block. That is, with root partition  $(w - r, \dots)$ , the pattern is stable by the column block containing elements of the form  $(w - 2(r - 1), \dots)$ .

### Example 5.3

Setting  $r = 4$  we have the following table

$$\begin{array}{llll}
 b = 0 : & (w - 4, 4, 0) & \rightarrow (w - 5, 5, 0) & \rightarrow (w - 6, 6, 0) \rightarrow \dots \\
 b = 1 : & (w - 4, 3, 1) & \rightarrow (w - 5, 4, 1) & \rightarrow (w - 6, 5, 1) \rightarrow \dots \\
 b = 2 : & (w - 4, 2, 2) & \rightarrow (w - 5, 3, 2) & \rightarrow (w - 6, 4, 2) \rightarrow \dots \\
 b = 2 : & (w - 4, 2, 1, 1) & \rightarrow (w - 5, 3, 1, 1) & \rightarrow (w - 6, 4, 1, 1) \rightarrow \dots \\
 b = 3 : & (w - 4, 1, 1, 1, 1) & \rightarrow (w - 5, 2, 1, 1, 1) & \rightarrow (w - 6, 3, 1, 1, 1) \rightarrow \dots
 \end{array}$$

The corresponding  $X$  elements in the 5th column block are underscored in Table 5.2. ■

### 5.1.1.1 The Principal Theorem

The proof of Theorem 5.9 depends on the following simple lemmas, the proofs of which are omitted. The hypergeometric expansion  $(1 - x)^{-1/\alpha}$  provides several relationships for Jack polynomials, for an example of which see (2.7) on p. 29.

#### Lemma 5.4

Letting  ${}_1F_0(1/\alpha; y) = (1 - y)^{-1/\alpha} = \sum_{j=0}^{\infty} \beta_j y^j$ , we have

$$\beta_j = \prod_{k=0}^{j-1} \left( \frac{1}{\alpha} + k \right) / j! \quad \text{and} \quad \frac{\beta_j}{\beta_{j-1}} = \frac{\alpha j + 1 - \alpha}{\alpha j} \quad \blacksquare$$

#### Lemma 5.5

Suppose that  $\lambda_{r,b,\nu} > \kappa > \lambda_{t,b,\nu}$ . Then  $\kappa = \lambda_{s,b,\nu}$ , for some  $s$  such that  $r < s < t$ .  $\blacksquare$

#### Lemma 5.6

Provided that  $s < t$ ,

$$\omega^m(\lambda_{s,b,\nu}, \lambda_{t,b,\nu}) = w + b - 2s$$

#### Proof 5.6

Proof is from Theorem 3.18 on p. 94.  $\blacksquare$

#### Lemma 5.7

For  $w \geq 2t$ ,

$$c(\lambda_{r,b,\nu}) - c(\lambda_{t,b,\nu}) = (t - r) [1 + \alpha \{w + b - (t + r)\}]$$

#### Proof 5.7

Proof is from (3.4) on p. 81.  $\blacksquare$

#### Lemma 5.8

$$\frac{c(\lambda_{r,b,\nu}) - c(\lambda_{t-1,b,\nu}) + \omega^m(\lambda_{t-1,b,\nu}, \lambda_{t,b,\nu})}{c(\lambda_{r,b,\nu}) - c(\lambda_{t,b,\nu})}$$

$$= 1 + \frac{(\alpha - 1)(2t - w - b - 1)}{(t - r)[1 + \alpha(w + b) - \alpha(t + r)]}$$

**Theorem 5.9**

Let  $(1 - y)^{-1/\alpha} = \sum_{j=0}^{\infty} \beta_j y^j$ , define  $\lambda_{s,b,\nu}$  as in Definition 5.2, let  $r \leq t \leq w/2$ , and set  $x = w - 2r + b$ . Then, for  $r < t$ ,

$$j^m(\lambda_{r,b,\nu}, \lambda_{t,b,\nu}) = \beta_{t-r} \frac{\prod_{j=1}^{t-r} (x - j + 1)}{\prod_{j=1}^{t-r} (x - j + \frac{1}{\alpha})} \quad (5.2)$$

while for  $r = t$ ,

$$j^m(\lambda_{r,b,\nu}, \lambda_{r,b,\nu}) = \beta_0 = 1$$

Note that the particular form of  $\nu$  is immaterial for this result.

**Proof 5.9**

In order to simplify notation, set  $j^m(\lambda_{r,b,\nu}, \lambda_{s,b,\nu}) = j_{r,s}$ ,  $\omega^m(\lambda_{s,b,\nu}, \lambda_{t,b,\nu}) = \omega_{s,t}$  and  $c(\lambda_{s,b,\nu}) = c_s$ .

As in (4.7) on p. 105, the constraint on the  $j_{\rho,\lambda}^m$  coefficients from the column in  $g\Omega^m$  indexed by  $\lambda_{s,b}$  becomes

$$\sum_{j=0}^{s-r-1} j_{r,r+j} \omega_{r+j,s} = (c_r - c_s) j_{r,s} \quad (5.3)$$

For  $s = r + 1$ , an application of Lemmas 5.6 and 5.7 to (5.3) shows that

$$j_{r,r+1} = \frac{w - 2r + b}{\alpha(w + b) - \alpha(2r + 1) + 1} = \frac{1}{\alpha} \times \frac{x}{x - 1 + \frac{1}{\alpha}}$$

thereby establishing (5.2) when  $t = r + 1$ , since  $\beta_1 = \frac{1}{\alpha}$ .

For the general inductive proof, suppose that

$$j_{r,s} = \beta_{s-r} \frac{\prod_{j=1}^{s-r} (x - j + 1)}{\prod_{j=1}^{s-r} (x - j + \frac{1}{\alpha})},$$

for  $s = r + 1, \dots, t - 1$ , inclusive.

Setting  $s = t - 1$  and  $s = t$  in (5.3) leads respectively to

$$\sum_{j=0}^{t-r-2} j_{r,r+j} \omega_{r+j,t-1} = (c_r - c_{t-1}) j_{r,t-1} \quad (5.4)$$

and

$$\sum_{j=0}^{t-r-1} j_{r,r+j} \omega_{r+j,t} = (c_r - c_t) j_{r,t} \quad (5.5)$$

Noting from Lemma 5.6 that  $\omega_{su}$  does not depend on  $u$ , we can substitute (5.4) into (5.5) to obtain

$$(c_r - c_{t-1}) j_{r,t-1} + j_{r,t-1} \omega_{t-1,t} = (c_r - c_t) j_{r,t} ;$$

from which, after simplification using Lemmas 5.8 and 5.4,

$$\frac{j_{r,t}}{j_{r,t-1}} = \frac{c_r - c_{t-1} + \omega_{t-1,t}}{c_r - c_t} = \frac{\beta_{t-r}}{\beta_{t-r-1}} \frac{x - (t - r - 1)}{x - (t - r - \frac{1}{\alpha})} ,$$

thereby establishing the result. ■

To this more general framework one can apply the slicing theorems given in §4.6.2.2. The following theorems are immediate corollaries of Theorems 4.22 and 4.24:

**Theorem 5.10**

Let  $\rho = (w - r) \cup \sigma$ ,  $\lambda = (w - r) \cup \tau$ , where  $w(\sigma) = w(\tau) = r$  and  $w - r \geq \max(h(\sigma), h(\tau))$ . Then

$$j^m(\rho, \lambda) = j^m(\sigma, \tau) \quad \blacksquare$$

**Theorem 5.11**

Let  $\rho = (w - r) \cup \sigma$ ,  $\lambda = (w - t) \cup \tau$ , with  $t = w(\tau) \geq r = w(\sigma) \geq \max(h(\sigma), h(\tau))$ ; and suppose that  $\ell(\sigma) = \ell(\tau) = p$ . Set  $\rho_1 = (n - r - 1) \cup [\sigma - (1^p)]$ ,  $\lambda_1 = (n - t - 1) \cup [\tau - (1^p)]$ . Then

$$j^m(\rho, \lambda) = j^m(\rho_1, \lambda_1)$$

**Example 5.11**

In order to illustrate the theorem, let  $\rho = (w - 2, 1, 1)$ ,  $\lambda = (w - 5, 3, 2)$ ,  $w' = w - 3$ . Then, for  $w \geq 5$ ,

$$j^m(\rho, \lambda) = j^m((w'), (w' - 3, 2, 1))$$

Setting  $w'$  to  $w$ , the elements indicated by  $W$  in Table 5.2 are equal. ■

### 5.1.1.2 Consistency of Theorem 5.9 with the Knop-Sahi result

One consequence of Theorem 1.1 in Knop & Sahi (1997) is that in the expansion  $J_\rho = \sum_\lambda j_{\rho,\lambda} m_\lambda$ , the coefficients  $j_{\rho,\lambda}$  are polynomials in  $\alpha$  with positive integral coefficients. We exhibit consistency between this result and Theorem 5.9.

In this thesis we have generally used  $J_\rho$  to denote the Jack polynomial with indeterminate normalisation. In the literature, however, the symbol  $J_\rho$  is generally understood to incorporate the normalisation given in Stanley (1989, p. 97) as

$$j_{\lambda,\lambda} = \prod_{(i,j) \in \rho} [l'_j - i + 1 + \alpha(l_i - j)] \quad (5.6)$$

where the product is taken over the points  $(i, j)$  of the Ferrers diagram for  $\rho$ .

For this section only, denote by  $\tilde{J}$  the Jack polynomial with conventional normalisation, and by  $\hat{J}$  the Jack polynomial with the normalisation used in this thesis. Then set

$$\tilde{J}_\rho = \sum_{\kappa} \tilde{j}_{\rho,\kappa} m_\kappa \quad \text{and} \quad \hat{J}_\rho = \sum_{\kappa} \hat{j}_{\rho,\kappa} m_\kappa$$

so that  $\tilde{j}_{\lambda,\lambda}$  is given by (5.6), and  $\hat{j}_{\lambda,\lambda} = 1$ . Moreover  $\tilde{j}_{\rho,\kappa} = \tilde{j}_{\rho,\rho} \times \hat{j}_{\rho,\kappa}$ . We verify that  $\tilde{j}_{\rho,\rho} \times \hat{j}_{\rho,\kappa}$  is a polynomial in  $\alpha$  for  $\rho = \lambda_{r,b,\nu}$  and  $\kappa = \lambda_{t,b,\nu}$ .

For the partition  $(w-r, r-b) \cup \nu$ , restrict the product in (5.6) to the final  $w-2r+b = x$  elements in the first row of the Ferrers diagram, which is the part of the first row jutting out from the lower parts of the diagram. Then

$$\tilde{j}(\lambda_{r,b,\nu}, \lambda_{r,b,\nu}) = L_1 \prod_{k=1}^x [1 + \alpha(x - k)] \quad (5.7)$$

where  $L_1$  is a polynomial in  $\alpha$ , containing the remaining elements of the product in (5.6). From (5.2), on the other hand,

$$\hat{j}(\lambda_{r,b,\nu}, \lambda_{t,b,\nu}) = L_2 \times \frac{\beta_{t-r}}{\prod_{j=1}^{t-r} (x - j + \frac{1}{\alpha})}$$

where  $L_2$  is a polynomial in  $\alpha$ .

From the hypergeometric expansion given in Lemma 5.4

$$\frac{(t-r)! \beta_{t-r}}{\prod_{j=1}^{t-r} (x-j+\frac{1}{\alpha})} = \frac{\prod_{j=0}^{t-r-1} (\frac{1}{\alpha} + j)}{\prod_{j=1}^{t-r} (x-j+\frac{1}{\alpha})} = \frac{\prod_{j=0}^{t-r-1} (1 + \alpha j)}{\prod_{j=1}^{t-r} [\alpha(x-j) + 1]}$$

one has

$$\widehat{j}(\lambda_{r,b,\nu}, \lambda_{t,b,\nu}) = \frac{L_3}{\prod_{j=1}^{t-r} [\alpha(x-j) + 1]} \quad (5.8)$$

where  $L_3$  is a polynomial in  $\alpha$ .

The maximum value of  $t$  for which the partition  $\lambda_{t,b,\nu}$  remains in standard listing, i.e. in non-increasing order of elements of the partition, is  $\lceil \frac{w+b}{2} \rceil$ , so that the maximum value of  $t-r$  is  $\lceil \frac{x}{2} \rceil$ . From (5.7) and (5.8) one sees that  $\widetilde{j}(\lambda_{r,b,\nu}, \lambda_{r,b,\nu}) \times \widehat{j}(\lambda_{r,b,\nu}, \lambda_{t,b,\nu})$  is a polynomial in  $\alpha$ .

## 5.2 The $j_{\rho,\lambda}^m$ coefficients for $\lambda$ low in the RLO

Our approach to the operator matrix  $\Omega^m$  started from the top left hand corner, and proceeded downwards by back substitution as in (4.7) on p. 105. For leading coefficient unity, the  $j_{\rho,\lambda}^m$  coefficients for low  $\lambda$  in the RLO are ratios of complicated polynomials in  $w$ .

In this section we shall commence from the bottom right extremity of  $\Omega^m$ , considering how some of the  $j_{\rho,\lambda}^m$  coefficients behave for very low  $\lambda$  without first working through the coefficients for higher  $\lambda$ . We shall see in Chapter 6 that the  $m_\lambda$  functions tend to be of higher order for low  $\lambda$  than for high  $\lambda$ , highlighting the problem that the terms for low  $\lambda$  in (1.1) may well dominate for large partition weights: but our simple approach here can only deal with a handful of coefficients.

We utilise the fact that the Jack polynomials are known when all the variables assume the value unity, and equate coefficients of the highest powers of  $n$ . Combining these results with constraints available from the final columns of the operator matrix allows one to identify three of the extreme rightmost coefficients, using the final coefficient as numeraire.

Let  $m_\lambda(I)$  denote the value of  $m_\lambda$  when all variables assume the value unity.

### Definition 5.12

Let  $e_j(a, d, n)$  denote the  $e_j$  function evaluated for when the  $n$  variables form an arithmetic progression, with first member  $a$  and common difference  $d$ : thus the  $n$  variables are  $a + jd$ , with  $j = 0, 1, \dots, n-1$ . ■

### Lemma 5.13

$$\frac{e_1(a, d, n)}{n} = a + \frac{d}{2} (n-1) \quad (5.9)$$

$$\frac{e_2(a, d, n)}{n(n-1)} = \frac{a^2}{2} + \frac{ad}{2} (n-1) + \frac{d^2}{24} (n-2)(3n-1) \quad (5.10)$$

### Proof 5.13

The proof requires the following relations (Abramowitz & Stegun (1964, eqns. 23.1.4 and 7, p. 804)):

$$\sum_{j=1}^r j = \frac{r(r+1)}{2}; \quad \sum_{j=1}^r j^2 = \frac{r(r+1)(2r+1)}{6}; \quad \sum_{j=1}^r j^3 = \left( \sum_{j=1}^r j \right)^2. \quad (5.11)$$

First note that

$$e_1(a, d, n) = a + (a + d) + (a + 2d) + \dots + (a + \{n-1\}d)$$

which easily reduces to (5.9) upon the application of the first of the relations (5.11).

As for the second relation,

$$\begin{aligned} e_2(a, d, n) &= \sum_{j=0}^{n-2} (a + jd) \sum_{k=j+1}^{n-1} (a + kd) \\ &= \sum_{j=0}^{n-2} (a + jd) e_1(a + [j+1]d, d, n-j-1) \end{aligned}$$

which simplifies to (5.10) upon applying (5.9) and (5.11). ■

Let  $\lambda_{11} = (1^w)$ ;  $\lambda_{21} = (2, 1^{w-2})$ ;  $\lambda_{22} = (2, 2, 1^{w-4})$ ; and  $\lambda_{31} = (3, 1^{w-3})$ ; and further abbreviate  $m_{\lambda_{ij}}(I)$  to  $m_{ij}(I)$ ,  $j_{\rho, \lambda_{ij}}$  to  $j_{ij}$  and  $c_{\lambda_{ij}}$  to  $c_{ij}$ .

**Lemma 5.14**

$$m_{11}(I) = \frac{1}{w!} \left[ n^w - \frac{w(w-1)}{2} n^{w-1} + \frac{w(w-1)(w-2)(3w-1)}{24} n^{w-2} + O(n^{w-3}) \right] \quad (5.12)$$

$$m_{21}(I) = \frac{1}{(w-2)!} \left[ n^{w-1} - \frac{(w-1)(w-2)}{2} n^{w-2} + O(n^{w-3}) \right] \quad (5.13)$$

$$m_{22}(I) = \frac{n^{w-2}}{2(w-4)!} + O(n^{w-3})$$

$$m_{31}(I) = \frac{n^{w-2}}{(w-3)!} + O(n^{w-3})$$

**Proof 5.14**

We have

$$m_{11}(I) = \binom{n}{w} = \binom{n}{w, n-w} = \frac{n(n-1) \dots (n-w+1)}{w!},$$

so that

$$w! m_{11}(I) = n^w - e_1(0, 1, w) n^{w-1} + e_2(0, 1, w) n^{w-2} + O(n^{w-3}),$$

which simplifies to (5.12) upon applying Lemma (5.13). In similar vein,

$$m_{21}(I) = \binom{n}{1, w-2, n-w+1} = \frac{n(n-1) \dots (n-w+2)}{(w-2)!}$$

$$(w-2)! m_{21}(I) = n^{w-1} - e_1(0, 1, w-1) n^{w-2} + O(n^{w-3}),$$

which likewise simplifies to (5.13). For the remainder of the proof, simply note that

$$m_{22}(I) = \binom{n}{2, w-4, n-w+2} = \frac{n(n-1) \dots (n-w+3)}{2(w-4)!}$$

$$m_{31}(I) = \binom{n}{1, w-3, n-w+2} = \frac{n(n-1) \dots (n-w+3)}{(w-3)!} \quad \blacksquare$$

Now, as  $n \rightarrow \infty$ ,

$$J_\rho(I) = j_{11}^m m_{11}(I) + j_{21}^m m_{21}(I) + j_{22}^m m_{22}(I) + j_{31}^m m_{31}(I) + O(n^{w-3}) \quad (5.14)$$

When the  $n$  variates are unity, Stanley (1989, p. 93) states that  $J_\rho$  has value

$$J_\rho(I) = \prod_{(i,j) \in \rho} [n - (i-1) + \alpha(j-1)] = n^w + g_1 n^{w-1} + g_2 n^{w-2} + O(n^{w-3}) \quad (5.15)$$

where the first equation is taken from Stanley (1989, p. 93), with the product taken over the Ferrers diagram for  $\rho$ . The highest order coefficients  $g_j$  are found in Theorem 5.20, and in fact  $g_1$  is given in Stanley (1989, p. 106). One notes some resemblance between (5.6) and (5.15).

We shall use the coefficient  $j_{11}^m$  as numeraire, and define

$$\zeta_{ij} = \frac{j_{ij}^m}{j_{11}^m} \quad (5.16)$$

Equating coefficients between (5.14) and (5.15) will give us relations (5.17), (5.18) and (5.19) concerning the four  $j_{\rho,\lambda}$  coefficients singled out above. Further relations (5.22) and (5.23) are obtained from the constraints implicit in the final 2 columns of  $\Omega_{\rho,1}$ .

### 5.2.0.1 Equating coefficients

Equating coefficients between (5.14) and (5.15), and utilising Lemma 5.14, one obtains the following equations:

$$\frac{j_{11}}{w!} = 1 \quad (5.17)$$

$$-\frac{j_{11}}{2(w-2)!} + \frac{j_{21}}{(w-2)!} = g_1 \quad (5.18)$$

$$\frac{j_{11}(3w-1)}{24(w-3)!} - \frac{j_{21}(w-1)}{2(w-3)!} + \frac{j_{22}}{2(w-4)!} + \frac{j_{31}}{(w-3)!} = g_2 \quad (5.19)$$

Simplification of (5.18) and (5.19) leads respectively to the following equations.

$$\zeta_{21} = \frac{1}{2} + \frac{g_1}{w_{(2)}} \quad (5.20)$$

$$\frac{3w-1}{12} - \zeta_{21}(w-1) + \zeta_{22}(w-3) + 2\zeta_{31} = \frac{2g_2}{w_{(3)}} \quad (5.21)$$

### 5.2.0.2 Constraints implicit in the two final columns of $\Omega_{\rho,1}^m$

#### Lemma 5.15

$$c_{11} = wn - \frac{w(w+1)}{2} \quad \text{and} \quad c_{21} = c_{11} + w + \alpha - 1$$

#### Proof 5.15

Proof is from (3.4) on p. 81. ■

The constraints implicit in the  $\lambda_{21}$ th and  $\lambda_{11}$ th columns of  $\Omega_{\rho,1}^m$  are

$$3(w-2)j_{31} + (w-2)(w-3)j_{22} = j_{21}(c_{\rho} - c_{21}) \quad (5.22)$$

$$w(w-1)j_{21} = j_{11}(c_{\rho} - c_{11}) \quad (5.23)$$

Applying (5.29) and Lemma 5.15, the equation (5.23) reproduces (5.20), while (5.22) becomes

$$3\zeta_{31} + (w-3)\zeta_{22} = \zeta_{21} \left( \frac{g_1 - \alpha}{w-2} + \frac{w-1}{2} \right) \quad (5.24)$$

### 5.2.0.3 Reassembling the jigsaw

Solving (5.21) and (5.24) leads to the following theorem.

#### Theorem 5.16

The coefficients  $\zeta_{21}$ ,  $\zeta_{22}$  and  $\zeta_{31}$  are given by (5.20),

$$\frac{1}{2}(w-3)\zeta_{22} = \frac{w-3}{8} + \frac{g_1}{w} - \frac{g_1 - \alpha}{2(w-2)} + \frac{3g_2 - g_1(g_1 - \alpha)}{w_{(3)}} \quad (5.25)$$

where  $w_{(k)}$  is the falling factorial; and

$$\zeta_{31} = \frac{1}{6} - \frac{g_1}{2w} + \frac{g_1 - \alpha}{2(w-2)} - \frac{2g_2 - g_1(g_1 - \alpha)}{w_{(3)}} \quad (5.26)$$

The solutions in Theorem 5.16 become rather more tractable for partitions  $\rho$  of length less than 3. ■

**Theorem 5.17**

When  $\ell(\rho) \leq 2$ ,

$$\frac{\zeta_{31}}{\alpha + 1} - \zeta_{21} = -\frac{\alpha + 2}{6}$$

or equivalently, from (5.16),

$$j_{31} - \alpha j_{21} = j_{21} - \frac{(\alpha + 1)(\alpha + 2)}{6} j_{11}$$

**Proof 5.17**

Subtracting (5.21) from (5.24) yields

$$\zeta_{31} - \frac{3w - 1}{12} + \zeta_{21}(w - 1) = \zeta_{21} \left( \frac{g_1 - \alpha}{w - 2} + \frac{w - 1}{2} \right) - \frac{2g_2}{w_{(3)}}$$

From Lemmas 5.18 and 5.21 one may write

$$\begin{aligned} & \zeta_{31} - \frac{3w - 1}{12} \\ &= \frac{\alpha + 1}{4} [\alpha(w + 1) - (w - 1)] + \frac{\phi}{w_{(3)}} [-\alpha w_{(2)} + \phi + \alpha + 1 - w] \\ & \quad - \frac{2}{w_{(3)}} \left[ \frac{1}{24} \alpha^2 w_{(3)} (3w - 1) - \frac{1}{2} \alpha w_{(2)} \phi + \frac{1}{2} \phi (\phi + \alpha w - \alpha - 1) \right] \end{aligned}$$

Simplifying, one gets

$$\begin{aligned} & \zeta_{31} + (\alpha^2 - 1) \frac{3w - 1}{12} \\ &= \frac{\alpha + 1}{4} [\alpha(w + 1) - (w - 1)] + \frac{\phi}{w_{(3)}} [\alpha + 1 - w] - \frac{\phi}{w_{(3)}} (\alpha w - \alpha - 1) \\ &= \frac{\alpha + 1}{4} [\alpha(w + 1) - (w - 1)] - (\alpha + 1) \frac{\phi}{w_{(2)}} \end{aligned} \tag{5.27}$$

From (5.20) and (5.38) one has

$$\frac{\phi}{w_{(2)}} = \frac{\alpha + 1}{2} - \zeta_{21},$$

substitution of which in (5.27) completes the proof. ■

**Lemma 5.18**

When  $\ell(\rho) \leq 2$ ,

$$\zeta_{21} \left( \frac{g_1 - \alpha}{w - 2} - \frac{w - 1}{2} \right)$$

$$= \frac{a+1}{4} [a(w+1) - (w-1)] + \frac{\phi}{w_{(3)}} [-aw_{(2)} + \phi + a + 1 - w] \quad (5.28)$$

**Proof 5.18**

The proof employs (5.20) and (5.38). The result may be verified using the code in §5.2.2. ■

**Theorem 5.19**

When  $\ell(\rho) \leq 2$ ,

$$\zeta_{22} = \frac{1}{4}(\alpha+1)^2 + \frac{\phi^2}{w_{(4)}} - \frac{\phi(\alpha+1)}{w_{(4)}} \left[ (w-2)^2 - \frac{w-2}{\alpha+1} + 1 \right]$$

**Proof 5.19**

Summing (5.25) and (5.26), simplification yields the result.

**Corollary 5.19**

When  $\rho = (w)$ ,

$$\zeta_{22} = \frac{1}{4}(\alpha+1)^2. \quad \blacksquare$$

Theorems 5.16, 5.17 and 5.19 extend results in Roberts (1998, §6.5) to general values of  $\alpha$ .

### 5.2.1 Finding the values of $g_j$

We shall give two ways of finding the  $g_1$  and  $g_2$  coefficients, the second of which is to be applied when  $\ell(\rho) = 2$ .

### 5.2.1.1 General approach

#### Theorem 5.20

For  $\rho = (r_1, r_2, \dots)$ , define

$$A = \sum_{s=1}^{\infty} sr_s; \quad B = \sum_{s=1}^{\infty} r_s^2; \quad C = \sum_{s=1}^{\infty} s^2 r_s; \quad D = \sum_{s=1}^{\infty} sr_s^2; \quad E = \sum_{s=1}^{\infty} r_s^3$$

For  $\rho' = (r'_1, r'_2, \dots)$ , define similarly  $A' = \sum sr'_s$  etc.

Then, in the expansion (5.15),

$$g_1 = c_\rho - (n-1)w = w(1-\alpha) + \alpha A' - A \quad (5.29)$$

$$g_2 = g_1^2 - (C + \alpha^2 C') + \alpha(D + A) + (\alpha-1)[2g_1 + (\alpha-1)w] \quad (5.30)$$

Also we note that

$$2A' = B + w \quad C' = \frac{1}{3}E + \frac{1}{2}B + \frac{1}{6}w \quad (5.31)$$

#### Proof 5.20

From (5.15)

$$g_1 = \sum_{(i,j) \in \lambda} [-(i-1) + \alpha(j-1)]$$

where the summation is over the cells of the Ferrers diagram. The final equation of (5.29) follows from the relations

$$\sum_{(i,j) \in \lambda} i = A \quad \text{and} \quad \sum_{(i,j) \in \lambda} j = A'$$

For (5.30), note the following relations

$$\sum_{\substack{(i,j) \in \lambda \\ \text{distinct pairs of cells}}} \sum_{(k,l) \in \lambda} i = A(w-1) \quad \text{and} \quad \sum_{\substack{(i,j) \in \lambda \\ \text{distinct pairs of cells}}} \sum_{(k,l) \in \lambda} j = A'(w-1) \quad (5.32)$$

$$\sum_{\substack{(i,j) \in \lambda \\ \text{distinct pairs of cells}}} \sum_{(k,l) \in \lambda} ik = A^2 - C \quad \text{and} \quad \sum_{\substack{(i,j) \in \lambda \\ \text{distinct pairs of cells}}} \sum_{(k,l) \in \lambda} il = AA' - \frac{1}{2}(A + D). \quad (5.33)$$

Now (5.30) and (5.31) follow from (5.32), (5.33) and (5.11). ■

### 5.2.1.2 Approach suitable for short $\rho$

We define

$$\begin{aligned} g_{11} &= \alpha \frac{r_1(r_1 - 1)}{2}; \quad g_{12} = -r_2 + \alpha \frac{r_2(r_2 - 1)}{2}; \\ g_{13} &= -2r_3 + \alpha \frac{r_3(r_3 - 1)}{2}, \text{ etc.} \end{aligned} \quad (5.34)$$

and set

$$G(x, \alpha, r) = \sum_{j=0}^{r-2} (x + j\alpha) \sum_{k=j+1}^{r-1} (x + k\alpha)$$

which can be evaluated from Lemma 5.13 as

$$\frac{G(x, \alpha, r)}{r(r-1)} = \frac{1}{2}\alpha x(r-1) + \frac{1}{2}x^2 + \frac{1}{24}\alpha^2(r-2)(3r-1) \quad (5.35)$$

One has in general

$$g_1 = g_{11} + g_{12} + g_{13} + \dots \quad (5.36)$$

and

$$\begin{aligned} g_2 &= G(0, \alpha, r_1) + G(-1, \alpha, r_2) + G(-2, \alpha, r_3) + \dots \\ &\quad + g_{11}g_{12} + g_{11}g_{13} + g_{12}g_{13} + \dots \end{aligned} \quad (5.37)$$

But with  $r_3 = 0$ ,

$$g_1 = g_{11} + g_{12} = \frac{\alpha}{2}w(w-1) - r_2(1 + \alpha r_1) = \frac{\alpha}{2}w_{(2)} - \phi \quad (5.38)$$

where

$$\phi = r_2(1 - \alpha r_1). \quad (5.39)$$

The basic result for  $g_2$  is the following.

#### Lemma 5.21

When  $\ell(\rho) \leq 2$ ,

$$g_2 = \frac{1}{24}\alpha^2 w_{(3)}(3w-1) - \frac{1}{2}\alpha w_{(2)}\phi + \frac{1}{2}\phi(\phi + \alpha w - \alpha - 1) \quad (5.40)$$

#### Proof 5.21

The proof is a straightforward but tedious manipulation employing (5.35). Maple code enabling verification of the result is included in §5.2.2. ■

In Chapter 6 some evidence is presented to show that  $m_\lambda$  functions may be of higher order when  $\lambda$  is low in the RLO, so that the  $j_{\rho,\lambda}^m$  coefficients may assume greater importance when  $\lambda$  is low. The simple methodology in §5.2 would however be hard pushed to provide more coefficients  $j_{\rho,\lambda}^m$  for  $\lambda$  at the lower extremity of the RLO.

## 5.2.2 Supporting Maple code

Some obvious abbreviations are made, e.g.  $\alpha$  is written as  $a$  in the code.

### 5.2.2.1 Maple code for Lemma 5.21

The code utilises successively (5.35), (5.34), (5.38), (5.39) and (5.37). The variable  $z$  is defined from (5.40). The final two lines in each case indicate the result of inputting the preceding code into Maple.

```
G:=(x,a,r)->(a*x*(r-1)/2+x^2/2+a^2*(r-2)*(3*r-1)/24)*r*(r-1);
g11:=a*r1*(r1-1)/2;
g12:=-r2+a*r2*(r2-1)/2;
g1:=g11+g12;
phi:=r2+a*r1*r2;
w:=r1+r2;
g2:=G(0,a,r1)+G(-1,a,r2)+g11*g12;
z:=24*g2-a^2*w*(w-1)*(w-2)*(3*w-1)+12*a*w*(w-1)*phi
-12*phi*(phi+a*w-a-1);
simplify(z);
> simplify(z);
```

## 5.2.2.2 Maple code for Lemma 5.18

The code utilises successively (5.34), (5.38) and (5.39). The variables *lhs* and *rhs* are an obvious notation from (5.28), save that they have been multiplied by a constant.

```

g11:=a*r1*(r1-1)/2;
g12:=-r2+a*r2*(r2-1)/2;
g1:=g11+g12;
phi:=r2+a*r1*r2;
w:=r1+r2;
lefths:=w*(w-1)*(w-2)*(1/2+g1/w/(w-1))*((g1-a)/(w-2)-(w-1)/2);
righths:=(1+a)*w*(w-1)*(w-2)/4*(a*(w+1)-(w-1))
+phi*(-a*w*(w-1)+phi+a+1-w);
simplify(lefths-righths);
> simplify(lefths-righths);

```

0

# Chapter 6

## Evaluation of the $m_\lambda$ functions

### 6.1 Summary

Most of the work in this chapter is effected by an expansion of  $am_\lambda$  functions in  $p_\lambda$  functions, proved as Theorem 6.1. This facilitates the calculation of  $m_\lambda$  functions when the variates assume a geometric progression, for the  $p_\lambda$  functions are then quickly found. But it also enables us to investigate the  $m_\lambda$  functions when the variates stand in arithmetic progression, at least as far as evaluating the higher order behaviour is concerned.

From Corollary 6.9, for variates in arithmetic progression, the asymptotic behaviour of  $m_\lambda$  functions is given by  $m_\lambda = O(n^{\ell(\lambda)})$  as  $n \rightarrow \infty$ . For variates in geometric progression, and letting  $n \rightarrow \infty$  for a fixed ratio between successive terms, the vector  $M$  of  $m_\lambda$  functions behaves independently of  $\ell(\lambda)$  at the highest order; but then letting the ratio of one term to the next  $\rightarrow 1$  means that the dominant  $m_\lambda$  functions are those for long  $\lambda$  (low in the RLO), according to Theorem 6.15. For more general variate behaviour one has no particular results. A little tentatively, one can perhaps say that there is some evidence that  $m_\lambda$  functions are of higher order when  $\lambda$  is low in the RLO.

## 6.2 Preliminary

According to standard results on transformations from one set of basis functions for the symmetric homogeneous polynomials to another basis, the expression of  $m_\lambda$  functions in terms of  $p_\lambda$  functions can be accomplished in two steps: for instance, the transformation from  $m_\lambda$  functions to  $s_\lambda$  functions involves the Kostka numbers, and the transformation from  $s_\lambda$  functions to  $p_\lambda$  functions involves the character table of the symmetric group (e.g., Macdonald (1995, §I.6), Stanley (1999, Ch. 7)).

Macdonald (1995, §I.6, Ex. 10) mentions the transition matrix from the  $p_\lambda$  basis to the  $am_\lambda$ , essentially derived as a variant on a matrix of Kostka numbers, but our Theorem 6.1 seems simple enough to derive directly.

Our first result Theorem 6.1 utilises a simple application of the inclusion exclusion principle to prove an expansion of the  $am_\lambda$  functions in terms of the  $p_\lambda$  functions. The expansion has the additional virtue of being relatively parsimonious, at least in comparison with the standard determinantal expansion of  $e_\lambda$  functions in terms of the  $p_\lambda$  functions (e.g., Macdonald (1995, §I.2), Aigner (1979, p. 163)). The latter expansion contains terms indexed by partitions of  $\lambda$ , whereas our result contains terms indexed by partitions of  $\ell(\lambda)$ . The main theorem is proved for indeterminate variates taken to indeterminate powers, although in the applications in later sections of this chapter we shall restrict those powers to be integers.

It is an easy application of Theorem 6.1 to find the  $m_\lambda$  functions when the variates stand in geometric progression, since the  $p_\lambda$  functions are then readily calculated; and it is an easy extension to let the number of variates become infinite. Before then, in a less obvious application, Theorem 6.1 will also be used to evaluate  $m_\lambda$  functions when variates assume an arithmetic progression, restricting ourselves to finding higher order terms only.

Let  $a, a+d, a+2d, \dots, a+(n-1)d$  be the  $n$  variates in arithmetic progression. The first situation investigated held  $d$  fixed as  $n \rightarrow \infty$ , so that the range of the  $n$  variates also becomes infinite; while in the second case considered  $nd$  was kept fixed, so that

the range of the variates remained constant while their number became infinite. In the former case  $m_\lambda = O(n^r)$  and in the latter case  $m_\lambda = O(n^{r+w})$ , where  $r = \ell(\lambda)$ .

In either case, then, one can expect that for large  $n$  the most important of the  $j_{\rho,\lambda}^m$  coefficients will be those for long partitions  $\lambda$ . That is unfortunate, in that the most difficult coefficients  $j_{\rho,\lambda}^m$  to calculate are those with short  $\rho$  (high in the RLO) and long  $\lambda$  (low in the RLO), since these are the end product of a long sequence of back substitutions; otherwise expressed,  $[\rho, \lambda]$  supports many d2 chains.

Before starting the subject matter of this chapter, it is perhaps worth noting that while variates are anticipated to be non-negative in most statistical applications of Jack polynomials, calculation of both the  $m_\lambda$  functions and the  $e_\lambda$  functions simplifies greatly when variates are symmetrically placed about 0: Theorem 7.1 on p. 166 refers.

### 6.3 The expansion of $am_\lambda$ in $p_\lambda$ functions

In this section it is convenient to extend the definition of  $p_j = \sum_{i=1}^n x_i^j$  to quantities  $j$  which are not necessarily non-negative integers. For a vector  $v = (i, j, k, \dots)$ , we write interchangeably  $p_v = p_{(i,j,k,\dots)} = p(i, j, k, \dots) = p(i)p(j)p(k) \dots = p_i p_j p_k \dots$ . Further, let  $S_r$  denote the symmetric group on  $r$  symbols, say the integers  $1, 2, \dots, r$ , so that the action of  $s \in S_r$  on  $i$  is given by  $si = s(i)$ . Finally, when  $\lambda$  is a partition of length  $r$ , the definition of  $am_{(l_1, l_2, \dots, l_r)}$  in (6.2) is consistent with that of  $am_\lambda$  given in §1.5 on p. 15.

Theorem 6.1 is valid for  $x_i$  and  $l_j$  being indeterminants such that the usual rules of exponents are followed. That is,  $\prod_j x_{k_j}^{l_j}$  is well defined,  $\prod_j x^{l_j} = x^{\sum_j l_j}$ ,  $x^0 = 1$  etc.

Formally one can proceed to define a  $\mathbb{Z}$ -module  $X$  containing finite linear combinations (with integral coefficients) of elements of the form  $x_1^{j_1} x_2^{j_2} \dots x_n^{j_n}$ , in which  $j_i \in J$ ; and where  $J$  is an additive abelian semi-group with identity, viz. zero. Thus  $i, j, k \in J$  implies that  $i + j = j + i \in J$ ,  $(i + j) + k = i + (j + k)$  and  $i + 0 = 0 + i = i$ .

A product in  $X$  is defined by

$$x_1^{i_1} x_2^{i_2} \dots x_n^{i_n} \times x_1^{j_1} x_2^{j_2} \dots x_n^{j_n} = x_1^{i_1+j_1} x_2^{i_2+j_2} \dots x_n^{i_n+j_n} \quad (6.1)$$

When  $j_i = 0$  for  $i \in I \subseteq \{1, 2, \dots, n\}$ , we abbreviate by writing

$$\prod_{i=1}^n x_i^{j_i} = \prod_{\substack{i=1 \\ i \notin I}}^n x_i^{j_i}$$

Thus  $X$  is the commutative ring with identity generated by the elements  $x_i^j$  for  $i = 1, \dots, n$  and  $j \in J$ . The identity is  $\prod_{i=1}^n x_i^0$ , and we adopt the convention that  $0^j = 0$  for all  $j \in J$ .

We work with  $X \otimes_{\mathbb{Z}} \mathbb{Q}$ , thereby allowing the coefficients to be rational numbers (see Atiyah & Macdonald (1969, pp. 1, 24); and Curtis & Reiner (1962) and Cohn (1989), say, for fuller treatments of the tensor product). In fact the only elements of  $X$  arising in this section are those with at most  $r$  indices non-zero. Should  $r > n$ , one can augment the set of indeterminants by  $r - n$  zeroes.

### Theorem 6.1

Let  $X$  be a commutative ring with identity generated by elements of the form  $x_i^j$  for  $i = 1, \dots, n$  and  $j \in J$ , where  $J$  is an additive abelian semi-group with identity. The product in  $X$  is defined as in (6.1).

Define “augmented monomial symmetric functions” as follows:

$$am_{(l_1, l_2, \dots, l_r)} = \sum_{\substack{\gamma_1, \gamma_2, \dots, \gamma_r=1 \\ d}}^n x_{\gamma_1}^{l_1} x_{\gamma_2}^{l_2} \dots x_{\gamma_r}^{l_r} \quad (6.2)$$

where the “d” underneath the summation sign indicates that only distinct values of the indices  $\gamma_j$  are to be used.

For  $\kappa = (k_1, k_2, \dots) \vdash r$ , where  $\kappa$  is listed in non-increasing order, set  $\sigma_{\kappa}(i) = \sum_{j=1}^i k_j$  for  $i \leq \ell(\kappa)$ , with  $\sigma_{\kappa}(0) = 0$ . Then

$$am_{(l_1, l_2, \dots, l_r)} = \sum_{\kappa \vdash r} \frac{(-1)^{\ell(\kappa)+r}}{z_{\kappa}} \sum_{s \in S_r} \prod_{j=1}^{\ell(\kappa)} p(l_{s(\sigma_{\kappa}(j-1)+1)} + l_{s(\sigma_{\kappa}(j-1)+2)} + \dots + l_{s(\sigma_{\kappa}(j))}) \quad (6.3)$$

in which  $p(j) = p_j = \sum_{i=1}^n x_i^j$ ; terms in the right side of (6.2) and (6.3) are to be interpreted as elements of  $X \otimes_{\mathbb{Z}} \mathbb{Q}$ ; and where  $z_{\kappa}$  is defined in §1.5 on p. 15, and is in any case given below in (6.4). ■

The  $am_{(l_1, l_2, \dots, l_r)}$  function reduces to the falling factorial when all variate values are unity: see Example 6.4. In particular,  $am_{(l_1, l_2, \dots, l_r)}$  vanishes when  $r > n$ .

It is convenient to decompose  $z_{\kappa}$  as follows.

$$\kappa = (1^{m_1}, 2^{m_2}, \dots) \Rightarrow z_{\kappa} = \prod i^{m_i} m_i! = \frac{\prod (i!)^{m_i} m_i!}{\prod ((i-1)!)^{m_i}} = \frac{y_{\kappa}}{\theta_{\kappa}} \quad (6.4)$$

in which the multiplicity  $m_i = m_i(\kappa)$ , as in §1.5.1 on p. 15. The constant  $z_{\kappa}$  is expressed as a ratio of  $\theta_{\kappa}$  and  $y_{\kappa}$  because in the conventional setting out as a sum of brackets of similar terms of “type”  $\kappa \vdash r$ , the brackets contain  $r!/y_{\kappa}$  terms, while  $\theta_{\kappa}$  is retained outside the  $\kappa$ th bracket as a weighting factor.

Some definitions and examples are given before commencing the proof.

### 6.3.1 Definitions

#### 6.3.1.1 The concave map associated with a partition $\kappa \vdash r$

For  $\kappa = (k_1, k_2, \dots, k_t) \vdash r$ , where  $\kappa$  is listed in standard non-increasing order with  $\ell(\kappa) = t$ , define a map

$$\sigma_{\kappa} : \{0, 1, 2, \dots, t\} \rightarrow \{0, 1, 2, \dots, r\}$$

by

$$\sigma_{\kappa} i = \sigma_{\kappa}(i) = \sum_{j=1}^i k_j \quad \text{for } 1 \leq i \leq t; \quad \text{and } \sigma_{\kappa} 0 = 0$$

In particular  $\sigma_{\kappa} t = r$ . Associated with such a map and a vector  $(l_1, l_2, \dots, l_r)$  one defines the function

$$q_{\sigma_{\kappa}} = p(l_1 + \dots + l_{\sigma_{\kappa} 1}) p(l_{\sigma_{\kappa} 1 + 1} + \dots + l_{\sigma_{\kappa} 2}) \dots p(l_{\sigma_{\kappa}(t-1) + 1} + \dots + l_r)$$

$$= \prod_{j=1}^{\ell(\kappa)} p(L_{\kappa,j}) \quad (6.5)$$

in which, for  $1 \leq j \leq \ell(\kappa)$ ,

$$L_{\kappa,j} = l_{\sigma_{\kappa}(j-1)+1} + \dots + l_{\sigma_{\kappa}j}.$$

For  $s \in S_r$ , and writing  $s\sigma i$  for  $s(\sigma i)$ , one further defines

$$\begin{aligned} L_{s\kappa,j} &= l_{s(\sigma_{\kappa}(j-1)+1)} + \dots + l_{s\sigma_{\kappa}j} \\ q_{s\sigma_{\kappa}} &= p(l_{s1} + \dots + l_{s\sigma_{\kappa}1}) p(l_{s(\sigma_{\kappa}1+1)} + \dots + l_{s\sigma_{\kappa}2}) \dots p(l_{s(\sigma_{\kappa}(t-1)+1)} + \dots + l_{sr}) \\ &= \prod_{j=1}^{\ell(\kappa)} p(L_{s\kappa,j}) \end{aligned}$$

### 6.3.1.2 The $\kappa$ th bracket $B_{\kappa}$

The  $\kappa$ th bracket is defined as:

$$B_{\kappa} = \frac{1}{y_{\kappa}} \sum_{s \in S_r} q_{s\sigma_{\kappa}}. \quad (6.6)$$

The number of monomial terms in  $B_{\kappa}$  is  $r!/y_{\kappa}$ , which is the number of ways of choosing subsets of sizes  $k_1, k_2, \dots$  from  $w(\kappa) = r$  objects. There are accordingly no duplicates of monomials in  $B_{\kappa}$ , unless different sums of distinct terms of the vector  $(l_1, l_2, \dots, l_r)$  happen to coincide, as illustrated in Examples 6.4 and 6.5.

Further definitions are given at the beginning of §6.4.

### 6.3.1.3 Rephrasing Theorem 6.1

The equation (6.3) can now be rewritten more compactly as

$$\begin{aligned} am_{(l_1, l_2, \dots, l_r)} &= \sum_{\kappa \vdash r} \frac{(-1)^{\ell(\kappa)+r}}{z_{\kappa}} \sum_{s \in S_r} \prod_{j=1}^{\ell(\kappa)} p(L_{s\kappa,j}) = \sum_{\kappa \vdash r} \frac{(-1)^{\ell(\kappa)+r}}{z_{\kappa}} \sum_{s \in S_r} q_{s\sigma_{\kappa}} \\ &= \sum_{\kappa \vdash r} (-1)^{\ell(\kappa)+r} \theta_{\kappa} B_{\kappa} \end{aligned} \quad (6.7)$$

### 6.3.2 Examples

Before proceeding to a proof of Theorem 6.1, we illustrate with examples.

#### Example 6.2

$$am_{(6,1)} = m_{(6,1)} = \sum_{\substack{i,j=1 \\ i \neq j}}^n x_i^6 x_j = \sum_{i,j=1}^n x_i^6 x_j - \sum_{i=1}^n x_i^7 = p_6 p_1 - p_7 \quad \blacksquare$$

#### Example 6.3

$$m_{(3,2,1)} = \sum_{\substack{i,j,k \\ d}} y_i^3 y_j^2 y_k = \sum_{i,j,k} y_i^3 y_j^2 y_k - \sum_{i,j} (y_i^5 y_j^1 + y_i^4 y_j^2 + y_i^3 y_j^3) + 2 \sum_i y_i^6,$$

which becomes

$$m_{(3,2,1)} = p_{(3,2,1)} - (p_{(5,1)} + p_{(4,2)} + p_{(3,3)}) + 2p_6 \quad \blacksquare$$

#### Example 6.4

$$\begin{aligned} am_{(i,j,k,l)} &= p_i p_j p_k p_l \\ &- (p_{i+j} p_k p_l + p_{i+k} p_j p_l + p_{i+l} p_j p_k + p_{j+k} p_i p_l + p_{j+l} p_i p_k + p_{k+l} p_i p_j) \\ &+ (p_{i+j} p_{k+l} + p_{i+k} p_{j+l} + p_{i+l} p_{j+k}) \\ &+ 2(p_{i+j+k} p_l + p_{i+j+l} p_k + p_{i+k+l} p_j + p_{j+k+l} p_i) - 6p_{i+j+k+l} \end{aligned} \quad (6.8)$$

The bracketed expressions, together with the first and last terms, are indexed by

$$\kappa = (1^4), (2, 1, 1), (2, 2), (3, 1), (4)$$

respectively. Recalling the definition of the  $\kappa$ th bracket in (6.6), equation (6.8) is recast as

$$am_{(i,j,k,l)} = B_{(1,1,1,1)} - B_{(2,1,1)} + B_{(2,2)} + 2B_{(3,1)} - 6B_{(4)} \quad (6.9)$$

In the given (lexicographic) order of terms in (6.9), the values of  $y_\kappa$  are  $4!, 4, 8, 6, 4!$  and those of  $\theta_\kappa$  are  $1, 1, 1, 2, 6$  respectively. The number of terms within each bracket is  $4!/y_\kappa$ , and the constants outside are  $\theta_\kappa$ .

The term  $p_{i+j+k}p_l$  in the  $(3,1)$ th bracket, for instance, is correcting for the terms above it of the form  $p_{i+j}p_kp_l$ ,  $p_{i+k}p_jp_l$  and  $p_{j+k}p_ip_l$  in the  $(2,1,1)$ th bracket as well as  $p_ip_jp_kp_l$  in the  $(1,1,1,1)$ th bracket.

Note that the coefficients in (6.8) and (6.9) are consistent with the falling factorial expansion

$$n_{(4)} = n(n-1)(n-2)(n-3) = n^4 - 6n^3 + 11n^2 - 6n. \quad (6.10)$$

The expansion (6.8) reduces to (6.10) when all variate values are unity, and naturally vanishes when  $n < 4$ . In (6.9) there are 6 entries in  $B_{(2,1,1)}$ , 11 entries in  $B_{(2,2)} + 2B_{(3,1)}$ , etc. ■

### Example 6.5

Set  $\lambda = (4, 2, 2, 1)$  in Example 6.4. Then

$$\begin{aligned} am_{\lambda} &= 2m_{(4,2,2,1)} = \sum_{\substack{i,j,k,l \\ d}} y_i^4 y_j^2 y_k^2 y_l = p_{(4,2,2,1)} \\ &- (2p_{(6,2,1)} + p_{(5,2,2)} + p_{(4,4,1)} + 2p_{(4,3,2)}) + (2p_{(6,3)} + p_{(5,4)}) \\ &+ 2(p_{(8,1)} + 2p_{(7,2)} + p_{(5,4)}) - 6p_9 \end{aligned} \quad \blacksquare$$

### 6.3.3 Proof of Theorem 6.1

Returning to the proof of the principal theorem, the logic of the proof is that of the inclusion exclusion principle. The left side of (6.8), viz.  $am_{(i,j,k,l)}$ , for instance, contains monomial elements  $x_1^i x_2^j \dots$  in which no  $x_{\gamma}$  variate repeats, so that there are  $n_{(4)} = n(n-1) \dots (n-3)$  monomial elements. The first term on the right side has elements of the same kind, but the variates can repeat, so that that term contains  $n^4$  monomial elements upon expansion. The succeeding terms on the right side serve to subtract those elements which appear in the first term on the right but not in the left side, and then to compensate for the over-correction etc.

The proof proceeds as if the vector  $(l_1, l_2, \dots, l_r)$  had distinct components  $l_j$ . But the equation (6.3) is an identity involving  $l_1, l_2, \dots$  etc., and is valid whether those

components are distinct or not. The expression (6.8), for instance, is valid for general  $i, j, k$  and  $l$  regardless of whether they happen to coincide or not.

### Proof 6.1

From (6.7), let

$$am_{(l_1, l_2, \dots, l_r)} = \sum_{\kappa \vdash r} (-1)^{\ell(\kappa)+r} \psi_\kappa B_\kappa, \quad (6.11)$$

so that the aim is to show that  $\psi_\kappa = \theta_\kappa$ . The starting point is that  $\psi_\kappa = \theta_\kappa$  when  $w(\kappa) \leq 4$ , as is evident from Examples 6.2, 6.3 and 6.4. Suppose now that  $\psi_\kappa = \theta_\kappa$  for  $w(\kappa) \leq r$ .

For  $1 \leq u \leq r$  write

$$am_{(l_1, l_2, \dots, l_{r+1})} = am_{(l_1, l_2, \dots, l_u)} \times am_{(l_{u+1}, l_{u+2}, \dots, l_{r+1})} + X_u. \quad (6.12)$$

All terms in  $X_u$  have exponents combining at least one element of  $\{l_1, l_2, \dots, l_u\}$  with at least one element of  $\{l_{u+1}, l_{u+2}, \dots, l_{r+1}\}$ .

First set  $u = r$ . The equation (6.12) becomes

$$am_{(l_1, l_2, \dots, l_{r+1})} = am_{(l_1, l_2, \dots, l_r)} \times p_{l_{r+1}} + X_r. \quad (6.13)$$

Regarding  $l_v$  and  $l_w$  as distinct for unequal  $v$  and  $w$ , as discussed in the preamble to this proof, one can state that  $p_{l_{r+1}}$  is not a factor of  $X_r$ .

The equation (6.13) becomes

$$\sum_{\kappa_1 \vdash r+1} (-1)^{\ell(\kappa_1)+r+1} \psi_{\kappa_1} B_{\kappa_1} = \left( \sum_{\kappa \vdash r} (-1)^{\ell(\kappa)+r} \psi_\kappa B_\kappa \right) \times p_{l_{r+1}} + X_r \quad (6.14)$$

Equating coefficients of  $q_{\sigma_\kappa} \times p_{l_{r+1}}$  in (6.14), for  $\kappa \vdash r$  one has

$$\psi_{\kappa \cup (1)} = \psi_\kappa.$$

But  $\psi_\kappa = \theta_\kappa$  by the inductive hypothesis; and  $\theta_{\kappa \cup (1)} = \theta_\kappa$ , so that

$$\psi_{\kappa \cup (1)} = \theta_{\kappa \cup (1)}.$$

The argument generalises immediately. Write (6.12) as

$$am_{(l_1, l_2, \dots, l_{r+1})} = am_{(l_1, l_2, \dots, l_u)} \times [(r-u)! p(l_{u+1} + l_{u+2} + \dots + l_{r+1}) + Y_u] + X_u$$

where  $p(l_{u+1} + l_{u+2} + \dots + l_{r+1})$  is not a factor of  $Y_u$ .

For  $\kappa \vdash u$ , we equate coefficients of  $q_{\sigma_\kappa} \times p(l_{u+1} + l_{u+2} + \dots + l_{r+1})$  to see that

$$\psi_{\kappa \cup (r-u+1)} = (r-u)! \psi_\kappa$$

But  $\psi_\kappa = \theta_\kappa$  by the inductive hypothesis; and  $\theta_{\kappa \cup (r-u+1)} = (r-u)! \theta_\kappa$ , so that

$$\psi_{\kappa \cup (r-u+1)} = \theta_{\kappa \cup (r-u+1)}$$

for all  $u$ ,  $1 \leq u \leq r$ . ■

## 6.4 Behaviour of the $am_\lambda$ functions when the underlying variables form an arithmetic progression

For the remainder of this chapter we assume that  $\lambda = (l_1, l_2, \dots, l_r) \vdash w$ , so that  $l_j \in \mathbb{N}_+$  for each  $j$ . In other words, the set  $J$  in the last section becomes the set of non-negative integers.

### Definitions 6.6

The limiting processes for the arithmetic progressions in §6.4 involve  $n \rightarrow \infty$  with  $a$  and  $d$  constant on the one hand; and  $a$  and  $dn$  constant on the other. In the former case, the parameters arising are  $C_{0,j}$  and  $C_{1,j}$ , which simplify substantially, as shown for instance in Theorem 6.11. The second case sees the analogous emergence of  $D_{0,j}$  and  $D_{1,j}$ , for which there is no obvious simplification.

For  $j \in \mathbb{N}_+$ , and with  $\kappa \vdash r$  and  $t = \ell(\kappa)$ ,

$$C_{0,j} = \frac{d^j}{j+1} \quad \text{and} \quad C_{1,j} = d^{j-1} \left( a - \frac{d}{2} \right)$$

$$D_{0,j} = \sum_{i=0}^j \binom{j}{i} a^{j-i} (nd)^i \frac{1}{i+1} \quad \text{and} \quad D_{1,j} = \frac{1}{2} [(a+nd)^j - a^j]$$

$$\begin{aligned}
 G_{0,\kappa} &= \sum_{s \in S_r} \prod_{j=1}^t C_{0,L_{s\kappa},j} & \text{and} & & G_{1,\kappa} &= \sum_{s \in S_r} \sum_{i=1}^t C_{1,L_{s\kappa},i} \prod_{\substack{j=1 \\ j \neq i}}^t C_{0,L_{s\kappa},j} \\
 H_{0,\kappa} &= \sum_{s \in S_r} \prod_{j=1}^t D_{0,L_{s\kappa},j} & \text{and} & & H_{1,\kappa} &= \sum_{s \in S_r} \sum_{i=1}^t D_{1,L_{s\kappa},i} \prod_{\substack{j=1 \\ j \neq i}}^t D_{0,L_{s\kappa},j}
 \end{aligned}$$

■

**Lemma 6.7**

 For  $j \in \mathbb{N}_+$ ,

$$\sum_{k=0}^{n-1} k^j = \frac{n^{j+1}}{j+1} - \frac{n^j}{2} + O(n^{j-1}) \quad (6.15)$$

The proof is immediate from Abramowitz & Stegun (1964, eqns. 23.1.4 and 7, p. 804). When  $j = 1$ , the final term in (6.15) is redundant. ■

**Lemma 6.8**

We consider two limiting processes: for the first  $a$  and  $d$  are fixed, and  $n \rightarrow \infty$ ; while for the second  $a$  and  $nd$  are fixed, so that  $d \rightarrow 0$  and  $n \rightarrow \infty$ .

**Lemma 6.8(a)**

Let  $a$  and  $d$  be fixed, and  $j \in \mathbb{N}_+$ . Then, as  $n \rightarrow \infty$ ,

$$\frac{p_j}{n^{j+1}} = \frac{d^j}{j+1} + \frac{d^{j-1}}{n} \left[ a - \frac{d}{2} \right] + O(n^{-2}) = C_{0j} + \frac{C_{1j}}{n} + O(n^{-2}).$$

**Lemma 6.8(b)**

Let  $a$  and  $dn$  be fixed, and  $j \in \mathbb{N}_+$ . Then, as  $n \rightarrow \infty$ ,

$$\begin{aligned}
 \frac{p_j}{n} &= \sum_{i=0}^j \binom{j}{i} a^{j-i} (nd)^i \frac{1}{i+1} - \frac{1}{2n} \sum_{i=1}^j \binom{j}{i} a^{j-i} (nd)^i + O(n^{-2}) \\
 &= \sum_{i=0}^j \binom{j}{i} a^{j-i} (nd)^i \frac{1}{i+1} - \frac{1}{2n} [(a+nd)^j - a^j] + O(n^{-2}) \\
 &= D_{0j} + \frac{D_{1j}}{n} + O(n^{-2}).
 \end{aligned}$$

**Proof 6.8**

$$p_j = \sum_{k=0}^{n-1} (a + kd)^j = \sum_{k=0}^{n-1} \sum_{i=0}^j \binom{j}{i} a^{j-i} d^i k^i$$

in which  $0^0 = 1$ . Reversing the order of summation, an application of Lemma 6.7 yields

$$\begin{aligned} p_j &= \sum_{i=0}^j \binom{j}{i} a^{j-i} d^i \sum_{k=0}^{n-1} k^i \\ &= na^j + \frac{1}{2} ja^{j-1} dn(n-1) + \sum_{i=2}^j \binom{j}{i} a^{j-i} d^i \left[ \frac{n^{i+1}}{i+1} - \frac{n^i}{2} + O(n^{i-1}) \right] \\ &= na^j + \frac{1}{2} ja^{j-1} dn(n-1) + \sum_{i=2}^j \binom{j}{i} a^{j-i} (nd)^i \left[ \frac{n}{i+1} - \frac{1}{2} + O(n^{-1}) \right] \quad \blacksquare \end{aligned}$$

**Theorem 6.9**

Let  $\lambda \vdash w$ ,  $\ell(\lambda) = r$ . For  $a$  and  $d$  fixed, and  $n \rightarrow \infty$ ,

$$\frac{am_\lambda}{n^{w+r}} = \frac{1}{r!} G_{0,(1^r)} + \frac{1}{n} \left( \frac{1}{r!} G_{1,(1^r)} - \frac{1}{2!(r-2)!} G_{0,(2,1^{r-2})} \right) + O(n^{-2}) \quad (6.16)$$

For  $a$  and  $dn$  fixed, and  $n \rightarrow \infty$ ,

$$\frac{am_\lambda}{n^r} = \frac{1}{r!} H_{0,(1^r)} + \frac{1}{n} \left( \frac{1}{r!} H_{1,(1^r)} - \frac{1}{2!(r-2)!} H_{0,(2,1^{r-2})} \right) + O(n^{-2}) \quad (6.17)$$

**Proof 6.9**

First, for  $d$  constant and with  $t = \ell(\kappa)$ , one has from (6.5) and Lemma 6.8(a) that

$$\begin{aligned} \frac{q_{\sigma_\kappa}}{n^{w+t}} &= \frac{p(L_{\kappa,1}) p(L_{\kappa,2}) \cdots p(L_{\kappa,t})}{n^{w+t}} \\ &= \prod_{j=1}^t \left\{ C_{0,L_{\kappa,j}} + \frac{C_{1,L_{\kappa,j}}}{n} + \cdots \right\} \\ &= \prod_{j=1}^t C_{0,L_{\kappa,j}} + \frac{1}{n} \sum_{i=1}^t C_{1,L_{\kappa,i}} \prod_{\substack{j=1 \\ j \neq i}}^t C_{0,L_{\kappa,j}} + O(n^{-2}) \end{aligned}$$

From (6.6), one may write

$$\frac{y_\kappa B_\kappa}{n^{w+t}} = \sum_{s \in S_r} \frac{q_{s\sigma_\kappa}}{n^{w+t}}$$

$$\begin{aligned}
 &= \sum_{s \in S_r} \prod_{j=1}^t C_{0, L_{s\kappa, j}} + \frac{1}{n} \sum_{s \in S_r} \sum_{i=1}^t C_{1, L_{s\kappa, i}} \prod_{\substack{j=1 \\ j \neq i}}^t C_{0, L_{s\kappa, j}} + O(n^{-2}) \\
 &= G_{0, \kappa} + \frac{1}{n} G_{1, \kappa} + O(n^{-2})
 \end{aligned}$$

Finally, from (6.11) one has

$$\begin{aligned}
 am_\lambda &= \sum_{\kappa \vdash r} (-1)^{\ell(\kappa)+r} \theta_\kappa B_\kappa \\
 \frac{am_\lambda}{n^{w+r}} &= \sum_{\kappa \vdash r} (-1)^{\ell(\kappa)+r} \frac{\theta_\kappa}{y_\kappa n^{r-t}} \times \frac{y_\kappa B_\kappa}{n^{w+t}} \\
 &= \frac{\theta_{(1^r)}}{y_{(1^r)}} \left( G_{0, (1^r)} + \frac{1}{n} G_{1, (1^r)} \right) - \frac{1}{n} \frac{\theta_{(2, 1^{r-2})}}{y_{(2, 1^{r-2})}} G_{0, (2, 1^{r-2})} + O(n^{-2}) \\
 &= \frac{1}{r!} \left( G_{0, (1^r)} + \frac{1}{n} G_{1, (1^r)} \right) - \frac{1}{2n(r-2)!} G_{0, (2, 1^{r-2})} + O(n^{-2})
 \end{aligned} \tag{6.18}$$

which simplifies to (6.16). The proof of (6.17) is analogous, utilising Lemma 6.8(b).

### Corollary 6.9

Let  $\lambda \vdash w$ . For  $a$  and  $d$  fixed, and as  $n \rightarrow \infty$ ,

$$m_\lambda = O(n^{w+\ell(\lambda)}).$$

For  $a$  and  $nd$  fixed, and as  $n \rightarrow \infty$ , this becomes

$$m_\lambda = O(n^{\ell(\lambda)}).$$

Thus in either case the longer partitions  $\lambda$  have  $m_\lambda$  functions which are of higher order than shorter partitions as  $n \rightarrow \infty$ . ■

### Lemma 6.10

Define

$$\begin{aligned}
 \tau_\kappa &= \frac{1}{y_\kappa} \sum_{s \in S_r} \prod_{j=1}^r [1 + l_{s(\sigma_\kappa(j-1)+1)} + l_{s(\sigma_\kappa(j-1)+2)} + \dots + l_{s(\sigma_\kappa(j))}]^{-1} \\
 &= \frac{1}{y_\kappa} \sum_{s \in S_r} \prod_{j=1}^t (1 + L_{s\kappa, j})^{-1}
 \end{aligned} \tag{6.19}$$

Then

$$G_{0, \kappa} = d^w y_\kappa \tau_\kappa$$

and

$$\frac{G_{1,\kappa}}{G_{0,\kappa}} = \frac{w + \ell(\kappa)}{d} \left( a - \frac{d}{2} \right) \quad (6.20)$$

**Proof 6.10**

First note from Definitions 6.6 that

$$\prod_{j=1}^t C_{0,L_{s\kappa},j} = \prod_{j=1}^t \frac{d^{L_{s\kappa},j}}{L_{s\kappa},j + 1} = \frac{d^w}{\prod_{j=1}^t (L_{s\kappa},j + 1)}$$

and

$$\frac{C_{1,j}}{C_{0,j}} = \frac{j + 1}{d} \left( a - \frac{d}{2} \right)$$

Then

$$G_{0,\kappa} = \sum_{s \in S_r} \prod_{j=1}^t C_{0,L_{s\kappa},j} = \sum_{s \in S_r} \frac{d^w}{\prod_{j=1}^t (L_{s\kappa},j + 1)} = d^w y_\kappa \tau_\kappa$$

and

$$\begin{aligned} G_{1,\kappa} &= \sum_{s \in S_r} \sum_{i=1}^t C_{1,L_{s\kappa},i} \prod_{\substack{j=1 \\ j \neq i}}^t C_{0,L_{s\kappa},j} \\ &= \sum_{s \in S_r} \sum_{i=1}^t \frac{C_{1,L_{s\kappa},i}}{C_{0,L_{s\kappa},i}} \prod_{j=1}^t C_{0,L_{s\kappa},j} \\ &= \sum_{s \in S_r} \sum_{i=1}^t \frac{L_{s\kappa},i + 1}{d} \left( a - \frac{d}{2} \right) \frac{d^w}{\prod_{j=1}^t (L_{s\kappa},j + 1)} \\ &= d^{w-1} (w + t) \left( a - \frac{d}{2} \right) y_\kappa \tau_\kappa \end{aligned}$$

which reduces to (6.20). ■

**Theorem 6.11**

For  $a$  and  $d$  fixed, and as  $n \rightarrow \infty$ ,

$$\frac{am_{(l_1, l_2, \dots, l_r)}}{d^w n^{w+r}} = \tau_{(1^r)} \left[ 1 + \frac{w + r}{dn} \left( a - \frac{d}{2} \right) \right] - \frac{1}{n} \tau_{(2, 1^{r-2})} + O(n^{-2}) \quad (6.21)$$

**Proof 6.11**

The proof follows from (6.18) and Lemma 6.10. ■

**Example 6.12**

For the vector  $(l_1, l_2, l_3)$ ,

$$\begin{aligned}\tau_{(1,1,1)} &= \frac{1}{(l_1 + 1)(l_2 + 1)(l_3 + 1)} \\ \tau_{(2,1)} &= \frac{1}{(l_1 + l_2 + 1)(l_3 + 1)} + \frac{1}{(l_1 + l_3 + 1)(l_2 + 1)} + \frac{1}{(l_2 + l_3 + 1)(l_1 + 1)} \\ \tau_{(3)} &= \frac{1}{(l_1 + l_2 + l_3 + 1)}\end{aligned}$$

As in the definition of  $B_\kappa$  in (6.6), the factor of  $y_\kappa^{-1}$  ensures that there are no duplicates in  $\tau_\kappa$  unless different sums of distinct terms of the vector  $(l_1, l_2, \dots, l_r)$  happen to coincide. Terms arising here naturally mirror those arising in Example 6.3. ■

## 6.5 Behaviour of the $am_\lambda$ functions when the underlying variables form a geometric progression

It is easy to apply Theorem 6.1 when the underlying variates form a geometric progression, since the values of  $p_j$  are readily found for this case.

Let the variables be  $x_j = x^{j-1}$  for  $j = 1, 2, \dots, n$ . Using the standard formula for summing a geometric progression, from (6.3) one has

$$\begin{aligned}am_{(l_1, l_2, \dots, l_r)} &= \sum_{\kappa \vdash r} \frac{(-1)^{\ell(\kappa)+r}}{z_\kappa} \sum_{s \in S_r} \prod_{j=1}^{\ell(\kappa)} p(L_{s\kappa, j}) \\ &= \sum_{\kappa \vdash r} \frac{(-1)^{\ell(\kappa)+r}}{z_\kappa} \sum_{s \in S_r} \prod_{j=1}^{\ell(\kappa)} \frac{1 - x^{nL_{s\kappa, j}}}{1 - x^{L_{s\kappa, j}}}\end{aligned} \tag{6.22}$$

**Example 6.13**

Let the  $n$  variables be  $\{x^j : j = 0, \dots, n-1\}$ . Then, from Example 6.2

$$m_{(6,1)} = p_6 p_1 - p_7 = \left( \frac{1 - x^{6n}}{1 - x^6} \right) \left( \frac{1 - x^n}{1 - x} \right) - \left( \frac{1 - x^{7n}}{1 - x^7} \right) \quad \blacksquare$$

**Theorem 6.14**

Let  $\lambda = (l_1, l_2, \dots, l_r)$  be written in standard non-increasing order, so that in particular  $\ell(\lambda) = r$ . Let the variates be  $\{x^j : j = 0, \dots, n-1\}$ , and set  $u = \sum_{j=1}^r (j-1)l_j$ . Then, as  $x \rightarrow 0$ ,

$$am_\lambda = x^u + O(x^{u+1})$$

**Proof 6.14**

From (6.2) the lowest power arising in the summation will arise when  $x_1 = 1$  is associated with  $l_1$ ,  $x_2 = x$  with  $l_2$ , etc., giving rise to the power  $u$ . ■

### 6.5.1 Infinite number of variates

It is convenient simply to let the number of underlying variates become infinite, for which the numerators of the expressions in (6.22) corresponding to  $p(L_{s\kappa, j})$  reduce to unity. We simply note the following.

**Theorem 6.15**

Let  $\lambda = (l_1, l_2, \dots, l_r) \vdash w$ . Suppose that the variates form an infinite geometric progression  $x^i$ , for  $i = 0, 1, 2, \dots$ , with  $0 < x < 1$ . Then

$$(1-x)^{\ell(\lambda)} am_\lambda \rightarrow \frac{1}{\prod_i l_i} \quad \text{as } x \nearrow 1.$$

**Proof 6.15**

Noting that

$$\frac{1}{1-x^j} \sim \frac{1}{j} \frac{1}{1-x} \quad \text{as } x \nearrow 1,$$

one sees that the dominant bracket in (6.7) on p. 154 is that for the longest partition  $\kappa = (1^r)$ , which contains the single term  $p_{l_1} p_{l_2} p_{l_3} \dots$ . Then

$$\begin{aligned} p_{l_1} p_{l_2} p_{l_3} \dots &= \frac{1}{1-x^{l_1}} \frac{1}{1-x^{l_2}} \frac{1}{1-x^{l_3}} \dots \\ &\sim \frac{1}{l_1} \frac{1}{l_2} \frac{1}{l_3} \dots \left( \frac{1}{1-x} \right)^{\ell(\lambda)} \end{aligned}$$

as  $x \nearrow 1$ . ■

## Chapter 7

# Evaluation of the $e_\lambda$ functions when variates form a geometric progression

All of the  $e_\lambda$  functions remain of the same order as  $n \rightarrow \infty$ , at least when the variates are positive, and bounded above and below away from zero: for then  $e_r = O(n^r)$  as  $n \rightarrow \infty$ , and  $e_\lambda = O(n^w)$ . The vector  $E$  of  $e_\lambda$  functions behaves in the highest order as a constant vector times  $n^w$ . For the  $m_\lambda$  functions, in contrast, there is circumstantial evidence to say that  $m_\lambda$  functions are of higher order for  $\lambda$  lower in the RLO: see §6.1.

The  $e_\lambda$  functions are relatively easy to calculate once the variates are assumed to follow a geometric progression. Less complete results are available otherwise, except for the artificial case in which variates are symmetric about the origin, when the  $e_\lambda$  and  $m_\lambda$  functions simplify. This situation is quickly disposed of in the first section, which is the only part of this chapter in which variates are not assumed to stand in geometric progression. The evaluation of  $e_\lambda$  functions when variates are assumed to follow an arithmetic progression does not seem particularly tractable.

## 7.1 A tangent: when the variates are symmetric about the origin

As a prelude we consider briefly the perhaps artificial case for which the variates are  $\pm x_i$  for  $i = 1, \dots, m$ ; where  $m = n/2$ , with  $n$  even, and  $n$  is as usual the number of variates. In this section the variates are not assumed to stand in geometric progression.

The  $e_\lambda$  functions simplify substantially, and to a lesser extent so do the  $m_\lambda$  functions, as noted in the following result.

### Theorem 7.1

Let  $x_j > 0$  for  $j = 1, 2, \dots, m = n/2$ , with  $n$  an even integer; and with  $\kappa = (k_1, k_2, \dots)$ , let  $2\kappa = (2k_1, 2k_2, \dots)$ . Then

$$e_{2\kappa}(\pm x_1, \pm x_2, \dots, \pm x_m) = e_\kappa(-x_1^2, -x_2^2, \dots, -x_m^2)$$

For partitions  $\lambda$  not of the form  $2\kappa$  for some  $\kappa$ , i.e. for which not all constituent elements are even,

$$e_\lambda(\pm x_1, \pm x_2, \dots, \pm x_m) = 0$$

$$m_\lambda(\pm x_1, \pm x_2, \dots, \pm x_m) = 0$$

### Proof 7.1

For the variates which are symmetric about zero, any monomial  $x_1^k Y$ , where  $Y$  is a monomial in the variates  $x_2, \dots, x_m$  and  $k$  is an odd integer, is offset by  $(-x_1)^k Y$ . The occurrence of  $x_1$  must be balanced by an occurrence of  $-x_1$  to contribute to the  $e_\lambda$  function. ■

## 7.2 Calculating $e_r$ : one can assume that $r \leq n/2$

For variates standing in geometric progression, Theorem 7.3 allows us to assume that  $r \leq n/2$  without loss of generality. That theorem is in turn a direct corollary

of Theorem 7.2, and is illustrated in Example 7.4.

In this section we consider the variates to be  $x^i$  for  $i \in I_1$  where the index set  $I_1 = I$  is a multiset (that is, multiple values are allowed):  $I = \{i_1, i_2, \dots, i_n\}$ , where  $i_j \in \mathbb{R}$  and where without loss of generality we assume that  $i_j \leq i_k$  when  $j < k$ . We set  $I + a = \{i_1 + a, i_2 + a, \dots, i_n + a\}$  for  $a \in \mathbb{R}$ ; and similarly for  $I - a$ ,  $a \pm I$  etc., on the understanding that the order of elements in  $a - I$  is to be reversed.

### Theorem 7.2

For variates  $x^i$  for  $x > 0$  and  $i \in I = \{i_1, i_2, \dots, i_n\}$ , where  $i_j \in \mathbb{R}$  and  $i_j \leq i_k$  when  $j < k$ , let  $I$  satisfy the symmetry condition

$$I - i_1 = i_n - I$$

Further define  $N = i_1 + i_2 + \dots + i_n$ ;  $N_{0r} = i_1 + i_2 + \dots + i_r$  and  $N_{1r} = i_{n-r+1} + i_{n-r+2} + \dots + i_n$ . Then

$$e_{n-r} = x^k e_r$$

where  $k = N - N_{0r} - N_{1r}$ .

### Proof 7.2

Define the elementary symmetric functions for the variates in the statement of the theorem:

$$e_r = e_r^{(1)} = \sum_{\substack{i_1, i_2, \dots, i_r=1 \\ i_1 < i_2 < \dots < i_r}}^n x^{i_1} x^{i_2} \dots x^{i_r} = \sum_{i \in I_r} x^i$$

where the minimum number in  $I_r$  is  $N_{0r}$  and the maximum  $N_{1r}$ .

In view of the equality

$$\{i_{j_1}, i_{j_2}, \dots, i_{j_r}\} - N_{0r} = N_{1r} - \{i_{n+1-j_1}, i_{n+1-j_2}, \dots, i_{n+1-j_r}\}$$

the index set  $I_r$  satisfies the symmetry condition

$$I_r - N_{0r} = N_{1r} - I_r$$

For variates inverse to the above, viz.  $x^{-i}$  for  $i \in I_1$ , define the corresponding

elementary symmetric function

$$e_r^{(2)} = \sum_{i \in -I_r} x^i$$

Now

$$\begin{aligned} e_{n-r}^{(1)} &= x^N e_r^{(2)} = x^N \sum_{i \in I_r} x^{-i} \\ &= x^{N-N_{1r}} \sum_{i \in I_r} x^{N_{1r}-i} = x^{N-N_{1r}} \sum_{i \in N_{1r}-I_r} x^i = x^{N-N_{1r}} \sum_{i \in I_r-N_{0r}} x^i = x^{N-N_{0r}-N_{1r}} e_r^{(1)} \quad \blacksquare \end{aligned} \quad (7.1)$$

### Theorem 7.3

For variates  $x^i$  for  $x > 0$  and  $i \in I = \{1, 2, \dots, n\}$ , one has

$$e_{n-r} = x^k e_r$$

where

$$k = (n+1) \left( \frac{n}{2} - r \right)$$

### Proof 7.3

Apply Theorem 7.2, with

$$N = \frac{n(n+1)}{2}, \quad N_{0r} = \frac{r(r+1)}{2} \quad \text{and} \quad N_{1r} = r(n+1) - N_{0r} \quad \blacksquare$$

### Example 7.4

For  $n = 5$  one has

$$e_3^{(1)} = x^1 x^2 x^3 + \dots + x^3 x^4 x^5 = x^{15} (x^{-4} x^{-5} + \dots + x^{-1} x^{-2}) = x^{15} e_2^{(2)}$$

which corresponds to (7.1); and then

$$\begin{aligned} e_3^{(1)} &= x^{15} (x^{-4} x^{-5} + \dots + x^{-1} x^{-2}) = x^3 (x^{6-4} x^{6-5} + \dots + x^{6-1} x^{6-2}) \\ &= x^3 (x^2 x^1 + \dots + x^5 x^4) = x^3 e_2^{(1)} \end{aligned}$$

The multisets in this case are

$$I_2 = \{3, 4, 5, 5, 6, 6, 7, 7, 8, 9\} \quad \text{and} \quad I_3 = \{6, 7, 8, 8, 9, 9, 10, 10, 11, 12\}$$

### 7.3 The basic recursion for $e_r$

In Theorem 7.8 we obtain a recursion for the  $e_j$  functions when the variates stand in geometric progression. This recursion can be solved to provide an expression for the  $e_j$  functions, given as Theorem 7.12, which turns out to be quite suitable for finding approximations and upper and lower bounds for  $e_j$ . These results provide expressions, approximations and bounds for the  $e_\lambda$  functions, and in turn for the Jack polynomials  $J_\lambda$ .

Denote the variates by  $x^i$  for  $i = 1, 2, \dots, n$ , with  $x > 0$  and  $x \neq 1$ ; and for  $j \in \mathbb{N}_+$  set

$$a_j = \frac{x^j}{x^j - 1} \quad (7.2)$$

so that  $p_j = a_j (x^{jn} - 1)$ . There are in fact standard determinantal expansions for  $p_j$  in terms of  $e_j$  and vice-versa: see e.g. Macdonald (1995, p. 28). While the  $p_j$  functions are easy to calculate when the variates stand in geometric progression, the method we propose is more parsimonious and facilitates finding approximations.

In order to set the scene, we calculate a few of the  $e_j$  functions for small  $j$ .

$$e_1 = \sum_{i=1}^n x^i = a_1 (x^n - 1)$$

$$e_2 = \sum_{\substack{i,j=1 \\ i < j}}^n x^i x^j = \sum_{i=1}^{n-1} \sum_{j=i+1}^n x^i x^j = a_1 \sum_{i=1}^n x^i (x^n - x^i) = a_1 x^n e_1 - a_1 a_2 (x^{2n} - 1)$$

The next example will be treated in some detail, since it illustrates the general approach.

$$\begin{aligned} e_3 &= \sum_{\substack{i,j,k=1 \\ i < j < k}}^n x^i x^j x^k = \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^n x^i x^j x^k \\ &= \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} x^i x^j a_1 (x^n - x^j) = \sum_{i=1}^{n-1} \sum_{j=i+1}^n x^i x^j a_1 (x^n - x^j) \\ &= a_1 x^n e_2 - a_1 \sum_{i=1}^{n-1} \sum_{j=i+1}^n x^i x^{2j} = a_1 x^n e_2 - a_1 \sum_{i=1}^{n-1} x^i a_2 (x^{2n} - x^{2i}) \end{aligned}$$

$$\begin{aligned}
 &= a_1 x^n e_2 - a_1 \sum_{i=1}^n x^i a_2 (x^{2n} - x^{2i}) = a_1 x^n e_2 - a_1 a_2 x^{2n} e_1 + a_1 a_2 \sum_{i=1}^n x^{3i} \\
 &= a_1 x^n e_2 - a_1 a_2 x^{2n} e_1 + a_1 a_2 a_3 (x^{3n} - 1)
 \end{aligned}$$

In similar vein one can prove the following:

$$e_4 = a_1 x^n e_3 - a_1 a_2 x^{2n} e_2 + a_1 a_2 a_3 x^{3n} e_1 - a_1 a_2 a_3 a_4 (x^{4n} - 1)$$

and Theorem 7.8 states that this generalises. We proceed as follows.

### Definition 7.5

For  $1 \leq r \leq n$  and  $1 \leq t$ , let

$$\eta_{rt} = \sum_{i_1=1}^{n-r+1} x^{i_1} \sum_{i_2=i_1+1}^{n-r+2} x^{i_2} \dots \sum_{i_{r-1}=i_{r-2}+1}^{n-1} x^{i_{r-1}} \sum_{i_r=i_{r-1}+1}^n x^{i_r} \quad (7.3)$$

Note that

$$p_t = \eta_{1t} = a_t (x^{nt} - 1) \quad \text{and} \quad \eta_{r1} = e_r \quad (7.4)$$

■

### Lemma 7.6

$$\frac{\eta_{rt}}{a_t} = x^{tn} \eta_{r-1,1} - \eta_{r-1,t+1} = x^{tn} e_{r-1} - \eta_{r-1,t+1}$$

### Proof 7.6

The final summation in (7.3) evaluates as

$$a_t (x^{tn} - x^{ti_{r-1}})$$

whence the result follows upon increasing the ranges of the remaining summations by 1, as in the example for  $e_3$  above Definition 7.5. ■

### Lemma 7.7

For  $i, j \in \mathbb{N}_+$  and  $i < j$ ,

$$a_i - a_j = \frac{a_i a_j}{x^i a_{j-i}}$$

■

### Theorem 7.8

Let  $A_j = \prod_{k=1}^j a_k$ , with  $A_0 \equiv 1$ . Then

$$e_r = \sum_{s=1}^{r-1} (-1)^{s-1} A_s x^{sn} e_{r-s} + (-1)^{r-1} A_r (x^{rn} - 1) \quad (7.5)$$

### Proof 7.8

The proof consists of  $(r-1)$  applications of Lemma 7.6, as well as noting (7.4):

$$e_r = \eta_{r1} = a_1 x^n e_{r-1} - a_1 \eta_{r-1,2} = a_1 x^n e_{r-1} - a_1 a_2 x^{2n} e_{r-2} + a_1 a_2 \eta_{r-2,3} = \dots \quad \blacksquare$$

### Lemma 7.9

For  $0 < t < r$ ,

$$A_{r-t} y_t - A_{r-t} \frac{A_{t-1}}{x^{(t)}_{(2)}} a_r = A_{r-t-1} \frac{A_t}{x^{(t+1)}_{(2)}} a_r \quad (7.6)$$

where

$$y_t = A_t / x^{(t)}_{(2)}.$$

### Proof 7.9

$$\begin{aligned} \text{Left side of (7.6)} &= A_{r-t} \left( \frac{A_t}{x^{(t)}_{(2)}} - \frac{A_{t-1}}{x^{(t)}_{(2)}} a_r \right) = A_{r-t} \frac{A_{t-1}}{x^{(t)}_{(2)}} (a_t - a_r) \\ &= A_{r-t} \frac{A_{t-1}}{x^{(t)}_{(2)}} \frac{a_t a_r}{x^t a_{r-t}} = \text{right side of (7.6)} \end{aligned}$$

where we have used Lemma 7.7. \blacksquare

When  $x > 1$ , it is tempting to replace the bracket  $(x^{rn} - 1)$  in (7.5) by  $x^{rn}$ , partly because the omitted term should be small, at least when  $x$  is not too close to 1; and partly because  $e_j \sim x^{jn}$  as  $n \rightarrow \infty$ , whereupon all terms in (7.5) are rendered of the same order by the change.

To this end we pursue the following series of recursions. First set  $b_s = (-1)^{s-1} A_s$ ;  $c_s = b_s x^{sn}$ ;  $f_0 = 1$ ; and since  $e_1 = a_1 (x^n - 1)$ , also let  $f_1 = a_1 x^n$ . Then, for  $r > 1$ , let

$$f_r = \sum_{s=1}^{r-1} b_s x^{sn} f_{r-s} + b_r x^{rn} = \sum_{s=1}^r c_s f_{r-s}. \quad (7.7)$$

The intention is to have  $f_r$  pick up the highest order terms of  $e_r$  in (7.5).

The symbols  $f_j$  and  $f_\lambda$  in this section should not be confused with the “forgotten symmetric functions”  $f_\lambda$  in Macdonald (1995, p. 22).

### Theorem 7.10

The solution to (7.7) is given by

$$f_r = A_r x^{rn-u} \quad \text{where} \quad u = \binom{r}{2}$$

### Proof 7.10

Setting  $z_j = f_j/x^{jn}$ , one may rewrite (7.7) as

$$z_r = \sum_{s=1}^{r-1} b_s z_{r-s} + b_r \quad \text{with} \quad z_1 = a_1$$

We need to prove that

$$z_r = A_r/x^{\binom{r}{2}} \tag{7.8}$$

for  $r > 1$ .

Now the use of Lemma 7.7 establishes the result for  $r = 2$ . Assume then that

$$z_t = A_t/x^{\binom{t}{2}}$$

for  $t < r$ . Repeated use of Lemma 7.9 establishes the truth of (7.8) for  $t = r$ . ■

Our approach to the principal theorem, Theorem 7.12, is a trifle indirect. For  $e_r$  we consider  $(r + 1)$  levels of functions, each picking up higher order terms than the next. We illustrate for  $e_3$ .

### Example 7.11

We rewrite (7.5) as

$$e_r = c_1 e_{r-1} + c_2 e_{r-2} + \dots + c_{r-1} e_1 + c_r e_0 - b_r \tag{7.9}$$

and taking our cue from (7.7) one may write

$$\begin{aligned} f_3 &= c_1 f_2 + c_2 f_1 + c_3 f_0 & g_3 &= c_1 g_2 + c_2 g_1 & h_3 &= c_1 h_2 & i_3 &= -b_3 \\ f_2 &= c_1 f_1 + c_2 f_0 & g_2 &= c_1 g_1 & h_2 &= -b_2 \\ f_1 &= c_1 f_0 & g_1 &= -b_1 \\ f_0 &= 1 \end{aligned}$$

Then given the last of the following relations, the remainder follow easily.

$$\begin{aligned} e_3 &= c_1 e_2 + c_2 e_1 + c_3 e_0 - b_3 = f_3 + g_3 + h_3 + i_3 = \phi_{03} + \phi_{13} + \phi_{23} + \phi_{33} \\ e_2 &= c_1 e_1 + c_2 e_0 - b_2 = f_2 + g_2 + h_2 = \phi_{02} + \phi_{12} + \phi_{22} \\ e_1 &= c_1 e_0 - b_1 = f_1 + g_1 = \phi_{01} + \phi_{11} \\ e_0 &= 1 = f_0 = \phi_{00} \end{aligned}$$

in which  $f_r = \phi_{0r}$ ,  $g_r = \phi_{1r}$  etc.; and  $\phi_{jr} = O(x^{(r-j)n})$  as  $n \rightarrow \infty$ . The functions  $\phi_{jr}$  are defined formally in the proof of Theorem 7.12. ■

Following the same procedure with a greater “depth” of functions, we have the following result.

### Theorem 7.12

For  $r \leq n$ ,

$$e_r = f_r - a_1 f_{r-1} + a_1 a_2 f_{r-2} - a_1 a_2 a_3 f_{r-3} + \dots = \sum_{t=0}^r (-1)^t A_t f_{r-t} \quad (7.10)$$

in which, when  $j \geq 0$ ,

$$f_j = A_j x^{jn-u} \quad \text{where} \quad u = \binom{j}{2}.$$

### Proof 7.12

Define functions  $\phi_{jr}$  by  $\phi_{jr} = 0$  if  $j > r$ ;  $\phi_{00} = 1$ ; and for  $j = 0, \dots, r-1$ ,

$$\phi_{jr} = c_1 \phi_{j,r-1} + c_2 \phi_{j,r-2} + \dots + c_{r-j} \phi_{jj} = \sum_{k=1}^{r-j} c_k \phi_{j,r-k}$$

$$\phi_{jj} = -b_j \quad \text{for } j = 1, \dots, r$$

Then  $\phi_{jr}$  and  $\phi_{0,r-j}$  satisfy the same recursion, whence

$$\phi_{jr} = k_j \phi_{0,r-j}$$

Comparing the  $\phi_{jj}$  values, one sees that  $k_j = -b_j$ .

For the induction needed to complete the proof, assume

$$e_t = \phi_{0t} + \phi_{1t} + \dots + \phi_{tt} = \sum_{j=0}^t \phi_{jt} \quad (7.11)$$

for  $t < r$ . In fact (7.11) is valid for  $t \leq 3$  from Example 7.11.

$$\begin{aligned} & \phi_{0r} + \phi_{1r} + \dots + \phi_{rr} \\ &= \sum_{k=1}^r c_k \phi_{0,r-k} + \sum_{k=1}^{r-1} c_k \phi_{1,r-k} + \dots + \sum_{k=1}^2 c_k \phi_{r-2,r-k} + \sum_{k=1}^1 c_k \phi_{r-1,r-k} + \phi_{rr} \\ &= c_1 (\phi_{0,r-1} + \phi_{1,r-1} + \dots + \phi_{r-1,r-1}) \\ & \quad + c_2 (\phi_{0,r-2} + \phi_{1,r-2} + \dots + \phi_{r-2,r-2}) + \dots + c_r \phi_{00} + \phi_{rr} \\ &= c_1 e_{r-1} + c_2 e_{r-2} + \dots + c_{r-1} e_1 + c_r e_0 - b_r \end{aligned}$$

so verifying (7.11) for  $t = r$ , in view of (7.9).

We then have

$$e_r = \phi_{0,r} + \phi_{1,r} + \dots = \phi_{0,r} - b_1 \phi_{0,r-1} - \dots = \phi_{0,r} - A_1 \phi_{0,r-1} + A_2 \phi_{0,r-2} - \dots$$

in which

$$\phi_{0,j} = A_j x^{j^{n-u}}$$

from Theorem 7.10, where  $u = \binom{j}{2}$ ;  $\phi_{00} = 1$  and  $\phi_{0k} = 0$  whenever  $k > 0$ . Finally we identify  $\phi_{0j}$  and  $f_j$ . ■

## 7.4 Magnitude of successive terms in the expansion of $e_r$

It is clearly of some interest to investigate the magnitude of successive terms in (7.10).

**Definition 7.13**

Setting  $\zeta_t$  to be the ratio of the magnitude of the  $(t+1)$ th term to that of the  $t$ th term in (7.10), one has

$$\zeta_t = \frac{A_t f_{r-t}}{A_{t-1} f_{r-t+1}} = \frac{a_t x^{r-t}}{a_{r-t+1} x^n} = \frac{x^{r+1} - x^t}{x^{n+1} (x^t - 1)} = \frac{x^{r+1-t}}{x^{n+1}} \frac{1 - x^{-r-1+t}}{1 - x^{-t}} \quad (7.12)$$

for  $0 < t \leq r$ . ■

In particular, since terms in (7.10) alternate in sign and  $\zeta_t$  is non-increasing in  $t$  from Lemma 7.14(a), finer and finer bounds are found for  $e_r$  by truncating further and further after  $\zeta_t$  falls below unity.

**Lemma 7.14(a)**

For  $x, n, r$  constant,  $\zeta_t$  is strictly monotonically decreasing in  $t$ . Symbolically,

$$t \nearrow \Rightarrow \zeta_t \searrow$$

**Lemma 7.14(b)**

For  $x, n, t$  constant,

$$r \nearrow \Rightarrow \zeta_t \nearrow$$
■

Setting  $\gamma = r + 1 - 2t$ , it is readily shown from Definition 7.13 that

$$\zeta_t = \frac{x^{r+1-t}}{x^{n+1}} \frac{r+1-t}{t} \left[ 1 - \frac{\gamma}{2}(x-1) - \frac{\gamma}{12}(-2\gamma-t-3)(x-1)^2 + \gamma \times O((x-1)^3) \right]$$

This expression leads immediately to Lemmas 7.15(a) and 7.16.

**Lemma 7.15(a)**

$$t = \frac{r+1}{2} \Rightarrow \zeta_t = x^{-(2n-r+1)/2} = \frac{x^{r+1-t}}{x^{n+1}}$$

**Lemma 7.15(b)**

When  $x > 1$ ,

$$t < \frac{r+1}{2} \Rightarrow \frac{x^{r+1-t}}{x^{n+1}} \frac{r+1-t}{t} > \zeta_t > \frac{x^{r+1-t}}{x^{n+1}}$$

$$t > \frac{r+1}{2} \Rightarrow \frac{x^{r+1-t}}{x^{n+1}} \frac{r+1-t}{t} < \zeta_t < \frac{x^{r+1-t}}{x^{n+1}}$$

**Proof 7.15(b)**

From (7.12) and Lemma 7.19. ■

**Lemma 7.16**

$$x \searrow 1 \Rightarrow \zeta_t \rightarrow \frac{r+1-t}{t}$$
■

**Lemma 7.17**

When  $x \geq 2$ , then  $\zeta_1 < 1$ .

**Proof 7.17**

$$\zeta_1 = \frac{x^r - 1}{x^n(x-1)} < \frac{x^r}{x^n(x-1)} \leq x^{r-n} \leq 1$$

in which for the last step we have used  $r \leq n$ .

**Corollary 7.17**

If  $x \geq 2$  and  $r \leq n/2$ , then  $\zeta_1 < x^{-n/2}$ . ■

**Lemma 7.18**

$$x > 1 \Rightarrow \frac{\partial}{\partial x} \zeta_t < 0.$$

**Proof 7.18**

$$\zeta_t = \frac{x^{r+1} - x^t}{x^{n+1}(x^t - 1)} = \frac{u}{v},$$

say. Some simplification yields

$$\frac{v^2}{x^{n+t}} \frac{d}{dx} \zeta_t = -(n-r)(x^t - 1)(x^{r+1-t} - 1) - t(x^{r+1} - 1) + (r+1)(x^t - 1)$$

whence the result follows from Lemma 7.19, with  $a = r+1$  and  $b = t$ . ■

**Lemma 7.19**

Let  $\psi(x) = a(x^b - 1) - b(x^a - 1)$ , with  $a > b$  and  $x > 0$ . Then  $\psi(x) \leq 0$ ; and  $\psi(x) = 0$  iff  $x = 1$ .

**Proof 7.19**

$$\frac{d}{dx} \psi(x) = abx^{b-1} (1 - x^{a-b}) \begin{cases} < 0 \text{ when } x > 1 \\ > 0 \text{ when } x < 1 \end{cases} \quad \blacksquare$$

**Lemma 7.20**

Let  $a, b \in \mathbb{N}_+$ , with  $a > b$ ; and set  $m = \left\lfloor \frac{a}{b} \right\rfloor$  and  $c = a - mb$ . Then

$$\frac{x^a - 1}{x^b - 1} = 1 + x^b + x^{2b} + \dots + x^{(m-1)b} + \frac{x^c - 1}{x^b - 1} x^{mb} \quad \blacksquare$$

**Lemma 7.21**

Let  $m = \left\lfloor \frac{r+1-t}{t} \right\rfloor$  and  $c = r + 1 - t - mt$ . Then from Definition 7.13 on p. 175 and Lemma 7.20 one has

$$\zeta_t = \frac{x^t}{x^{n+1}} \left( 1 + x^t + x^{2t} + \dots + x^{(m-1)t} + \frac{x^c - 1}{x^t - 1} x^{mt} \right) \quad (7.13) \quad \blacksquare$$

From (7.13) an obvious way in which to obtain approximations to  $\zeta_t$  is to develop the final term in the bracket. Approximations may for instance be found by using the following inequalities:

$$\frac{x^c - 1}{x^t} < \frac{x^c - 1}{x^t - 1} < \frac{x^c}{x^t} < \frac{c}{t}$$

There are other possibilities. For  $x > 1$ , the basic coefficients  $a_j$  defined in (7.2) enjoy the property that  $a_1 > a_2 > a_3 > \dots$ , with  $a_j \rightarrow 1$  as  $j \rightarrow \infty$ . This may provide further opportunities for approximations, but the point is not pursued further.

## 7.5 Approximate evaluation of $e_r$

The first of the approximations in Proposition 7.22 leads to a convenient means of obtaining good approximations to  $e_r$ , by summing an infinite series which is related to that used to define the  $\theta$  functions (e.g. Whittaker & Watson (1927)). Partial sums of these series are conveniently found by defining recursive functions, so that the same simple computer code can be iterated in a “do” loop to evaluate  $e_r$  functions.

The result achieving this end is Theorem 7.28 on p. 183, based on approximating  $e_r$  by the infinite series (7.28), which ultimately behaves as a series of the form (7.15).

### Proposition 7.22

*Suppose that  $t$  is not too close to 0 or  $r$ , and that  $x$  is not too close to 1. Then*

$$\zeta_t \approx \frac{x^{r+1-t}}{x^{n+1}} \quad (7.14)$$

*If  $t$  is close to  $r$ , and  $x$  is not too close to 1, then*

$$\zeta_t \approx \frac{x^{r+1-t} - 1}{x^{n+1}} \quad \blacksquare$$

Although (7.28) is based on two approximations, viz. the application of (7.14) and then extending the series obtained to infinity, the bracketed term in (7.28) can be evaluated very efficiently to high accuracy. The principal result to this end is Corollary 7.27, but preliminary work is needed before then. The starting point is to simplify notation, transforming the bracket in (7.28) into the function  $g(a, b)$  in (7.15).

### 7.5.1 Definitions

Set

$$g(a, b) = 1 - ab + a^2b^3 - a^3b^6 + a^4b^{10} - a^5b^{15} + \dots \quad (7.15)$$

for  $a, b \in \mathbb{C}$ ; and where  $|a| \leq 1$ ,  $|b| \leq 1$  and  $|ab| < 1$ , so that  $g(a, b)$  converges absolutely. We shall in fact be dealing with the case in which  $a, b \in (0, 1)$ , for which  $|g(a, b)| < 1$ .

The function  $g(a, b)$  can be given in explicit form when  $a = -1$  (Jansen (1975, pp. 81, 89, 492, 494)). While this special case is of little interest here, we shall see that the general expansion (7.15) can be calculated very efficiently in a recursive manner. To this end, define the individual terms in (7.15) by

$$u^{(j)}(a, b) = (-1)^j a^j b^{j(j+1)/2} \quad (7.16)$$

and the following variants:

$$u_i^{(j)}(a, b) = u^{(j)}(ab^i, b) \quad (7.17)$$

as well as partial sums

$$g_i^{(j)}(a, b) = \sum_{k=0}^j u^{(k)}(ab^i, b) = \sum_{k=0}^j u_i^{(k)}(a, b) \quad (7.18)$$

for  $0 \leq i < \infty$ ,  $0 \leq j \leq \infty$ . We have already implicitly used the abbreviation

$$g_0^{(\infty)}(a, b) = g(a, b)$$

and further simplify notation by setting

$$g_i^{(\infty)}(a, b) = g_i(a, b) \quad \text{and} \quad g_0^{(j)}(a, b) = g^{(j)}(a, b)$$

We note that

$$(g_i)_j(a, b) = g_i(ab^j, b) = g(ab^{j+i}, b) = g_{i+j}(a, b) \quad (7.19)$$

It is further convenient to define

$$h_i^{(j)}(a, b, v) = g_i^{(j-1)}(a, b) + u_i^{(j)}(a, b)v \quad (7.20)$$

so that

$$h_i^{(j)}(a, b, 1) = g_i^{(j)}(a, b)$$

For  $g_i^{(j)}(a, b)$ ,  $u_i^{(j)}(a, b)$  and  $h_i^{(j)}(a, b, v)$  we occasionally write  $g_i^{(j)}$ ,  $u_i^{(j)}$  and  $h_i^{(j)}(v)$  respectively; and we further sometimes set

$$h_k^{(l)}(h_i^{(j)}(v)) = h_k^{(l)} h_i^{(j)} v$$

with obvious extensions of this notation to higher levels of recursion.

## 7.5.2 Preliminary work

### Lemma 7.23

$$u_j^{(1)}(a, b)u^{(j)}(a, b) = u^{(j+1)}(a, b)$$

### Proof 7.23

$$\prod_{k=0}^{j-1} u_k^{(1)}(a, b) = \prod_{k=0}^{j-1} u^{(1)}(ab^k, b) = \prod_{k=0}^{j-1} (-1)ab^{k+1} = (-1)^j a^j b^{\sum(k+1)}$$

Now

$$\sum_{k=0}^{j-1} (k+1) = \sum_{k=1}^j k = \frac{j(j+1)}{2}$$

so that from (7.16)

$$\prod_{k=0}^{j-1} u_k^{(1)}(a, b) = (-1)^j a^j b^{j(j+1)/2} = u^{(j)}(a, b)$$

from which one finally has

$$u_j^{(1)}(a, b)u^{(j)}(a, b) = u_j^{(1)}(a, b) \prod_{k=0}^{j-1} u_k^{(1)}(a, b) = \prod_{k=0}^j u_k^{(1)}(a, b) = u^{(j+1)}(a, b) \quad \blacksquare$$

### Lemma 7.24

$$g_i = g_i^{(j-1)} + u_i^{(j)} g_{i+j} = h_i^{(j)}(g_{i+j}) \quad (7.21)$$

where  $i \geq 0$  and  $j$  is any positive integer.

### Proof 7.24

From (7.15) and (7.16) one has

$$\begin{aligned} g(a, b) &= \sum_{j=0}^{\infty} u^{(j)}(a, b) = \sum_{j=0}^{\infty} (-1)^j a^j b^{j(j+1)/2} \\ &= 1 + \sum_{j=1}^{\infty} (-1)^j a^j b^{j(j+1)/2} \end{aligned}$$

$$\begin{aligned}
 &= 1 + \sum_{k=0}^{\infty} (-1)^{k+1} a^{k+1} b^{(k+1)(k+2)/2} \\
 &= 1 - ab \sum_{k=0}^{\infty} (-1)^k (ab)^k b^{k(k+1)/2} \\
 &= 1 - ab \, g(ab, b)
 \end{aligned} \tag{7.22}$$

$$= g^{(0)}(a, b) + u^{(1)}(a, b) \, g_1(a, b) \tag{7.23}$$

As a corollary of (7.23), and bearing (7.19) in mind,

$$g_i(a, b) = g_i^{(0)}(a, b) + u_i^{(1)}(a, b) \, g_{1+i}(a, b) \tag{7.24}$$

Induction on (7.23) proceeds as follows. Consider the equation

$$g = g^{(k-1)} + u^{(k)} \, g_k, \tag{7.25}$$

which we know from (7.23) to be satisfied for  $k = 1$ . Assume that (7.25) is true for  $k \leq j$ .

Then, from Lemma 7.23 and (7.24)

$$\begin{aligned}
 g &= g^{(j-1)} + u^{(j)} g_j = g^{(j-1)} + u^{(j)} \left[ g_j^{(0)} + u_j^{(1)} g_{j+1} \right] \\
 &= g^{(j-1)} + u^{(j)} + u^{(j)} u_j^{(1)} g_{j+1} = g^{(j)} + u^{(j+1)} g_{j+1}
 \end{aligned}$$

Thus (7.25) is satisfied for  $k = j + 1$ , and for all  $k > 0$ . Finally, from (7.19)

$$g_i = g_i^{(j)} + u_i^{(j+1)} g_{i+j+1}$$

from which the first result in (7.21) follows. The second result then follows from (7.20). ■

### Lemma 7.25

$$h_i^{(j)}(v + w) = h_i^{(j)}(v) + u_i^{(j)} w$$

### Proof 7.25

Proof is immediate from (7.20).

**Corollary 7.25**

$$h_{i_1}^{(j_1)} h_{i_2}^{(j_2)} \dots h_{i_\nu}^{(j_\nu)} (v + w) = h_{i_1}^{(j_1)} h_{i_2}^{(j_2)} \dots h_{i_\nu}^{(j_\nu)} v + u_{i_1}^{(j_1)} u_{i_2}^{(j_2)} \dots u_{i_\nu}^{(j_\nu)} w$$

where  $i_k, j_k \in \mathbb{N}_+$  for  $1 \leq k \leq \nu$ . ■

**Lemma 7.26**

For each  $k$ , let  $j_k$  be any positive integer, and set  $i_{k+1} = j_k + i_k$  for  $1 \leq k \leq \nu$ . Then

$$u_{i_1}^{(j_1)} u_{i_2}^{(j_2)} \dots u_{i_\nu}^{(j_\nu)} = (-1)^A a^A b^B$$

where

$$A = \sum_{k=1}^{\nu} j_k \quad \text{and} \quad B = A \left( i_1 + \frac{1}{2} \right) + \sum_{\substack{k,l=1 \\ k < l}}^{\nu} j_k j_l + \frac{1}{2} \sum_{k=1}^{\nu} j_k^2 \quad (7.26)$$

**Proof 7.26**

From (7.16) and (7.17) we have

$$u_{i_1}^{(j_1)} u_{i_2}^{(j_2)} \dots u_{i_\nu}^{(j_\nu)} = (-1)^A a^A b^C$$

where

$$A = \sum_{k=1}^{\nu} j_k \quad C = \sum_{k=1}^{\nu} j_k [2i_k + j_k + 1] / 2$$

Now impose the conditions that  $i_{k+1} = j_k + i_k$ ; i.e.  $i_{k+1} = i_1 + \sum_{l=1}^k j_l$ . We find that  $C = B$ . ■

### 7.5.3 Principal results

**Theorem 7.27**

With notation defined in §7.5.1,

$$g_{i_1} = h_{i_1}^{(j_1)} h_{i_2}^{(j_2)} \dots h_{i_{\nu-1}}^{(j_{\nu-1})} g_{i_\nu}$$

where for each  $k$ ,  $j_k$  is any positive integer, and  $i_{k+1} = j_k + i_k$  for  $1 \leq k \leq \nu - 1$ .

### Corollary 7.27

With the same notation as in the theorem,

$$g_{i_1} = h_{i_1}^{(j_1)} h_{i_2}^{(j_2)} \dots h_{i_{\nu-1}}^{(j_{\nu-1})} g_{i_{\nu}}^{(j_{\nu}-1)} + (-1)^A a^A b^B g_{i_{\nu}+j_{\nu}}$$

where  $j_{\nu}$  is any positive integer;  $A$  and  $B$  are given in (7.26); and  $0 < g_{i_{\nu}+j_{\nu}} < 1$  when  $0 < a < 1$  and  $0 < b < 1$ .

### Proof 7.27

The theorem follows from repeated application of Lemma 7.24. For the corollary, write from Lemma 7.24:

$$g_{i_{\nu}} = g_{i_{\nu}}^{(j_{\nu}-1)} + u_{i_{\nu}}^{(j_{\nu})} g_{i_{\nu}+j_{\nu}}$$

so that the theorem implies

$$g_{i_1} = h_{i_1}^{(j_1)} h_{i_2}^{(j_2)} \dots h_{i_{\nu-1}}^{(j_{\nu-1})} \left( g_{i_{\nu}}^{(j_{\nu}-1)} + u_{i_{\nu}}^{(j_{\nu})} g_{i_{\nu}+j_{\nu}} \right)$$

We now apply Corollary 7.25 to write

$$g_{i_1} = h_{i_1}^{(j_1)} h_{i_2}^{(j_2)} \dots h_{i_{\nu-1}}^{(j_{\nu-1})} g_{i_{\nu}}^{(j_{\nu}-1)} + u_{i_1}^{(j_1)} u_{i_2}^{(j_2)} \dots u_{i_{\nu-1}}^{(j_{\nu-1})} u_{i_{\nu}}^{(j_{\nu})} g_{i_{\nu}+j_{\nu}}$$

and finally apply Lemma 7.26 to obtain

$$g_{i_1} = h_{i_1}^{(j_1)} h_{i_2}^{(j_2)} \dots h_{i_{\nu-1}}^{(j_{\nu-1})} g_{i_{\nu}}^{(j_{\nu}-1)} + (-1)^A a^A b^B g_{i_{\nu}+j_{\nu}} \quad \blacksquare$$

### Theorem 7.28

Let  $Y = (-1)^{t_0-1} A_{t_0-1} f_{r-t_0+1}$  and  $Z = \zeta_{t_0} \leq 1$ . Assuming the approximation (7.14) on p. 178 to be applicable for  $t \geq t_0$ , then

$$e_r \approx \sum_{t=0}^{t_0-2} (-1)^t A_t f_{r-t} + Y$$

$$-YZ \left[ h_{i_1}^{(j_1)} \left( h_{i_2}^{(j_2)} \left( \dots \left( h_{i_{\nu-1}}^{(j_{\nu-1})} \left( g_{i_{\nu}}^{(j_{\nu}-1)} \right) \dots \right) \right) \right) + (-1)^A Z^A x^{-B} \xi \right]$$

for arbitrary positive integers  $\nu$ , and  $j_k$  for  $k = 1, \dots, \nu - 1$ ; where  $i_{k+1} = j_k + i_k$ , with  $i_1 = 0$ ; and where  $0 < \xi < 1$ . The functions  $g_i^{(j)}$  and  $h_i^{(j)}$  are defined in (7.18) and (7.20) respectively; and  $A$  and  $B$  are defined in (7.26).

### Note 7.28

In practice the integers  $j_k$  would be set to a low common value, and the resulting

small amount of computer coding would be applied recursively to obtain any desired level of accuracy.

**Proof 7.28**

From (7.10) one has

$$\begin{aligned}
 e_r &= \sum_{t=0}^{t_0-2} (-1)^t A_t f_{r-t} + (-1)^{t_0-1} A_{t_0-1} f_{r-t_0+1} + \dots + (-1)^r A_r \\
 &\approx \sum_{t=0}^{t_0-2} (-1)^t A_t f_{r-t} + Y - YZ + YZ^2 \frac{1}{x} - YZ^3 \frac{1}{x^3} + YZ^4 \frac{1}{x^6} - \dots \\
 &\quad + (-1)^{r-t_0+1} Y Z^{r-t_0+1} x^{-(r-t_0+1)(r-t_0)/2}
 \end{aligned} \tag{7.27}$$

where we have applied (7.14). Further approximating by taking the series in (7.27) to infinity, we write

$$\begin{aligned}
 e_r &\approx \sum_{t=0}^{t_0-2} (-1)^t A_t f_{r-t} + Y - YZ \left( 1 - \frac{Z}{x} + \frac{Z^2}{x^3} - \frac{Z^3}{x^6} + \dots \right) \\
 &= \sum_{t=0}^{t_0-2} (-1)^t A_t f_{r-t} + Y - YZ g \left( Z, \frac{1}{x} \right) \\
 &= \sum_{t=0}^{t_0-2} (-1)^t A_t f_{r-t} + Y \\
 &\quad - YZ \left[ h_{i_1}^{(j_1)} \left( h_{i_2}^{(j_2)} \left( \dots \left( h_{i_{\nu-1}}^{(j_{\nu-1})} \left( g_{i_{\nu}}^{(j_{\nu}-1)} \right) \dots \right) \right) \right) + (-1)^A Z^A x^{-B} g_{i_{\nu}+j_{\nu}} \right]
 \end{aligned} \tag{7.28}$$

from Corollary 7.27. Noting that  $|g(a, b)| < 1$  when  $0 < a \leq 1$  and  $0 < b < 1$  completes the proof. ■

The first approximate term in (7.27) is that containing  $Z^2$ . Although Proposition 7.22 does not apply when  $t$  is close to its upper limit of  $r$ , the inaccuracy in the approximation in the final terms of (7.27) is unlikely to be material in the overall result.

## 7.6 Calculating $e_\lambda$ : the main theorem

### Theorem 7.29

Let  $\lambda = (r, s, t, \dots)$  and  $A_j = \prod_{k=1}^j a_k$ , with  $A_0 \equiv 1$ . Then

$$e_\lambda = \sum_{m=0}^{\infty} (-1)^m \sum_{\substack{i,j,k,\dots=0 \\ i+j+k+\dots=m}}^m A_i A_j A_k \dots f_{(r-i, s-j, t-k, \dots)}$$

in which  $f_\lambda = f_r f_s f_t \dots$ ; and, when  $q > 0$ ,

$$f_q = A_q x^{qn-u} \quad \text{where} \quad u = \binom{q}{2};$$

$f_0 = 1$ ; and  $f_q \equiv 0$  when  $q < 0$ .

### Proof 7.29

$$e_\lambda = e_r e_s e_t \dots = \left( \sum_{i=0}^{\infty} (-1)^i A_i f_{r-i} \right) \left( \sum_{j=0}^{\infty} (-1)^j A_j f_{s-j} \right) \left( \sum_{k=0}^{\infty} (-1)^k A_k f_{t-k} \right) \dots$$

from Theorem 7.12. ■

### Example 7.30

We illustrate Theorem 7.29 when  $\ell(\lambda) = 3$ , say  $\lambda = (r, s, t)$ . Then since  $e_\lambda = e_r e_s e_t$  (because of which property Macdonald (1995, p. 306) calls the  $e_\lambda$  functions a multiplicative family of polynomials), it is natural to define  $f_\lambda = f_r f_s f_t$ .

$$\begin{aligned} e_\lambda &= e_{(r,s,t)} = e_r e_s e_t \\ &= (f_r - a_1 f_{r-1} + a_1 a_2 f_{r-2} - \dots) (f_s - a_1 f_{s-1} + a_1 a_2 f_{s-2} - \dots) \\ &\quad \times (f_t - a_1 f_{t-1} + a_1 a_2 f_{t-2} - \dots) \end{aligned} \tag{7.29}$$

$$\begin{aligned} &= f_\lambda - a_1 (f_{r-1} f_s f_t + f_r f_{s-1} f_t + f_r f_s f_{t-1}) \\ &\quad + a_1^2 (f_{r-1} f_{s-1} f_t + f_{r-1} f_s f_{t-1} + f_r f_{s-1} f_{t-1}) \\ &\quad + a_1 a_2 (f_{r-2} f_s f_t + f_r f_{s-2} f_t + f_r f_s f_{t-2}) + \dots \end{aligned} \tag{7.30}$$

■

## 7.7 Approximations to $e_\lambda$

One can truncate at the level of each individual  $e_r$  function, i.e. using (7.29), or at the composite level  $e_\lambda$  using (7.30) or Theorem 7.29.

When  $\zeta_t < 1$ , Lemma 7.14 on p. 175 shows that terms succeeding the  $t$ th in (7.14) change sign and decrease in magnitude, so that bounds are easily obtained for  $e_r$  by truncating after the  $t$ th term. Obtaining bounds for the  $e_\lambda$  functions seems more difficult.

Since the  $e_\lambda$  functions are a multiplicative family of polynomials, one should presumably operate approximations at the level of the individual  $e_r$  functions. Upon multiplying the  $e_r$  functions to get the  $e_\lambda$  functions, one could attempt to balance the numbers of positive and negative errors; one would watch for accumulating errors; and one would seek to balance the accuracy required in approximating an  $e_\lambda$  function with its importance in the overall determinantal expansion of the Jack polynomial.

## Chapter 8

### Conclusion

In some respects this thesis has been a little disappointing for the author. It is frustrating to set something up and not get it past the drawing board, in the sense that no calculations of Jack polynomials have yet been effected by the methods advocated. Not even that: there have been no experiments carried out to evaluate how well the arithmetic and geometric progression approximations to the variates translate into the variates which arise in practice; nor have the methods suggested for calculation of  $e_\lambda$  and  $m_\lambda$  functions when variates assume a geometric progression been tried out. We have been rather vague in stating how easy it will be to shorten the process of using the determinantal expansions, although it is clear that there will be shortcuts available. The thesis title may have been misleading, as stressed at the beginning of the thesis; but in its main premise it is accurate, in that one is only groping towards the calculation of Jack polynomials. That step is only a means to an end, viz. the applications of Jack polynomials to practical problems, whether in statistics or elsewhere.

Nevertheless, the disappointments are only skin deep. With the determinantal expansions in Chapter 3 one is close now to calculating Jack polynomials, or at least to obtaining workable approximations to them for practical purposes. The dictates of time meant that I could not carry out the later stages of the programme; but those developments are tantalisingly close.

More importantly, however, I have enjoyed working in this area. My enthusiasm for the general areas of invariant polynomials, group representation theory, Lie groups and algebras and other mathematical areas on the one hand; and the applications of factorisation of measures and differential geometry to statistical theory on the other, is undimmed.

## 8.1 A wishlist

At the end of any enterprise comes the inevitable wishlist of what might have been.

1. I should like to have drawn more heavily on combinatorial mathematics. The algorithm for calculating Jack polynomials in Knop & Sahi (1997) seems at first glance to be a prime candidate for the application of Polya's theory of counting, as outlined in say de Bruijn (1963) (see de Bruijn (1964) or Riordan (1964) for more elementary accounts); but the problem at hand seemed not quite to fit into that framework. On the other hand, the principle of inclusion exclusion played a vital part in Chapter 6, and there is hardly any more basic theory in combinatorics.

If one could apply combinatorial methodology to the Knop-Sahi algorithm, it would be likely to apply equally well to the asymmetric Jack polynomials as to the symmetric Jack polynomials.

2. James (1968) drew some of his inspiration from Helgason (1962), and it might have been desirable to have placed the thesis more firmly in the framework of Lie groups and algebras, and symmetric spaces. Hannan (1965a, p. 50) also summarises the nub of Helgason's material on symmetric spaces.
3. As far as practical, it would seem worthwhile treating the asymmetric polynomials and the symmetric polynomials simultaneously.

At the conceptual end of things, the asymmetric polynomials are hardly any more difficult to define than the symmetric polynomials, and they can give rise to more general models. The problem of calculation for applications may or may not be more severe than for the symmetric polynomials: that is untested.

According to a recent private communication from Peter Forrester, future research by physicists in this area may well emphasise the asymmetric polynomials, deriving the symmetric polynomials by symmetrisation as needed. This emphasis is certainly reflected in recent papers of Forrester, most saliently perhaps in Baker & Forrester (1999).

## 8.2 Relative merits of expanding Jack polynomials in the $e_\lambda$ and $m_\lambda$ functions

Expansion of Jack polynomials in terms of  $m_\lambda$  functions vis-a-vis that in the  $e_\lambda$  function provides a natural counterbalance. There is firstly the verification of theorems and calculations from consistency of results using two distinct but related methodologies; in fact the methodologies are sufficiently similar that many algorithms could apply to expansion in either  $e_\lambda$  or  $m_\lambda$  functions.

But the complementarity of these two approaches lies more deeply than duplication of calculations. The expansion of Jack polynomials in  $e_\lambda$  functions produces coefficients of both sign, potentially allowing cancelling so that fewer terms need be used; while expansion in  $m_\lambda$  functions produces coefficients of the one sign, potentially facilitating finding lower or upper bounds if one can find dominant terms in the expansion. One or the other basis function may be more convenient to calculate under different scenarios: in particular, when the variates are eigenvalues of a matrix, Takemura (1984, §4.5) stresses that the  $e_\lambda$  functions are more readily computed than the  $m_\lambda$  functions, because they are calculated from the principal minors.

Both  $e_\lambda$  and  $m_\lambda$  functions may be convenient to work out when the variates form a geometric progression, to judge from the results in Chapters 6 and 7.

An advantage of the  $e_\lambda$  functions is that they belong to what Macdonald (1995, p. 306) calls the multiplicative family of polynomials, in that one can multiply together the  $e_{(j)} = e_j$  components to provide the composite function:  $e_\lambda = e_{l_1} e_{l_2} \dots$  when  $\lambda = (l_1, l_2, \dots)$ . No similar advantage accrues to the  $m_\lambda$  functions. There would seem

to be real advantages in this property, in that one can first calculate the component functions  $e_j$  to any level of approximation desirable, then multiply the components together in many different ways to obtain the  $e_\lambda$  functions: there are after all many more partitions  $\lambda$  than there are distinct component  $e_j$  functions.

On the other hand, theoretical developments of Jack polynomials have generally utilised the power symmetric functions  $p_\lambda$ , Schur functions  $s_\lambda$  or the  $m_\lambda$  functions, and it is the last named which has been used recently in definitions of Jack polynomials (Macdonald (1995, ), Stanley (1989, p. 77) etc.); and from the evidence in Chapter 4, expansion of Jack polynomials as determinants in  $m_\lambda$  functions involves fewer terms than the analogous expansion in  $e_\lambda$  functions.

Decisions as to the preference of one expansion over the other have to await numerical evidence, and will in any case depend on the circumstances of the particular modelling situation.

### **8.3 On the assumptions that variates form arithmetic or geometric progressions**

It is clearly convenient to make the assumption that the variates form a geometric progression, since both  $e_\lambda$  and  $m_\lambda$  functions are then easily calculated, according to results in Chapters 6 and 7. If one assumes that the variates are approximately a geometric progression, then fitting a regression line to the logarithms will produce a good fit (assuming the variates to be positive, as they often will be in statistical applications).

The distortion induced by making global assumptions that variates form an arithmetic progression or geometric progression is untested. In statistical hypothesis testing, moreover, one may be particularly interested in outliers, for which the assumption that the variates assume the form of an arithmetic progression or geometric progression may be highly questionable.

Insofar as the geometric progression approximation to the variates is globally suitable, the calculation of  $J_\lambda$  effectively becomes one dimensional, apart from a scaling factor. Estimation of Jack polynomials can then be effected by interpolation over that one parameter.

## 8.4 A beginning, not the end

Even if zonal polynomials have “gone off the boil” as claimed in Chapter 1, and even if statisticians have circumvented the calculation of zonal polynomials through the use of asymptotic expansions of hypergeometric functions, there is still interest in calculating zonal and Jack polynomials by statisticians, combinatorialists, electronic engineers and physicists, among others. If this thesis prompts a reawakening of the latent interest of mathematicians in these polynomials, I shall be well pleased.

Whatever the global impact of this thesis, however, the distribution of which must have an atom at zero with large probability, I am glad to have finished it. One reason for this is that I am now free to continue research in this area. I have learnt much, and it is again time to balance reading the literature with developing and applying ideas which have occurred to me but have not yet been developed.

The other reasons for which I am glad to finish are perhaps more obvious.

# Bibliography

- Abramowitz, M. and Stegun, I. E. (1964). *Handbook of Mathematical Functions*, Dover.
- Aigner, M. (1979). *Combinatorial Theory*, Springer Verlag.
- Akin, K., Buchsbaum, D. A. and Weyman, J. (1982). Schur functors and schur complexes, *Advances in Mathematics* **44**: 207–278.
- Anderson, G. A. (1965). An asymptotic expansion for the distribution of the latent roots of the estimated covariance matrix, *Annals of Mathematical Statistics* **36**(4): 1153–1173.
- Andrews, G. E. (1998). *The Theory of Partitions*, CUP. Updated reprinting of 1984 book of the same name.
- Antosik, P., Mikusinski, J. and Sikorski, R. (1973). *Theory of distributions: the sequential approach*, Elsevier.
- Apostol, T. M. (1976). *Introduction to Analytical Number Theory*, Springer Verlag.
- Atiyah, M. F. and Macdonald, I. G. (1969). *Introduction to Commutative Algebra*, Addison Wesley.
- Baker, T. H. and Forrester, P. J. (1997). The Calogero Sutherland model and generalized classical polynomials, *Communications in Mathematical Physics* **188**: 175–216.
- Baker, T. H. and Forrester, P. J. (1998). Nonsymmetric Jack polynomials and integral kernels, *Duke Mathematical Journal* **95**(1): 1–50.

- Baker, T. H. and Forrester, P. J. (1999). Symmetric Jack polynomials from non-symmetric theory, *Annals of Combinatorics* **3**: 159–170.
- Bhanu-Murti, T. S. (1960). Plancherel's measure for the factor space  $SL(n, r)/SO(n, r)$ , *Soviet Math. Doklady* **1**: 860–862.
- Bingham, C. (1974). An antipodally symmetric distribution on the sphere, *Annals of Statistics* **2**(6): 1201–1225.
- Bingham, C., Chang, T. and Richards, D. (1992). Approximating the matrix Fisher and Bingham distributions: applications to spherical regression and Procrustes analysis, *Journal of Multivariate Analysis* **41**: 314–337.
- Blokhuis, A. and Seidel, J. J. (1984). An introduction to multilinear algebra and some applications, *Philips Journal of Research* **39**(4/5): 111–120.
- Brylawski, T. (1973). The lattice of integer partitions, *Discrete mathematics* **6**: 201–219.
- Burrow, M. (1965). *Representation Theory of Finite Groups*, Academic Press.
- Chikuse, Y. (1990a). Distributions of orientations on Stiefel manifolds, *Journal of Multivariate Analysis* **33**: 247–264.
- Chikuse, Y. (1990b). The matrix angular central Gaussian distribution, *Journal of Multivariate Analysis* **33**: 265–274.
- Chikuse, Y. (1991a). Asymptotic expansions for distributions of the large sample matrix resultant and related statistics on the Stiefel manifold, *Journal of Multivariate Analysis* **39**(2): 270–283.
- Chikuse, Y. (1991b). High dimensional limit theorems and matrix decompositions on the Stiefel manifold, *Journal of Multivariate Analysis* **36**(2): 145–162.
- Chikuse, Y. (1993). High dimensional asymptotic expansions for the matrix Langevin distributions on the Stiefel manifold, *Journal of Multivariate Analysis* **44**: 82–101.

- Chikuse, Y. (1994). Invariant measures on Stiefel manifolds with applications to multivariate analysis, in T. W. Anderson, K. T. Fang and I. Olkin (eds), *Multivariate Analysis and its Applications*, Vol. 24 of *Lecture Notes - Monograph series*, IMS, pp. 177-194.
- Chikuse, Y. and Davis, A. W. (1986). A survey on the invariant polynomials with matrix arguments in relation to econometric distribution theory, *Econometric Theory* **2**: 232-248.
- Cohn, P. M. (1981). *Algebra*, Vol. 1, 2nd. edn, Wiley.
- Cohn, P. M. (1989). *Algebra*, Vol. 2, 2nd edn, Wiley.
- Constantine, A. G. (1963). Some non-central distribution problems in multivariate analysis, *Annals of Mathematical Statistics* **34**: 1270-1285.
- Curtis, C. W. and Reiner, I. (1962). *Representation Theory of Finite Groups and Associative Algebras*, Wiley.
- Davis, A. W. (1976). Statistical distributions in univariate and multivariate Edgeworth populations, *Biometrika* **63**: 661-670.
- Davis, A. W. (1979). Invariant polynomials with two matrix arguments extending the zonal polynomials: applications to multivariate distribution theory, *Journal of the Institute of Statistical Mathematics* **31A**: 465-485.
- Davis, A. W. (1980). Invariant polynomials with two matrix arguments, extending the zonal polynomials, in P. R. Krishnaiah (ed.), *Multivariate Analysis - V*, North Holland, pp. 287-299.
- Davis, A. W. (1981). On the construction of a class of invariant polynomials in several matrices, extending the zonal polynomials, *Annals of the Institute of Statistical Mathematics* **33**: 297-313.
- de Bruijn, N. G. (1963). Enumerative combinatorial problems concerning structures, *Nieuw Archief voor Wiskunde* **XI**(3): 142-161.
- de Bruijn, N. G. (1964). Polya's theory of counting, in E. F. Beckenbach (ed.), *Applied Combinatorial Mathematics*, Wiley, chapter 5, pp. 144-184.
- Downs, T. D. (1972). Orientation statistics, *Biometrika* **59**(3): 665-676.



- Grötschel, M. and Lovasz, L. (1995). Combinatorial optimization, in R. L. Graham, M. Grötschel and L. Lovasz (eds), *Handbook of Combinatorics*, Vol. 2, Elsevier/MIT Press, chapter 28, pp. 1541–1597.
- Hall, M. (1959). *The Theory of Groups*, 1st edn, Chelsea.
- Hannan, E. J. (1965a). Group representations and applied probability, *Journal of Applied Probability* **2**: 1–68.
- Hannan, E. J. (1965b). *Group Representations and Applied Probability*, Methuen. Monograph in Methuen's Review Series in Applied Probability, reprint of Hannan (1965a).
- Helgason, S. (1962). *Differential Geometry and Symmetric Spaces*, Academic Press.
- Herz, C. S. (1955). Bessel functions of a matrix argument, *Annals of Mathematics* **61**: 474–523.
- Hill, G. W. and Davis, A. W. (1968). Generalized asymptotic expansions of Cornish-Fisher type, *Annals of Mathematical Statistics* **39**(4): 1264–1273.
- Hillier, G. H. and Skeels, C. L. (1996). ET Interview: A T James, *Econometric Theory* **12**(1): 155–185.
- Hua, L. K. (1963). *Harmonic Analysis of Functions of Several Complex Variables in the Classical Domains*, Vol. 6 of *Translations of Mathematical Monographs*, American Mathematical Society.
- James, A. T. (1954). Normal multivariate analysis and the orthogonal group, *Annals of Mathematical Statistics* **25**: 40–75.
- James, A. T. (1955a). A generating function for averages over the orthogonal group, *Proceedings of the Royal Society (London), Series A* **229**: 367–375.
- James, A. T. (1955b). The non-central Wishart distribution, *Proceedings of the Royal Society (London), Series A* **229**: 364–366.
- James, A. T. (1957). The relationship algebra of an experimental design, *Annals of Mathematical Statistics* **28**: 993–1002.

- James, A. T. (1960). The distribution of the latent roots of the covariance matrix, *Annals of Mathematical Statistics* **31**: 151–158.
- James, A. T. (1961). Zonal polynomials of the real positive definite symmetric matrices, *Annals of Mathematics* **74**(3): 456–469.
- James, A. T. (1964). Distributions of matrix variables and latent roots derived from normal samples, *Annals of Mathematical Statistics* **35**(1): 475–501.
- James, A. T. (1968). Calculation of zonal polynomial coefficients by use of the Laplace Beltrami operator, *Annals of Mathematical Statistics* **39**: 1711–1718.
- James, A. T. (1982). Analyses of variance determined by symmetry and combinatorial properties of zonal polynomials, in G. Kallianpur, P. R. Krishnaiah and J. K. Ghosh (eds), *Statistics and Probability: Essays in honour of C R Rao*, North-Holland, pp. 329–341.
- Jansen, E. R. (1975). *A Table of Series and Products*, Prentice-Hall.
- Johnson, N. L., Kotz, S. and Balakrishnan, N. (1997). *Discrete Multivariate Distributions*, Wiley.
- Kent, J. T. (1987). Asymptotic expansions for the Bingham distribution, *Applied Statistics* **36**(2): 139–144.
- Khatri, C. G. (1966). On certain distribution problems based on positive definite quadratic functions in normal vectors, *Annals of Mathematical Statistics* **37**: 468–479.
- Khatri, C. G. and Mardia, K. V. (1977). The von-Mises Fisher matrix distribution in orientation statistics, *Jnl. of the Royal Statistical Society, Series B* **39**(1): 95–106.
- Knop, F. and Sahi, S. (1997). A recursion and a combinatorial formula for Jack polynomials, *Inventiones Mathematicae* **128**: 9–22.
- Lapointe, L., Lascoux, A. and Morse, J. (2000). Determinantal expression and recursion for Jack polynomials, *Electronic Journal of Combinatorics* **7**(1). #N1.

- Lapointe, L. and Vinet, L. (1995). A Rodrigues formula for the Jack polynomials and the Macdonald-Stanley conjecture, *International Mathematical Research Notices* **9**: 419–424.
- Lapointe, L. and Vinet, L. (1996). Exact operator solution of the Calogero-Sutherland model, *Communications in Mathematical Physics* **178**: 425–452.
- Lapointe, L. and Vinet, L. (1997a). Creation operators for the Macdonald and Jack polynomials, *Letters in Mathematical Physics* **40**: 269–286.
- Lapointe, L. and Vinet, L. (1997b). Rodrigues formulas for the Macdonald polynomials, *Advances in Mathematics* **130**: 261–279.
- Ledermann, W. (1977). *Introduction to Group Characters*, CUP.
- Littlewood, D. E. (1950). *The Theory of Group Characters and Matrix Representations of Groups*, OUP.
- Macdonald, I. D. (1968). *The Theory of Groups*, OUP.
- Macdonald, I. G. (1979). *Symmetric Functions and Hall Polynomials*, Clarendon Press.
- Macdonald, I. G. (1995). *Symmetric Functions and Hall Polynomials*, 2nd edn, Clarendon Press.
- Mann, H. B. (1960). The algebra of a linear hypothesis, *Annals of Mathematical Statistics* **31**: 1–15.
- Mardia, K. V., Kent, J. T. and Bibby, J. M. (1979). *Multivariate Analysis*, Academic Press.
- Mathai, A. M., Provost, S. B. and Hayakawa, T. (1995). *Bilinear Forms and Zonal Polynomials*, Springer Verlag.
- Mathai, A. M. and Saxena, R. K. (1973). *Generalised Hypergeometric Functions with Applications in Statistics and Physical Sciences*, Vol. 348 of *Lecture Notes in Mathematics*, Springer Verlag.
- McLaren, M. L. (1976). Coefficients of the zonal polynomials, *Applied Statistics* **25**: 82–87.

- Moran, P. A. P. (1968). *An Introduction to Probability Theory*, Clarendon Press.
- Muirhead, R. J. (1970). Systems of partial differential equations for hypergeometric functions of matrix argument, *Annals of Mathematical Statistics* **41**: 991–1001.
- Muirhead, R. J. (1975). Expressions for some hypergeometric functions of matrix argument with applications, *Journal of Multivariate Analysis* **5**: 283–293.
- Muirhead, R. J. (1978). Latent roots and matrix variates: A review of some asymptotic results, *Annals of Statistics* **6**: 5–33.
- Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*, Wiley.
- Opdam, E. M. (1995). Harmonic analysis for certain representations of graded Hecke algebras, *Acta Mathematica* **175**: 75–121.
- Read, R. C. (1968). The use of s-functions in combinatorial analysis, *Canadian Journal of Mathematics* **20**: 808–841.
- Riordan, J. (1964). Generating functions, in E. F. Beckenbach (ed.), *Applied Combinatorial Mathematics*, Wiley, chapter 3, pp. 67–95.
- Roberts, L. A. (1998). On the expansion of zonal polynomials in monomial symmetric functions, *Technical Report 98-31*, School of Mathematical and Computing Sciences, Victoria University, Wellington, New Zealand.
- Roberts, L. A. (2001). A unified view of determinantal expansions for Jack polynomials, *Electronic Journal of Combinatorics* **8**(1). #R3.
- Robinson, G. (1961). *Representation Theory of the Symmetric Group*, Edinburgh University Press.
- Saw, J. G. (1977). Zonal polynomials: an alternative approach, *Journal of Multivariate Analysis* **7**: 461–467.
- Shorrocks, R. W. and Zidek, J. V. (1976). An improved estimator of the generalized variance, *Annals of Statistics* **4**(3): 629–638.
- Smith, P. J. and Gao, H. (1999). The polynomial ratio distribution, *Technical Report 99-15*, School of Mathematical and Computing Sciences, Victoria University, Wellington, New Zealand.

- Smith, P. J. and Gao, H. (2000). A determinant representation for the distribution of a generalised quadratic form in complex normal vectors, *Journal of Multivariate Analysis* **73**(1): 41–54.
- Spiegel, M. R. (1959). *Vector Analysis and an Introduction to Tensor Analysis*, Schaum's Outline Series, McGraw Hill.
- Srivastava, M. S. (1968). On the distribution of a multiple correlation matrix: non-central multivariate beta distributions, *Annals of Mathematical Statistics* **39**(1): 227–232.
- Stanley, R. P. (1986). *Enumerative Combinatorics*, Vol. 1, Wadsworth and Brooks/Coles.
- Stanley, R. P. (1989). Some combinatorial properties of Jack symmetric functions, *Advances in Mathematics* **77**: 76–115.
- Stanley, R. P. (1999). *Enumerative Combinatorics*, Vol. 2, CUP.
- Szegő, G. (1939). *Orthogonal Polynomials*, American Mathematical Society Colloquium Publications No. 23, American Mathematical Society.
- Takemura, A. (1984). *Zonal Polynomials*, Vol. 4 of *Lecture Notes - Monograph Series*, Institute of Mathematical Statistics.
- Turnbull, H. W. and Aitken, A. C. (1932). *An Introduction to the Theory of Canonical Matrices*, Blackie.
- Vilenkin, N. J. (1968). *Special Functions and the Theory of Group Representations*, Vol. 22 of *Translations of Mathematical Monographs*, American Mathematical Society.
- Vilenkin, N. J. and Klimyk, A. U. (1991). *Representation of Lie Groups and Special Functions: Volume 1*, Vol. 72 of *Mathematics and its Applications (Soviet series)*, Kluwer Academic Publishers.
- Vilenkin, N. J. and Klimyk, A. U. (1992). *Representation of Lie Groups and Special Functions: Volume 3*, Vol. 75 of *Mathematics and its Applications (Soviet series)*, Kluwer Academic Publishers.

- Vilenkin, N. J. and Klimyk, A. U. (1993). *Representation of Lie Groups and Special Functions: Volume 2*, Vol. 74 of *Mathematics and its Applications (Soviet series)*, Kluwer Academic Publishers.
- Vilenkin, N. J. and Klimyk, A. U. (1995). *Representation of Lie Groups and Special Functions: Recent Advances*, Vol. 316 of *Mathematics and its Applications*, Kluwer Academic Publishers.
- Watson, G. S. (1983). *Statistics on Spheres*, Vol. 6 of *Lecture Notes in Mathematics*, Wiley.
- Weyl, H. (1931). *The Theory of Groups and Quantum Mechanics*, Methuen.
- Weyl, H. (1946). *The Classical Groups: Their Invariants and Representations*, 2nd. edn, Princeton University Press.
- Whittaker, E. T. and Watson, G. N. (1927). *A Course of Modern Analysis*, 4th. edn, CUP.
- Zidek, J. V. (1978). Deriving unbiased risk estimators of multinormal mean and regression coefficient estimators using zonal polynomials, *Annals of Statistics* 6: 769–782.