CHAPTER 5

Complex zonal polynomials

In this chapter we study complex zonal polynomials, i.e. zonal polynomials associated with the complex normal and the complex Wishart distributions. The complex multivariate normal distribution is used in the frequency analysis of multiple time series and complex zonal polynomials are useful for noncentral distributions arising in this setting. Other than that the practical applicability of complex zonal polynomials seems rather limited. Actually our main reason of studying them is that they are *simpler* than real zonal polynomials. If one compares Farrell (1980) and Chapter 1 of Macdonald (1979) it becomes apparent that complex zonal polynomials are the same as homogeneous symmetric polynomials called the Schur functions and the latter have been extensively studied. We will make this connection clear. Hopefully developing complex zonal polynomials gives further insights into the real case.

The theory of the complex normal and the Wishart distributions very closely parallels that of the real case (see Goodman (1963) or Brillinger (1975)) and it turns out that our development of Chapter 3 and Chapter 4 can be directly translated into the complex case. In the literature on zonal polynomials it seems customary to put a $\sim$ to denote corresponding objects in the complex case. For example we use $\tilde{\mathcal{Z}}_p, \tilde{\mathcal{C}}_p, \tilde{\mathcal{Y}}_p$, etc. With this convention the translation of the results in Chapter 3 and 4 are almost immediate.

§ 5.1 THE COMPLEX NORMAL AND THE COMPLEX WISHART DISTRIBUTIONS

We give a brief summary of the complex normal and the complex Wishart distributions. Let x, y be independently distributed according to $\mathcal{N}(0,1/2)$ and let $z = x + iy$. z is said to have *the standard complex normal distribution*. Or we say that z is a *standard complex normal (random) variable*. Now let $\boldsymbol{A}$ be an $n \times n$ matrix with complex elements and let

$$\boldsymbol{u} = (u_1, \ldots, u_k)' = \boldsymbol{A}\ (z_1, \ldots, z_k)', \tag{1}$$

where $z_1, \ldots, z_k$ are independent standard complex normal variables. This scheme generates a family of distributions called the multivariate complex normal distribution. Its density (with respect to $\prod_1^k d(\Re u_i)\ \prod_1^k d(\Im u_i)$) is given by

$$f(\boldsymbol{u}) = \frac{1}{\pi^k |\boldsymbol{\Sigma}|} \exp(-\boldsymbol{u}^* \boldsymbol{\Sigma}^{-1} \boldsymbol{u}) \tag{2}$$

where $*$ means conjugate transpose and $\boldsymbol{\Sigma} = \mathcal{E}\boldsymbol{u}\boldsymbol{u}^* = \boldsymbol{A}\boldsymbol{A}^*$. If $\boldsymbol{u}$ has the density (2) we denote this by $\boldsymbol{u} \sim C\mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$. Now suppose that $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ are independently distributed according to $C\mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$. Let $\tilde{\boldsymbol{W}} = \sum_{i=1}^n \boldsymbol{u}_i \boldsymbol{u}_i^*$. The distribution of $\tilde{\boldsymbol{W}}$ is called the complex Wishart distribution and its density (with respect to $\prod_{i=1}^k d\,\tilde{w}_{ii}\ \prod_{i<j} d(\Re\, \tilde{w}_{ij}) d(\Im\, \tilde{w}_{ij})$) is given by

$$f(\tilde{\boldsymbol{w}}) = \frac{|\tilde{\boldsymbol{w}}|^{n-k} \exp(-\operatorname{tr} \boldsymbol{\Sigma}^{-1}\, \tilde{\boldsymbol{w}})}{\pi^{p(p-1)/2} \prod_{i=1}^k \Gamma(n-i+1) |\boldsymbol{\Sigma}|^n} \tag{3}$$

This distribution is denoted by $C\mathcal{W}(\boldsymbol{\Sigma}, n)$.

Let $\tilde{\boldsymbol{W}} = \tilde{\boldsymbol{T}}\tilde{\boldsymbol{T}}^*$ be the (unique) triangular decomposition of a positive definite Hermitian matrix where $\tilde{\boldsymbol{T}} = (\tilde{t}_{ij})$ is a lower triangular matrix with positive diagonal elements. Analogous to Lemma 3.1.3 we have the following lemma.

Lemma 1. *Let $\tilde{\boldsymbol{W}}$ be distributed according to $C\mathcal{W}(\boldsymbol{I}_k, \nu)$. Let $\tilde{\boldsymbol{W}} = \tilde{\boldsymbol{T}}\tilde{\boldsymbol{T}}^*$. Then $\tilde{t}_{ij}$, $i \geq j$, are independently distributed. $2^{1/2}\, \tilde{t}_{ii} \sim \chi(2(\nu - i + 1))$ and $\tilde{t}_{ij}$, $i > j$, are standard complex normal variables.*

Proof. See Goodman (1963), formula(1.8) ∎

Remark 1. $\tilde{t}_{ii}^2$ has the gamma density $f(x) = (1/\Gamma(\nu - i + 1))x^{\nu-i}e^{-x}$.

With this lemma we are ready to translate the results of Chapter 3 and 4.

§ 5.2 DERIVATION AND PROPERTIES OF COMPLEX ZONAL POLYNOMIALS

For ease of comparison of the results here and the results of Chapter 3 and 4 we will consistently put $\sim$ on corresponding objects of the complex case. This sometimes results in somewhat unnatural notation, for example if $\boldsymbol{H}$ is orthogonal then $\tilde{\boldsymbol{H}}$ is unitary etc. So much for the notation; now let us follow the development of real zonal polynomials step by step for a while. All proofs will be omitted since they are the same for the real and the complex cases.

We consider the following transformation.

$$(\tilde{\tau}_\nu \mathcal{U}_p)(\tilde{\boldsymbol{A}}) = \mathcal{E}_{\tilde{W}}\{\mathcal{U}_p(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{W}})\}, \tag{1}$$

where $\tilde{\boldsymbol{A}}$ is Hermitian and $\tilde{\boldsymbol{W}} \sim \mathcal{CW}(\boldsymbol{I}_k, \nu)$.

Lemma 1. *(corresponding to Lemma 3.1.2)* $\tilde{\tau}_\nu \mathcal{U}_p \in V_n$.

Corollary 1. *(Corollary 3.1.1)*

$$(\tilde{\tau}_\nu \mathcal{U}_p)(\tilde{\boldsymbol{A}}) = \tilde{\lambda}_{\nu p} \mathcal{U}_p(\tilde{\boldsymbol{A}}) + \sum_{q<p} \tilde{a}_{pq} \mathcal{U}_q(\tilde{\boldsymbol{A}}). \tag{2}$$

Corollary 2. *(Corollary 3.1.2)*

$$\begin{aligned} \tilde{\lambda}_{\nu p} &= \prod_{i=1}^{\ell(p)} \frac{\Gamma(p_i + \nu + 1 - i)}{\Gamma(\nu + 1 - i)} \\ &= \prod_{i=1}^{\ell(p)} (\nu + 1 - i)_{p_i} \\ &= \nu(\nu+1)\cdots(\nu + p_1 - 1) \\ &\quad \cdot(\nu - 1)\nu\cdots(\nu - 1 + p_2 - 1) \\ &\quad \cdots \\ &\quad \cdot(\nu - \ell + 1)\cdots(\nu - \ell + p_\ell) \end{aligned} \tag{3}$$

where $\ell = \ell(p)$ *and* $(a)_k = a(a+1)\cdots(a+k-1)$.

Corollary 1 shows that

$$\tilde{r}_\nu(\mathcal{U}) = \tilde{\boldsymbol{T}}_\nu \mathcal{U}, \tag{4}$$

where $\tilde{\boldsymbol{T}}_\nu$ is an upper triangular matrix with diagonal elements $\tilde{t}_{pp} = \tilde{\lambda}_{\nu p}$.

Lemma 2. *(Lemma 3.1.4) There exists a nonsingular upper triangular matrix $\tilde{\boldsymbol{B}}$ such that*

$$\tilde{\boldsymbol{B}}\tilde{\boldsymbol{T}}_\nu = \tilde{\boldsymbol{\Lambda}}_\nu \tilde{\boldsymbol{B}} \qquad \text{for all } \nu, \tag{5}$$

where $\tilde{\boldsymbol{\Lambda}}_\nu = \mathrm{diag}(\tilde{\lambda}_{\nu p},\ p \in \mathcal{P}_n)$. $\tilde{\boldsymbol{B}}$ is uniquely determined up to a (possibly different) multiplicative constant for each row.

Using this $\tilde{\boldsymbol{B}}$ we define complex zonal polynomials.

Definition 1. *(Definition 3.1.1) Complex zonal polynomials*

Let $\tilde{\boldsymbol{B}}$ be as in Lemma 2. Complex zonal polynomials $\tilde{\mathcal{Y}}_p$, $p \in \mathcal{P}_n$ are defined by

$$\tilde{\mathcal{Y}} = \begin{pmatrix} \tilde{\mathcal{Y}}_{(n)} \\ \tilde{\mathcal{Y}}_{(n-1,1)} \\ \cdot \\ \cdot \\ \cdot \\ \tilde{\mathcal{Y}}_{(1^n)} \end{pmatrix} = \tilde{\boldsymbol{B}}\mathcal{U}. \tag{6}$$

Lemma 2 is a consequence of the fact that there exists ν_0 for which $\tilde{\lambda}_{\nu_0 p}$, $p \in \mathcal{P}_n$ are all different and the following lemma.

Lemma 3. *(Lemma 3.1.5)*

$$\tilde{\boldsymbol{T}}_\nu \tilde{\boldsymbol{T}}_\mu = \tilde{\boldsymbol{T}}_\mu \tilde{\boldsymbol{T}}_\nu . \tag{7}$$

We summarize these results in the following theorem.

Theorem 1. *(Theorem 3.1.1) Let $\tilde{\mathcal{Y}}_p$ be a complex zonal polynomial then*

$$\mathcal{E}_{\tilde{W}}\, \tilde{\mathcal{Y}}_p(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{W}}) = \tilde{\lambda}_{\nu p}\, \tilde{\mathcal{Y}}_p(\tilde{\boldsymbol{A}}), \tag{8}$$

where $\tilde{\boldsymbol{W}} \sim C\mathcal{W}(\boldsymbol{I}_k, \nu)$, $\tilde{\boldsymbol{A}}$ is Hermitian, and $\tilde{\lambda}_{\nu p}$ is given by (3). Conversely (8) (for all sufficiently large ν and for all Hermitian $\tilde{\boldsymbol{A}}$) implies that $\tilde{\mathcal{Y}}_p$ is a complex zonal polynomial.

Now we explore various integral identities satisfied by complex zonal polynomials. The uniform distribution of unitary matrices can be defined as in the case of orthogonal matrices. In particular we have

Lemma 4. *(Lemma 3.2.2) Let $\tilde{\boldsymbol{U}} = (\tilde{u}_{ij})$ be a $k \times k$ matrix such that $\tilde{u}_{ij}$ are independent standard complex normal variables. Then with probability* 1 *$\tilde{\boldsymbol{U}}$ can be uniquely expressed as*

$$\tilde{\boldsymbol{U}} = \tilde{\boldsymbol{T}}\tilde{\boldsymbol{H}}, \tag{9}$$

where $\tilde{\boldsymbol{T}} = (\tilde{t}_{ij})$ is lower triangular with positive diagonal elements and $\tilde{\boldsymbol{H}}$ is unitary. Furthermore (i) $\tilde{\boldsymbol{T}}$,$\tilde{\boldsymbol{H}}$ are independent, (ii) $\tilde{\boldsymbol{H}}$ is uniform,(iii) $\tilde{t}_{ij}$ are all independent and $2^{1/2}\, \tilde{t}_{ii} \sim \chi(2(k-i+1))$, $\quad \tilde{t}_{ij}$, $i > j$, $\sim C\mathcal{N}(0,1)$.

Now we obtain the "splitting property" of complex zonal polynomials.

Theorem 2. *(Theorem 3.2.1) Let $\tilde{\boldsymbol{A}}$,$\tilde{\boldsymbol{B}}$ be $k \times k$ Hermitian matrices. Then*

$$\mathcal{E}_{\tilde{H}}\, \tilde{\mathcal{Y}}_p(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{H}}\tilde{\boldsymbol{B}}\tilde{\boldsymbol{H}}^*) = \tilde{\mathcal{Y}}_p(\tilde{\boldsymbol{A}})\,\tilde{\mathcal{Y}}_p(\tilde{\boldsymbol{B}})/\,\tilde{\mathcal{Y}}_p(\boldsymbol{I}_k), \tag{10}$$

where $k \times k$ unitary $\tilde{\boldsymbol{H}}$ has the uniform distribution.

Definition 2. A random Hermitian matrix $\tilde{\boldsymbol{V}}$ is said to have a *unitarily invariant distribution* if for every unitary $\tilde{\boldsymbol{\Gamma}}$, $\tilde{\boldsymbol{\Gamma}}\tilde{\boldsymbol{V}}\tilde{\boldsymbol{\Gamma}}^*$ has the same distribution as $\tilde{\boldsymbol{V}}$.

As in the real case Theorem 1 generalizes to unitarily invariant distributions.

Theorem 3. *(Theorem 3.2.2) Suppose that $\tilde{V}$ has a unitarily invariant distribution, then for Hermitian $\tilde{A}$*

$$\mathcal{E}_{\tilde{V}}\, \tilde{y}_p(\tilde{A}\tilde{V}) = c_p\, \tilde{y}_p(\tilde{A}), \tag{11}$$

where

$$c_p = \mathcal{E}_{\tilde{V}}\{\tilde{y}_p(\tilde{V})\} / \tilde{y}_p(I_k). \tag{12}$$

Unitarily invariant distributions are characterized as follows.

Lemma 5. *(Lemma 3.2.3) Let $\tilde{V} = \tilde{H}\tilde{D}\tilde{H}^*$ where $\tilde{H}$ is unitary and $\tilde{D}$ is diagonal. Let $\tilde{H}$ and $\tilde{D}$ be independently distributed such that $\tilde{H}$ has the uniform distribution. (Diagonal elements of $\tilde{D}$ can have any distribution.) Then $\tilde{V}$ has a unitarily invariant distribution. Conversely all unitarily invariant distributions can be obtained in this way.*

We can replace $\tilde{H}$ in Theorem 2 by $\tilde{U}$ whose elements are independent standard complex normal variables.

Theorem 4. *(Theorem 3.2.3) Let $\tilde{U} = (\tilde{u}_{ij})$ be a $k \times k$ matrix such that $\tilde{u}_{ij}$ are independent standard complex normal variables. Then for Hermitian $\tilde{A}, \tilde{B}$*

$$\mathcal{E}_{\tilde{U}}\, \tilde{y}_p(\tilde{A}\tilde{U}\tilde{B}\tilde{U}^*) = \frac{\tilde{\lambda}_{kp}}{\tilde{y}_p(I_k)}\, \tilde{y}_p(\tilde{A})\, \tilde{y}_p(\tilde{B}). \tag{13}$$

As in the real case this leads to the following observation.

Theorem 5. *(Theorem 3.2.4) $\tilde{b}_p \equiv \tilde{\lambda}_{kp} / \tilde{y}_p(I_k)$ is a constant independent of k.*

Unitarily biinvariant distributions are defined in an obvious way.

Definition 3. A random matrix $\tilde{X}$ has a *unitarily biinvariant distribution* if for every unitary $\tilde{\Gamma}_1, \tilde{\Gamma}_2$, the distribution of $\tilde{\Gamma}_1 \tilde{X} \tilde{\Gamma}_2$ is the same as the distribution of $\tilde{X}$.

Now Theorem 2 and Theorem 4 generalize as follows.

Theorem 6. *(Theorem 3.2.5) If $\tilde{\boldsymbol{X}}$ has a unitarily biinvariant distribution then for Hermitian $\tilde{\boldsymbol{A}},\tilde{\boldsymbol{B}}$*

$$\mathcal{E}_{\tilde{X}}\, \tilde{\mathcal{Y}}_p(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{X}}\tilde{\boldsymbol{B}}\tilde{\boldsymbol{X}}^*) = \gamma_p\, \tilde{\mathcal{Y}}_p(\tilde{\boldsymbol{A}})\, \tilde{\mathcal{Y}}_p(\tilde{\boldsymbol{B}}), \tag{14}$$

where

$$\gamma_p = \mathcal{E}_{\tilde{X}}\{\tilde{\mathcal{Y}}_p(\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}^*)\}/\{\tilde{\mathcal{Y}}_p(\boldsymbol{I}_k)\}^2. \tag{15}$$

Characterization of unitarily biinvariant distributions can be given in an obvious way.

Lemma 6. *(Lemma 3.2.4) Let $\tilde{\boldsymbol{X}} = \tilde{\boldsymbol{H}}_1\tilde{\boldsymbol{D}}\tilde{\boldsymbol{H}}_2$ where $\tilde{\boldsymbol{H}}_1,\tilde{\boldsymbol{H}}_2$ are unitary and $\tilde{\boldsymbol{D}}$ is diagonal. Let $\tilde{\boldsymbol{H}}_1,\tilde{\boldsymbol{H}}_2,\tilde{\boldsymbol{D}}$ be independently distributed such that $\tilde{\boldsymbol{H}}_1$, $\tilde{\boldsymbol{H}}_2$ have the uniform distribution. ($\tilde{\boldsymbol{D}}$ can have any distribution.) Then $\tilde{\boldsymbol{X}}$ has a unitarily biinvariant distribution. Conversely all unitarily biinvariant distributions can be obtained in this way.*

Remark 1. *The notion of unitarily biinvariant distributions applies to rectangular matrices as well.*

Now we take a look at the integral representation of zonal polynomials in the complex case.

Definition 4. A particular normalization of a zonal polynomial denoted by $\tilde{Z}_p$ is defined by

$$\tilde{Z}_p(\boldsymbol{I}_k) = \tilde{\lambda}_{kp}, \tag{16}$$

or $\tilde{b}_p = 1$ in Theorem 5.

Theorem 7. *(Theorem 3.3.1) Let $p = (p_1,\ldots,p_\ell)$. For $k \times k$ Hermitian $\tilde{\boldsymbol{A}}$*

$$\tilde{Z}_p(\tilde{\boldsymbol{A}}) = \mathcal{E}_{\tilde{U}}\{\tilde{\Delta}_1^{p_1-p_2}\,\tilde{\Delta}_2^{p_2-p_3}\cdots\tilde{\Delta}_\ell^{p_\ell}\}, \tag{17}$$

where $\tilde{\Delta}_i = \tilde{U}\tilde{A}\tilde{U}^*(1,\ldots,i)$ is the determinant of the $i \times i$ upper left minor of $\tilde{U}\tilde{A}\tilde{U}^*$ and $\tilde{U}$ is a $k \times k$ random matrix whose entries are independent standard complex normal variables.

(17) implies that $\tilde{Z}_p(\tilde{A})$ is positive for positive definite $\tilde{A}$ and increasing in each root. Furthermore using the Gale-Ryser theorem (see Remark 4.1.1 and Remark 4.1.2) the coefficients of $\mathcal{M}_q$ in $\tilde{Z}_p$ are nonnegative and they are positive iff $p \succ q$.

As in the real case ${}_1\tilde{b}_p$ denotes the leading coefficient of $\tilde{Z}_p$, namely

$$\tilde{Z}_p = {}_1\tilde{b}_p \, {}_1\tilde{y}_p \,. \tag{18}$$

Theorem 8. *(Theorem 4.2.2)*

$$\begin{aligned} {}_1\tilde{b}_p &= \prod_{i=1}^{\ell(p)} \prod_{j=1}^{i} (i - j + 1 + p_j - p_i)_{p_i - p_{i+1}} \\ &= \frac{\prod_{i=1}^{\ell(p)} (p_i - i + \ell(p))!}{\prod_{i<j} (p_i - p_j - i + j)} . \end{aligned} \tag{19}$$

Other than mentioning Theorem 4.2.2 we will not follow the development of Chapter 4. Of course all the results of Chapter 4 can be translated into the complex case as has been done so far. However, it is pointless to go into numerical aspects of complex zonal polynomials because, as mentioned above, complex zonal polynomials are the Schur functions and the Schur functions are already well studied. Although the translation of the results in Chapter 4 presents an alternative "probabilistic" derivation of properties of the Schur functions, it is hardly more advantageous than a well developed standard approach to the subject. See Chapter 1 of Macdonald (1979) for example. The link between complex zonal polynomials and the Schur functions is given by Saw's generating function.

Saw's generating function in the complex case was introduced by Farrell (1980). Let $\tilde{u}_{ij}$ be a standard complex normal variable. Then $2|\tilde{u}_{ij}|^2 =$

$2\,\tilde{u}_{ij}\,\tilde{u}^*_{ij} \sim \chi^2(2)$ (i.e. $|\tilde{u}_{ij}|^2$ has the standard exponential distribution). Therefore by considering $\mathcal{E}_{\tilde{U}}\{\exp(\theta\,\mathrm{tr}\,\tilde{\boldsymbol{A}}\tilde{\boldsymbol{U}}\tilde{\boldsymbol{B}}\tilde{\boldsymbol{U}}^*)\}$ where $\tilde{\boldsymbol{A}}= \mathrm{diag}(\alpha_1, \dots, \alpha_k)$, $\tilde{\boldsymbol{B}} = \mathrm{diag}(\beta_1, \dots, \beta_k)$, and $\tilde{\boldsymbol{U}}$ is composed of independent standard complex normal variables, we obtain

Theorem 9. *(Theorem 3.4.1)*

$$\prod_{i,j}^{k}(1 - \theta\alpha_i\beta_j)^{-1} = \sum_{n=0}^{\infty}(\theta^n/n!) \sum_{p\in\mathcal{P}_n} \tilde{d}_p\, \tilde{Z}_p(\tilde{\boldsymbol{A}})\, \tilde{Z}_p(\tilde{\boldsymbol{B}}), \tag{20}$$

where $\tilde{d}_p$ is determined by

$$(\mathrm{tr}\,\tilde{\boldsymbol{A}})^n = \sum_{p\in\mathcal{P}_n} \tilde{d}_p\, \tilde{Z}_p(\tilde{\boldsymbol{A}}). \tag{21}$$

Coefficients of $\tilde{Z}_p$ can be obtained as in the real case, namely (i) compare the coefficients of θ^n in both sides of (20), (ii) express the left hand side as a quadratic form in $\mathcal{M}_p$ or $\mathcal{U}_p$, (iii) do the triangular decomposition to the resulting positive definite symmetric matrix of coefficients. Now it will be shown in the next section that the Schur functions $\mathcal{S}_p$ satisfy the same generating function(20) and $\mathcal{S}_p$ is a linear combination of lower order $\mathcal{M}_q$'s ($\mathcal{S}_p = \sum_{q\leq p} a_{pq}\mathcal{M}_q$). Therefore the Schur functions agree with the complex zonal polynomials by the uniqueness of the triangular decomposition of a positive definite symmetric matrix.

§ 5.3 SCHUR FUNCTIONS

In this section we give a definition of the Schur functions and show that they coincide with complex zonal polynomials by using Saw's generating function. In terms of the Schur functions Saw's generating function is given in Section 1.4 of Macdonald (1979) or in Section 7.6 of Weyl (1946).

Let $p = (p_1, \dots, p_\ell) \in \mathcal{P}_n$. The Schur function $\mathcal{S}_p(x_1, \dots, x_k)$ $(k \geq \ell)$

is defined by

$$
\begin{aligned}
&S_p(x_1,\dots,x_k)\\
&\quad= \det(x_j^{p_i+k-i})_{1\le i,j\le k} \;/\; \det(x_j^{k-i})_{1\le i,j\le k}
\end{aligned}
\tag{1}
$$

$$
= \begin{vmatrix} x_1^{p_1+k-1} & \dots & x_k^{p_1+k-1} \\ x_1^{p_2+k-2} & \dots & x_k^{p_2+k-2} \\ \vdots & & \vdots \\ x_1^{p_k} & \dots & x_k^{p_k} \end{vmatrix} \div \begin{vmatrix} x_1^{k-1} & \dots & x_k^{k-1} \\ \vdots & & \vdots \\ x_1 & \dots & x_k \\ 1 & \dots & 1 \end{vmatrix}.
$$

If $k < \ell$ we define $S_p(x_1,\dots,x_k) = 0$. See formula (3.1), Section 1.3 of Macdonald (1979). It is given in formula (35) of James (1964). In Weyl (1946) it is introduced as the primitive character of the unitary group and as the polynomial character of the general linear group (see Sections 7.5 and 7.6 of Weyl (1946)).

Note that the denominator of (1) is the *Vandermonde determinant*

$$
\det(x_j^{k-i}) = \prod_{i<j}(x_i - x_j). \tag{2}
$$

Clearly the numerator has $(x_i - x_j)$ as a factor because if $x_i = x_j$ then $\det(x_j^{p_i+k-i}) = 0$. Running (i,j) over all pairs we see that the numerator has the Vandermonde determinant as a factor. Furthermore if x_i and x_j are interchanged then both the numerator and the denominator change the sign and the ratio remains the same. Therefore $S_p(x_1,\dots,x_k)$ is a symmetric polynomial in x_i's. It is easy to see that it is homogeneous of degree $|p|$. Now we want to show that

$$
S_p(x_1,\dots,x_k,0) = S_p(x_1,\dots,x_k). \tag{3}
$$

The last column of $(x_j^{p_i+k+1-i})_{1\le i,j\le k+1}$ is

$$
(x_{k+1}^{p_1+k},\dots,x_{k+1}^{p_k+1},x_{k+1}^{p_{k+1}})'.
$$

If $x_{k+1} = 0$ it reduces to $(0,\dots,0,1)'$. (Note that $p_{k+1} = 0$ by definition.) Hence if $x_{k+1} = 0$ then $\det(x_j^{p_i+k+1-i}) = (\prod_{j=1}^k x_j)\det(x_j^{p_i+k-i})$, the right hand side being the determinant of the $k \times k$ principal minor of the matrix on the left hand side. Similarly $\det(x_j^{k+1-i}) = (\prod_{j=1}^k x_j)\det(x_j^{k-i})$. Therefore we have (3) and in general by induction

$$S_p(x_1,\ldots,x_k,0\ldots,0) = S_p(x_1,\ldots,x_k). \tag{4}$$

This shows that $S_p \in V_n$. Now let us look at the highest monomial in S_p of the form $ax_1^{q_1}\cdots x_k^{q_k}$ $((q_1,\ldots,q_k) \in P_n)$. In $\det(x_j^{p_i+k-i})$ and $\det(x_j^{k-i})$ the similar terms are obtained by the products of the diagonal elements. They are

$$x_1^{p_1+k-1}x_2^{p_2+k-2}\cdots x_k^{p_k}, \qquad x_1^{k-1}\cdots x_{k-1}$$

respectively. From $S_p(x_1,\ldots,x_k)\det(x_j^{k-i}) = \det(x_j^{p_i+k-i})$ we obtain

$$(ax_1^{q_1}\cdots x_k^{q_k})(x_1^{k-1}\cdots x_{k-1}) = x_1^{p_1+k-1}\cdots x_k^{p_k}$$

Therefore $a = 1$ and $q = (q_1,\ldots,q_k) = (p_1,\ldots,p_k) = p$. We summarize these results in a lemma.

Lemma 1.

$$S_p = M_p + \sum_{q<p} a_{pq}M_q, \tag{5}$$

and $\{ S_p, p \in P_n \}$ *forms a basis of* V_n.

Now we prove the following.

Lemma 2.

$$\prod_{i,j}^{k}(1-\theta x_i y_j)^{-1} = \sum_{n=0}^{\infty}\theta^n \sum_{p\in P_n} S_p(x_1,\ldots,x_k)S_p(y_1,\ldots,y_k). \tag{6}$$

Proof. Replacing x_i by θx_i we can assume $\theta = 1$ without loss of generality. We prove (6) in the following form:

$$\frac{\det(x_j^{k-i})\det(y_j^{k-i})}{\prod_{i,j=1}^k(1-x_iy_j)} = \sum_{n=0}^{\infty}\sum_{p\in P_n}\det(x_j^{p_i+k-i})\det(y_j^{p_i+k-i}). \tag{7}$$

We recognize the left hand side of (7) as Cauchy's determinant, i.e.

$$\frac{\det(x_j^{k-i})\det(y_j^{k-i})}{\prod(1-x_iy_j)} = \det\left(\frac{1}{1-x_iy_j}\right)_{1\le i,j\le k} \tag{8}$$

See Lemma 2 of Anderson and Mentz (1977) for example. To prove this directly consider the matrix on the right hand side. Now subtract appropriate multiples of the first row from other rows so that each element in the first column is converted to 0, except for that in the first row. Then the rest of the right hand matrix becomes

$$\frac{1}{1-x_iy_j} - \frac{1}{1-x_1y_j}\frac{1-x_1y_1}{1-x_iy_1} = \frac{x_1-x_i}{1-x_iy_1}\frac{y_1-y_j}{1-x_1y_j}\frac{1}{1-x_iy_j} \tag{9}$$

The first two factors of the right hand side come out of the determinant as common factors and (8) is now proved by induction on dimensionality. Now to derive the right hand side of (7) from the Cauchy's determinant, expand $(1-x_iy_j)^{-1}$ as $1+x_iy_j+x_i^2y_j^2+\dots$ for every element of the matrix and then expand the determinant. Consider the term of the form $cx_1^{\ell_1}\cdots x_k^{\ell_k}$, $\ell_1 \ge \cdots \ge \ell_k$. This term arises as follows. Take ℓ_i-*th* power term in each element of the *i-th* row, $i = 1,\dots,k$. Then $x_i^{\ell_i}$ comes out as a common factor and we obtain $x_1^{\ell_1}\cdots x_k^{\ell_k}\det(y_j^{l_i})$. Collecting permuted terms in x's we have $\det(x_j^{\ell_i})\det(y_j^{\ell_i})$. Therefore Cauchy's determinant can be expanded as

$$\sum_{\ell_1\ge\cdots\ge\ell_k} \det(x_j^{\ell_i})\det(y_j^{\ell_i}). \tag{10}$$

Now if $\ell_i = \ell_{i+1}$ for some i, then $\det(x_j^{\ell_i}) = 0$. Hence this summation is actually over the set $\{(\ell_1,\dots,\ell_k) : \ell_1 > \cdots > \ell_k\}$. Now letting $\ell_i = p_i + k - i$, $i = 1,\dots,k$ the summation becomes over all partitions $p = (p_1,\dots,p_k)$. This proves the lemma. ∎

This proof has been adapted from p.202, Section 7.6 of Weyl (1946).

Comparing (5.2.20) and (6) we have

$$\sum_{p\in\mathcal{P}_n} (\tilde{d}_p/n!)\,\tilde{Z}_p(\boldsymbol{A})\,\tilde{Z}_p(\boldsymbol{B}) = \sum_{z\in\mathcal{P}_n} S_p(\boldsymbol{A})S_p(\boldsymbol{B}) \tag{11}$$

Now when expressed with respect to the basis $\{\mathcal{M}_q\}$, both $\tilde{\mathcal{Z}}_p$ and $\mathcal{S}_p$ are linear combinations of $\mathcal{M}_q$ with $q \leq p$. Therefore if (11) is expressed in terms of $\mathcal{M}_q$'s, then two sides of (11) give the same triangular (: lower times upper) decomposition of a positive definite coefficient matrix. By the uniqueness of the triangular decomposition of a symmetric positive definite matrix we have $\mathcal{S}_p = c_p\, {}_1\tilde{\mathcal{Y}}_p$ for some c_p. Comparing the leading coefficient (see (5)) we obtain $c_p = 1$. Furthermore considering the leading coefficient of $\tilde{\mathcal{Z}}_p$ we obtain $\tilde{d}_p\, {}_1\tilde{b}_p^2 = n!$, $n = |p|$. This was mentioned at the end of Section 4.2.

We have proved

Theorem 1.

$$(12) \qquad \mathcal{S}_p = {}_1\tilde{\mathcal{Y}}_p \quad \textit{and} \quad \tilde{d}_p\, {}_1\tilde{b}_p^2 = n! \quad \textit{where} \quad n = |p|.$$

There are three more determinantal expressions stated in James (1964). One involving elementary symmetric functions (formula(37) in James (1964)) is found in (2.9') of Macdonald. One involving "complete symmetric functions" (formula(36) of James (1964)) is given in (3.4) of Macdonald. Formula (38) in James (1964) is not given in Macdonald.

Elementary symmetric functions and complete symmetric functions of the roots of a matrix are relatively easy to calculate. Determinants can be evaluated easily by computer as well. Hence from the viewpoint of numerical computation these determinantal expressions seem to be all we have to know. Namely we do not need to know the coefficients of $\mathcal{M}_p$ or $\mathcal{U}_p$ etc. to evaluate the Schur function. It might be worthwhile to look for an analogue of this for the real zonal polynomial. Another possibility is to express the real zonal polynomials in terms of the Schur functions.

§ 5.4 RELATION BETWEEN THE REAL AND THE COMPLEX ZONAL POLYNOMIALS

We finish this chapter by discussing some results which we were unable to derive by our elementary approach. James (1964) gives the following formula relating the complex and the real zonal polynomials:

$$\frac{Z_p(\boldsymbol{X}\boldsymbol{X}')}{Z_p(\boldsymbol{I}_k)} = \mathcal{E}_H S_{2p}(\boldsymbol{X}\boldsymbol{H}), \tag{1}$$

where the $k \times k$ $\boldsymbol{H}$ has the uniform distribution of orthogonal matrices, $p = (p_1, \ldots, p_\ell) \in \mathcal{P}_n$, and $2p = (2p_1, \ldots, 2p_\ell) \in \mathcal{P}_{2n}$. (Formula (34) in James (1964).) Furthermore he states

$$\mathcal{E}_H S_p(\boldsymbol{X}\boldsymbol{H}) = 0, \tag{2}$$

if one or more parts of p is odd. (Formula (40)). See also Theorem 12.11.6 and Remark 12.11.11 in Farrell (1976).

Given these results we can evaluate d_p in (3.4.1) as follows. First note that by replacing $\boldsymbol{H}$ by $\boldsymbol{U}$ where $\boldsymbol{U}$ is composed of independent standard (real) normal variables we obtain

$$\begin{aligned} \mathcal{E}_U S_{2p}(\boldsymbol{X}\boldsymbol{U}) &= \mathcal{E}_{T,H} S_{2p}(\boldsymbol{X}\boldsymbol{T}\boldsymbol{H}) \\ &= \mathcal{E}_T Z_p(\boldsymbol{X}\boldsymbol{T}\boldsymbol{T}'\boldsymbol{X}')/Z_p(\boldsymbol{I}_k) \\ &= Z_p(\boldsymbol{X}\boldsymbol{X}'). \end{aligned} \tag{3}$$

By (5.2.21) and (5.3.12)

$$\begin{aligned} (\operatorname{tr}\boldsymbol{A})^{2n} &= \sum_{p\in\mathcal{P}_{2n}} \tilde{d}_p \tilde{Z}_p(\boldsymbol{A}) \\ &= \sum_{p\in\mathcal{P}_{2n}} \tilde{d}_p \, {}_1\tilde{b}_p \, {}_1\tilde{\mathcal{Y}}_p(\boldsymbol{A}) \\ &= \sum_{p\in\mathcal{P}_{2n}} (2n)! \, {}_1\tilde{b}_p^{-1} \, {}_1\tilde{\mathcal{Y}}_p(\boldsymbol{A}). \end{aligned} \tag{4}$$

Now let $\boldsymbol{A} = \operatorname{diag}(\alpha_1, \ldots, \alpha_k)$ and replace $\boldsymbol{A}$ by $\boldsymbol{A}\boldsymbol{U}$. In this case

$$\operatorname{tr}\boldsymbol{A}\boldsymbol{U} = \sum_{i=1}^{k} \alpha_i u_{ii} \sim \mathcal{N}(0, \textstyle\sum \alpha_i^2).$$

Hence

$$\begin{aligned} \mathcal{E}_U(\operatorname{tr}\boldsymbol{A}\boldsymbol{U})^{2n} &= 1\cdot 3\cdots(2n-1)(\textstyle\sum \alpha_i^2)^n \\ &= \frac{(2n)!}{2^n n!}(\operatorname{tr}\boldsymbol{A}\boldsymbol{A}')^n. \end{aligned} \tag{5}$$

On the other hand by (3) and (2)

$$\mathcal{E}_U \sum_{p \in \mathcal{P}_{2n}} (2n)!\, {}_1\tilde{b}_p^{-1}\, {}_1\tilde{\mathcal{Y}}_p(\boldsymbol{AU}) = \sum_{p \in \mathcal{P}_n} (2n)!\, {}_1\tilde{b}_{2p}^{-1}\, Z_p(\boldsymbol{AA'}). \tag{6}$$

Hence comparing (4), (5), and (6) we obtain

$$(\operatorname{tr} \boldsymbol{AA'})^n = \sum_{p \in \mathcal{P}_n} {}_1\tilde{b}_{2p}^{-1}\, 2^n n!\, Z_p(\boldsymbol{AA'})$$

or

$$d_p = 2^n n!\, {}_1\tilde{b}_{2p}^{-1}\,. \tag{7}$$

Now (5.2.19) gives (3.4.12).