All Invariant Moments of the Wishart Distribution

GÉRARD LETAC

Université Paul Sabatier

HÉLÈNE MASSAM

York University

ABSTRACT. In this paper, we compute moments of a Wishart matrix variate U of the form $\mathbb{E}(Q(U))$ where $Q(u)$ is a polynomial with respect to the entries of the symmetric matrix u, invariant in the sense that it depends only on the eigenvalues of the matrix u. This gives us in particular the expected value of any power of the Wishart matrix U or its inverse U^{-1}. For our proofs, we do not rely on traditional combinatorial methods but rather on the interplay between two bases of the space of invariant polynomials in U. This means that all moments can be obtained through the multiplication of three matrices with known entries. Practically, the moments are obtained by computer with an extremely simple Maple program.

Key words: eigenvalues of random matrices, Schur polynomials, Wishart distribution, zonal polynomials

1. Introduction

The Wishart distribution is the distribution of the sample variance–covariance matrix of a multivariate Gaussian model. As such, it is fundamental to inference in multivariate statistics. In particular its moments and the moments of its inverse are needed to approximate the distribution of many test statistics. For example, Krishnamoorty & Mathew (1999) use the first and second moments of the Wishart and the inverse Wishart in approximation methods for computing tolerance factors for a multivariate Gaussian population. The inverse Wishart distribution has also been used in discriminant analysis (see Das Gupta, 1968; Haff, 1982), in obtaining moments of the maximum likelihood estimator on a growth curve model (von Rosen, 1988, 1991, 1997) and in spectral estimation (see Shaman, 1980). In Bayesian statistics, the inverse Wishart distribution is used as a prior distribution for the covariance parameter when sampling from a multivariate Gaussian distribution (see Anderson, 1984, section 8.7.7, or Dawid, 1981). Some distributions related to the Wishart have been introduced in the more recent theory of graphical Gaussian models. The moments of these distributions are needed in evaluating the strength of the correlation between variables. They can be deduced quite easily from the moments of the Wishart distribution. It is therefore necessary to be able to compute the moments of the Wishart and inverse Wishart distribution.

Our aim in this paper is to give the explicit expression of certain moments of the Wishart matrix U and its inverse U^{-1} and show that they can be obtained in an automatic way with a simple Maple program. We obtain in particular $\mathbb{E}(U^{\pm k})$ for $k = 1, 2, \ldots$. For instance, for a real Wishart matrix U with scale parameter Σ and shape parameter p, we obtain

$$\mathbb{E}(U^3) = n\Sigma(\mathrm{tr}\ \Sigma)^2 + (n^2 + n)(\Sigma\mathrm{tr}\ (\Sigma^2) + 2\Sigma^2\mathrm{tr}\ \Sigma) + (n^3 + 3n^2 + 4n)\Sigma^3.$$

The moments we consider are also given for Hermitian Wishart matrices with complex entries. In the last 40 years, there has been a steady interest in the extension of the Gaussian distribution from $\mathbb{R}^r$ to $\mathbb{C}^r$ (see Goodman, 1963). To these Gaussian distribution on $\mathbb{C}^r$ are naturally associated complex Wishart distributions on the cone of non-negative complex Hermitian (r, r) matrices. While the extension to complex numbers was originally motivated by the study of complex stationary Gaussian processes in signal processing (see for instance Maiwald & Kraus, 2000), a new motivation for the complex Wishart is found in the study of random matrices (Haagerup & Thorbjørnsen, 1998) used in physics, in free probability and even in number theory, where explicit formulae for the moments of complex Wishart are extremely useful (see Graczyk *et al.*, 2003). Going from the complex to the quaternionic case as done by Andersson (1975) and Hanlon *et al.* (1992) is a natural and relatively easy step. However, as we are not aware of applications of the Wishart distribution in the quaternionic case, we work here with real or complex matrices only, although all our results are valid in the quaternionic case as well, with minor modifications (see Letac & Massam, 2003) for a more general version of the present paper).

Our first result, theorem 1 gives an explicit formula for any moment of order k for the Wishart distribution, that is, it gives an explicit expression for the expected value $\mathbb{E}(\prod_{i=1}^{k}\mathrm{tr}(Uh_i))$ of a product of traces $\mathrm{tr}(Uh_i)$ where U is Wishart and h_i are given Hermitian matrices. But looking at formula (15) in theorem 1 quickly reveals the importance of certain polynomials $r_{(i)}$, indexed by the set I_k of k-tuples $(i) = (i_1,\ldots, i_k)$ of integers $i_1,\ldots, i_k$ such that $i_1 + 2i_2 + \cdots + ki_k = k$. This set of polynomials $r_{(i)}$ is known to form a basis of the space $P(V)_k^K$ of homogeneous K-invariant polynomials of order k where V is the space of $r \times r$ Hermitian matrices and K is the group of transformations on the cone of positive definite Hermitian matrices, of the form $s \mapsto usu^*$ for u orthogonal or unitary and u^* the conjugate transpose of u. There is another basis of this same space which is also very interesting, that is the basis of spherical polynomials $\Phi_{(\mathbf{m})}$ indexed by the set $M_{k,r}$ of sequences $(m) = (m_1, m_2,\ldots, m_r)$ of integers $m_1 \geq m_2 \geq \cdots \geq m_r$ with $m_1 + m_2 + \cdots + m_r = k$. These polynomials have been known to statisticians for a long time since, in the real case, they correspond to the zonal polynomials. Their very special property is that if U is Wishart distributed with scale parameter σ, then the expected value of $\Phi_{(\mathbf{m})}(U)$ is a scalar multiple of $\Phi_{(\mathbf{m})}(\sigma)$ and similarly the expected value of $\Phi_{(\mathbf{m})}(U^{-1})$ is a scalar multiple of $\Phi_{(\mathbf{m})}(\sigma^{-1})$. Then if the expected values of the $\Phi_{(\mathbf{m})}(U)$ and $\Phi_{(\mathbf{m})}(U^{-1})$ are known, to compute any invariant moment, that is, the expected value of any element of $P(V)_k^K$, we need only know the expression of that element in the basis of spherical polynomials. The expression of any $r_{(i)}$ in the basis of spherical polynomials can be obtained with a simple Maple program developed by Stembridge (1998) and, therefore, we can compute using this Maple program any invariant moment of the Wishart distribution and its inverse. We recall the properties of these two selected bases of $P(V)_k^K$ in section 4, where we also show that the computation of any invariant moment for U or U^{-1} is in fact reduced to the multiplication of three known matrices. The explicit expression for these moments is given in (24) and (29). These formulae are repeated and refined in theorem 2 which is our second important result. In section 5, through a process, which we call 'lifting', we show how, with the help of the invariant moments, we can also compute the expected value of $U^{\pm k}$ and $(U^{-1})^{\pm k}$ and other interesting expressions. The computation of all the moments we have studied can be implemented on the computer. In section 6, we give the algorithm and some examples. The appendix contains a number of proofs.

2. Notation and definitions

2.1. The Wishart distribution

In order to treat the real and complex Wishart distributions simultaneously, we introduce the parameter d, which takes the value $d = 1$ in the real case and $d = 2$ in the complex case.

We write K_d, $d = 1, 2$, respectively, for $K_1 = \mathbb{R}$ and $K_2 = \mathbb{C}$. A square matrix x of order r with entries in K_d is said to be Hermitian if it is equal to its conjugate transpose (that is its transpose in the real case). Throughout this paper, the order r of the matrix x is fixed and we denote by V the space of real Hermitian matrices of order r on K_d. The parameters r, d and $f = d/2$ are sometimes called, respectively, the rank, the Peirce constant and the half Peirce constant of V. We denote by Ω the open cone of positive definite elements of V, by $\bar{\Omega}$ the closed cone of the non-negative ones and by e the unit element of V. Let $\mathbb{O}_d(r)$ be the group of (r, r) matrices u with entries in $\mathbb{R}$, $\mathbb{C}$ such that $uu^* = e$, where u^* denotes the conjugate transpose of u. For $d = 1, 2$, $\mathbb{O}_d(r)$ is, respectively, the orthogonal and unitary group.

For the complex case, the Wishart distributions on V has been considered by Goodman (1963) and Carter (1975) and for the quaternionic case by Andersson (1975). A more general framework for these Wishart distributions would be the framework of Jordan algebras, as introduced by Jensen (1988), and subsequently studied by Casalis & Letac (1994) and Massam (1994). We will not use the framework of Jordan algebras in the present paper but rather confine ourselves to the real and complex matrix cases. We will however, in some circumstances, use Faraut & Koranyi (1994) (henceforth abbreviated as FK) written in the framework of Jordan algebras, as a reference because we find that it is, on some points, the most accessible reference even in the matrix case.

Let us now recall the definition of the Wishart distribution on V. Using the half Peirce constant $f = d/2$, the following subset of $\mathbb{R}^+$,

$$\Lambda = \{f, 2f, \ldots, (r-1)f\} \cup ((r-1)f, \infty) \tag{1}$$

is called the *Gyndikin set*. This set has the following remarkable property: for p positive, there exists a positive measure μ_p on $\bar{\Omega}$ such that for all $\theta \in -\Omega$, we have

$$\int_{\bar{\Omega}} e^{\langle \theta, u \rangle} \mu_p(du) = (\det(-\theta))^{-p} \tag{2}$$

if and only if p is in Λ (see Casalis & Letac, 1994, or Massam, 1994). If $p > (r-1)f$, then μ_p is absolutely continuous with respect to the Lebesgue measure of V and is equal to

$$\mu_p(dx) = \frac{1}{\Gamma_\Omega(p)} (\det x)^{p-(1+f(r-1))} \mathbf{1}_\Omega(x) dx, \tag{3}$$

where the constant $\Gamma_\Omega(p)$ is defined by

$$\Gamma_\Omega(p) = \pi^{dr(r-1)/4} \prod_{j=1}^{r} \Gamma(p - f(j-1)) \tag{4}$$

and the Lebesgue measure dx on V is chosen such that (3) is a probability (note that our choice of the Lebesgue measure is different from FK (p. 123) and agrees with the choice of statisticians and probabilists).

If $p = f, 2f, \ldots, (r-1)f$, the measure μ_p is singular and is concentrated on the boundary $\partial\Omega = \bar{\Omega} \backslash \Omega$ of the cone. The measure μ_f can easily be described using the image of the Lebesgue measure on $\mathbb{R}^{rd}$ (suitably identified with $\mathbb{C}^r$ for $d = 2$) by the map $z \mapsto zz^*$ from $\mathbb{R}^{rd}$ to $\bar{\Omega}$. For any integer $n = 2, \ldots, r-1$ the measure μ_{nf} is obtained as the nth power of convolution of μ_f.

The natural exponential family on V generated by μ_p is the set of probabilities on $\bar{\Omega}$ of the form $(\det(-\theta))^p \exp\langle \theta, x \rangle \mu_p(dx)$ where $\theta \in -\Omega$. A more convenient parametrization of this family of distributions is given by the map from $-\Omega$ to Ω defined by

$$\theta \mapsto \sigma(\theta) = (-\theta)^{-1}. \tag{5}$$

For $\sigma \in \Omega$ and $p \in \Lambda$, the distribution

$$\gamma_{p,\sigma}(du) = \frac{1}{(\det \sigma)^p} e^{-\langle \sigma^{-1}, u \rangle} \mu_p(du) \tag{6}$$

is called the *Wishart distribution* with *shape* parameter p and *scale* parameter σ. When $d = 1$, $p = n/2$ where n is a positive integer, and when $\Sigma = 2\sigma$ then $\gamma_{p,\sigma}$ is the classical Wishart distribution $W_r(n, \Sigma)$. As $2p$ is not necessarily an integer in (6), our definition of the Wishart distribution is slightly more general than the usual one.

The next proposition shows that the expectation operator for the Wishart distribution $\gamma_{p,\sigma}$ and denoted $T_p(\cdot)$ has certain properties we shall need later. If $n = r + fr(r - 1)$ is the dimension of the real space V, let $u_1, \ldots, u_n$ be the components of $u \in V$ in a given basis. We recall that a mapping $Q : V \mapsto \mathbb{R}$ is said to be a homogeneous polynomial of degree s if the mapping $(u_1, \ldots, u_n) \mapsto Q(u)$ is a homogeneous polynomial of degree s. It is trivial to check that this property does not depend upon the chosen basis. The proof of the following proposition is deferred to the appendix.

Proposition 1

Let a be an invertible (r, r) matrix with entries in K_d. Consider the automorphism g of Ω defined by $g(x) = axa^*$. If U follows the Wishart distribution $\gamma_{p,\sigma}$ then the distribution of $g(U)$ is $\gamma_{p,g(\sigma)}$. Furthermore, if $Q : V \to \mathbb{R}$ is any function such that

$$T_p(Q)(\sigma) \overset{\text{def}}{=} \int_V Q(u)\gamma_{p,\sigma}(du) \tag{7}$$

exists, then we have for all g as above

$$T_p(Q \circ g)(\sigma) = T_p(Q)(g(\sigma)). \tag{8}$$

Finally, if Q is a homogeneous polynomial with degree k, then $T_p(Q)$ has the same property.

2.2. The moments of the Wishart distribution

For the sake of simplicity, let us describe the problem of the computation of the moments in the space of symmetric matrices with real entries. We will give a precise definition of the moments we consider in sections 3 and 4 below. It is well known that for $p > (r - 1)/2$, $\mathbb{E}(U) = p\sigma$ (see Muirhead, 1982, or Eaton, 1983) and that for $p > (r + 1)/2$, $\mathbb{E}(U^{-1}) = \sigma^{-1}/[p - \frac{1}{2}(r+1)]$. This gives us the scalar moments $\mathbb{E}(U_{ij})$ and $\mathbb{E}((U^{-1})_{ij})$.

Given a positive integer k, if for $l = 1, \ldots, k$, (i_l, j_l) is in $\{1, \ldots, r\}^2$, we call the scalar

$$\mathbb{E}(U_{i_1,j_1} \ldots U_{i_k,j_k}) \tag{9}$$

a *moment of U* of order k. Similarly, if $p > \frac{1}{2}(r - 1) + k$ we can see that $U^{-1} = ((U^{-1})_{ij})_{1 \le i,j \le r}$ exists and so does

$$\mathbb{E}((U^{-1})_{i_1 j_1} \cdots (U^{-1})_{i_k j_k}). \tag{10}$$

The expected value (10) is called a *moment of* U^{-1}. We note that (9) and (10) can be expressed in a coordinate-free way if we consider the inner product on V defined by $\langle x, y \rangle = \mathrm{tr}(xy)$. Indeed, if $h_1, \ldots, h_k$ are in V, by choosing h_l appropriately, we have $\langle U, h_l \rangle = U_{i_l j_l}$ and the scalars

$$\mathbb{E}(\langle U, h_1 \rangle \cdots \langle U, h_k \rangle) \tag{11}$$

and

$$\mathbb{E}(\langle U^{-1}, h_1 \rangle \cdots \langle U^{-1}, h_k \rangle) \tag{12}$$

are moments of type (9) and (10), respectively. Any moment of the Wishart or its inverse can be expressed under these forms and therefore, to compute (9) and (10), it is sufficient to find an expression for these quantities in function of the parameters p and σ and of the indices (i_l, j_l), $l = 1,\ldots, k$. We give the explicit expression of (11) in theorem 1 below.

Let us note here that the calculation of moments is simplest in the complex case, and the real and quaternionic cases are in a sense dual of each other (for further details see Graczyk *et al.*, 2003 for the complex case and Hanlon *et al.*, 1992 for the real and quaternionic cases).

3. Expectation of product of traces

As mentioned in section 2, expected values of the type (9) can be computed from expected values of the type (11) by choosing the h_i's appropriately. In this section, we compute the expected value (11) when the Wishart random variable U belongs to the cone $\bar{\Omega}$ of Hermitian non-negative matrices for $d = 1, 2$. From now on $h_1,\ldots, h_k$ are fixed given elements of the space V of Hermitian (r, r) matrices. For a positive integer k, $\mathcal{S}_k$ denotes the group of permutations of $\{1,\ldots, k\}$. Any permutation π in $\mathcal{S}_k$ can be written in a unique way as a product of cycles. If $c = (j_1,\ldots, j_l)$ is one of these cycles, then, for $x_1, x_2,\ldots, x_k$ in V, we use the notation $\mathrm{tr}(\Pi_{j \in c} x_j)$ to mean

$$\mathrm{tr}(x_{j_1} x_{j_2} \cdots x_{j_l}).$$

The set of cycles of π is denoted $C(\pi)$. We define

$$r_\pi(x_1, \ldots, x_k) = \prod_{c \in C(\pi)} \mathrm{tr}\left(\prod_{j \in c} x_j \right) \tag{13}$$

and for u and h_j in V, we use the notation $r_\pi(u)(h_1,\ldots, h_k)$ for

$$r_\pi(u)(h_1, \ldots, h_k) = \prod_{c \in C(\pi)} \mathrm{tr}\left(\prod_{j \in c} u h_j \right). \tag{14}$$

To prove the theorem in this section, we state without proof the following differentiation result.

Proposition 2

Let Ω be the cone of positive definite matrices. For $\theta \in \Omega$ define $\sigma(\theta) = (-\theta)^{-1}$. Then

1. *The differential of $\theta \mapsto \sigma(\theta) = (-\theta)^{-1}$ is $h \mapsto \sigma h \sigma$.*
2. *For $d = 1, 2$ the differential of $\theta \mapsto \log \det \sigma(\theta)$ is $h \mapsto \mathrm{tr}(\sigma h)$.*

Theorem 1

Let U be a random variable on V following the Wishart $\gamma_{p,\sigma}$ distribution. Then for $d = 1, 2$ we have

$$\mathbb{E}(\langle U, h_1 \rangle \cdots \langle U, h_k \rangle) = \mathbb{E}(\mathrm{tr}(Uh_1) \cdots \mathrm{tr}(Uh_k)) = \sum_{\pi \in \mathcal{S}_k} p^{m(\pi)} r_\pi(\sigma)(h_1, \ldots, h_k), \tag{15}$$

where $m(\pi)$ is the number of cycles of π.

Remark. The reader is referred to theorem 5 below for a simpler form of (15) when $h_1 = \cdots = h_k$.

Proof. We will prove (15) by induction on k. Clearly, (15) is true for $k = 1$. Let us assume that it is true for k and show it also holds for $k + 1$. From the expression of the Laplace transform of the measure μ_p, i.e. $L_p(\theta) = \int_{\bar{\Omega}} e^{\langle\theta,u\rangle}\mu_p(du) = (\det(-\theta))^{-p} = \exp(p \log \det(\sigma(\theta)))$ where $\sigma(\theta) = (-\theta)^{-1}$, we see immediately that the kth differential of L_p for $d = 1, 2$ is

$$L_p^{(k)}(\theta)(h_1, \ldots, h_k) = L_p(\theta)\mathbb{E}(\operatorname{tr}(Uh_1)\ldots\operatorname{tr}(Uh_k)). \tag{16}$$

Let us now take the differential of these expressions in the direction h_{k+1} and use our induction assumption. The right-hand side of (16) is then the sum of products of functions of θ. The differential of $L_p(\theta)$ is $p\langle\sigma, h\rangle L_p(\theta)$. Using proposition 2 we find that the differentials of $\theta \mapsto \operatorname{tr}(\sigma h_1\sigma\cdots\sigma h_k)$ in the direction h_{k+1} is, for $d = 1, 2$,

$$\sum_{j=1}^{k} \operatorname{tr}(\sigma h_1 \sigma \cdots \sigma h_{j-1} \sigma h_{k+1} \sigma h_j \cdots \sigma h_k),$$

from which we deduce the differential of $\theta \mapsto \operatorname{tr}(_{j\in c}(\sigma h_j))$. From this we obtain the differentials of $r_\pi(\sigma h_1,\ldots, \sigma h_k)$. Combining these differentials we can see that (15) holds true for $k+1$.

The complete extension of theorem 1 to Jordan algebras (for h_i's not necessarily equal) is treated in Letac & Massam (2001).

4. Expectation of invariant polynomials in a Wishart matrix

We see in (15) that the polynomials r_π play an important role. Moreover when the arguments of r_π are equal, $r_\pi(u,\ldots, u)$, which, according to the notation given in (14), we will write $r_\pi(u)(e,\ldots, e)$ or even more simply $r_\pi(u)$, has a special property of invariance. We will use this property of invariance to compute a wide class of moments of the Wishart distribution, the invariant moments and other moments derived from the invariant ones. We first describe what we mean by invariance and then study two sets of bases for the space of invariant homogeneous polynomials of degree k.

4.1. Invariance

For $u \in \mathbb{O}_d(r)$ the endomorphism k_u of V defined by $k_u(x) = uxu^*$ maps Ω onto itself and satisfies $k_u(e) = e$. One can easily prove that any endomorphism k of V with these two properties has the form $k = k_u$ with $u \in \mathbb{O}_d(r)$. Note that k_u is an orthogonal transformation of the Euclidean space V. We define K as the group of isomorphisms of V of the form k_u for some $u \in \mathbb{O}_d(r)$ (the correspondence is almost one-to-one, as the kernel of the group homorphism $u \mapsto k_u$ is $\pm e$). By definition, two elements x and y of V belong to the same orbit of K if $y = k(x)$ for some $k \in K$ and this is so if and only if the Hermitian matrices x and y have the same eigenvalues.

Let us now define the real invariant homogeneous polynomials of degree s on V. Consider a homogeneous polynomial Q, as defined in section 2.1. We will say that Q is *K-invariant* (or simply that Q is invariant) if, for all $k \in K$ and all $x \in V$, we have $Q(k(x)) = Q(x)$. Thus Q is invariant if and only if for all $x \in V$, $Q(x)$ depends only on the eigenvalues of x. We denote by $P(V)_s^K$ the set of real invariant homogeneous polynomials of degree s on V.

In the sequel, we will actually always talk about invariant homogeneous polynomials of degree k where k is therefore an integer and not an element of K and we will thus use the notation $P(V)_k^K$ for the set of real invariant homogeneous polynomials of degree k on V. This should lead to no confusion at all as the meaning of k will always be clear from the context.

4.2. *The basis of the* $r_{(i)}$*'s*

Let us now introduce the polynomials $r_{(i)}$. In section 3, we have considered the polynomials

$$u \mapsto r_\pi(u)(h_1, \dots, h_k)$$

for π in the symmetric group $\mathcal{S}_k$ and for fixed $h_1,\dots, h_k$ in V. When all the h_i's are equal to the unit element $h = e \in V$, $r_\pi(u)(e,\dots, e)$ becomes a polynomial $Q(u) = r_{(i)}(u)$ such that $Q \in P(V)_k^K$, which we are going to describe now.

Let I_k be the set of all sequences $(i) = (i_1,\dots, i_k)$ of non-negative integers such that $i_1 + 2i_2 + \cdots + ki_k = k$. If $\pi \in \mathcal{S}_k$, let i_j be the number of cycles of π of length j. The corresponding element (i) of I_k is called the *portrait* of π. If π and π' have the same portrait, we clearly have $r_\pi(u)(h,\dots, h) = r_{\pi'}(u)(h,\dots, h)$ and so, for a given $(i) \in I_k$, it makes sense to consider

$$r_{(i)}(h) = \prod_{j=1}^{k} (\operatorname{tr} h^j)^{i_j}. \tag{17}$$

For instance $r_{(0,0,\dots,0,1)}(h) = \operatorname{tr}(h^k)$ and $r_{(k)}(h) = (\operatorname{tr} h)^k$ where from now on $r_{(k)}(h)$ is a short notation for $r_{(k,0,\dots,0)}(h)$.

For the sake of simplicity and consistently with the notation $r_{(k)}(h)$ just adopted, in the sequel we write $r_{(i)} = r_{(i_1,\dots,i_n)}$, where i_n is the last non-zero integer of the sequence (i). The $r_{(i)}$'s are important since, as we are going to see in proposition 4, they form a basis of $P(V)_k^K$. To show this, we need proposition 3, which tells us that the invariant polynomials in $u \in V$ are in 1–1 correspondence with the symmetric homogeneous polynomials in the eigenvalues of u.

Proposition 3

Let $P(V)_k^K$ be the space of K-invariant homogeneous polynomials $Q : V \to \mathbb{R}$ of degree k and let $\mathcal{S}_{k,r}$ be the space of symmetric homogeneous polynomials $S : \mathbb{R}^r \mapsto \mathbb{R}$ of degree k. Then, for all $Q \in P(V)_k^K$, there exists a unique S_Q in $\mathcal{S}_{k,r}$ such that for all $u \in V$, $Q(u) = S_Q(\lambda_1,\dots, \lambda_r)$ where $(\lambda_1,\dots, \lambda_r)$ are the eigenvalues of u. Furthermore $Q \mapsto S_Q$ is an isomorphism between the two spaces.

The proof, which is fairly easy, is omitted. It relies on the fact that any $u \in V$ can be written in an appropriate basis as a diagonal matrix with elements its eigenvalues λ_i, $i = 1,\dots, r$ and conversely any symmetric homogeneous polynomial in λ_i, $i = 1,\dots, r$ can be written in terms of the elementary symmetric functions

$$e_1(\lambda_1, \dots, \lambda_r) = \lambda_1 + \cdots + \lambda_r, \quad e_2(\lambda_1, \dots, \lambda_r) = \sum_{1 \le i \le j \le r} \lambda_i \lambda_j. \tag{18}$$

Proposition 4

Let $I_{k,r}$ be the set of $(i) = (i_1,\dots, i_k)$ in I_k such that $i_j = 0$ when $j > r$, r being the rank of V. Then $(r_{(i)})_{(i) \in I_{k,r}}$ is a basis of $P(V)_k^K$.

Proof. To show this result, we need to show that the $(r_{(i)})_{(i) \in I_{k,r}}$ are linearly independent and that the cardinality of this set is equal to the dimension of $P(V)_k^K$. If $u \in V$ has eigenvalues $(\lambda_1, \ldots, \lambda_r)$ then

$$r_{(i)}(u) = \prod_{i=1}^{r} (\lambda_1^j + \cdots + \lambda_r^j)^{i_j}$$

and therefore from the previous proposition, showing that the $(r_{(i)})_{(i) \in I_{k,r}}$ are linearly independent is equivalent to showing that the

$$p_{(i)}(\lambda_1, \ldots, \lambda_r) = S_{r_{(i)}}(\lambda_1, \ldots, \lambda_r) = \prod_{j=1}^{r} (\lambda_1^j + \cdots + \lambda_r^j)^{i_j} \tag{19}$$

are linearly independent for $(i) \in I_{k,r}$. This is a classical result and can be found, for example, in Macdonald (1995, p. 24).

Let us now show that the cardinality of $I_{k,r}$ is the dimension of $\mathcal{S}_{k,r}$, that is, from the previous proposition, the dimension of $P(V)_k^K$. Any element of $\mathcal{S}_{k,r}$ is a polynomial with respect to the elementary symmetric functions as given in (18). Since monomials of $e_1, \ldots, e_r$ in $\mathcal{S}_{k,r}$ must be of the form $e_1^{i_1} e_2^{i_2} \cdots e_r^{i_r}$ with $i_1 + \cdots + ri_r = k$, it is clear that the cardinality of $I_{k,r}$ is the dimension of $\mathcal{S}_{k,r}$ and therefore of $P(V)_k^K$.

4.3. The basis of spherical polynomials

We now introduce another basis of $P(V)_k^K$. Let $M_{k,r}$ be the set of sequences of integers $\mathbf{m} = (m_1, m_2, \ldots, m_r)$ such that $m_1 \geq m_2 \geq \cdots m_r \geq 0$ and $m_1 + \cdots + m_r = k$. The sets $M_{k,r}$ and $I_{k,r}$ as defined in proposition 4 have the same cardinality, since the correspondence $\mathbf{m} \mapsto (i)$ given by $i_j = \#\{l; m_l = j\}$ is clearly one-to-one. It is well known (see Macdonald, 1995, Chapter VI, especially sections 1 and 10) that for a given number $a > 0$, one can define a set of polynomials $(J_{\mathbf{m}}^{(a)}, \mathbf{m} \in M_{k.r})$ in $\lambda_1, \ldots, \lambda_r$, called the Jack polynomials of order $\mathbf{m}$ that form an orthogonal basis of $\mathcal{S}_{k,r}$ for a special scalar product depending on a (see Stanley, 1989, pp. 76–79; Macdonald, 1995, VI formulae (1.4) and (4.7); Lassalle, 1992, pp. 224 and 225 for a more compact presentation). For $a = 2$, the $J_{\mathbf{m}}^{(a)}$'s are the zonal polynomials, well known to statisticians; for $a = 1$ they are the Schur polynomials (see James, 1964; Takemura, 1984).

Coming back to V with rank r and Peirce constant $d = 2f = 1, 2$ and following the notation of proposition 3 we define on V the *spherical polynomial* $u \mapsto \Phi_{\mathbf{m}}(u)$ of order $\mathbf{m} \in M_{k,r}$ by

$$J_{\mathbf{m}}^{(1/f)} = S_{\Phi_{\mathbf{m}}},$$

that is, $J_{\mathbf{m}}^{(1/f)}$ is the polynomial in $\mathcal{S}_{k,r}$ corresponding to the symmetric polynomial $\Phi_{\mathbf{m}}$ in $P(V)_k^K$. (To see that they coincide with the spherical polynomials as defined in FK, p. 228, see FK, p. 239 and the references given therein.)

As mentioned before, our aim is to compute the expected value of $Q(U)$, that is $\int_V Q(u)\gamma_{p,\sigma}(du)$ for $Q \in P(V)_k^K$. As the $r_{(i)}$'s and the $\Phi_{\mathbf{m}}$'s form two different bases of $P(V)_k^K$, if we know the expectations of each element of one basis, we can compute the expectation of any element of the other basis provided we know the matrix of passage from one basis to the other. We can then compute the expectation of any $Q(U)$ with Q in $P(V)_k^K$. This is exactly what we are going to do. The expected value of $\Phi_{\mathbf{m}}(U)$ is known (proposition 5) and we will compute the expected value of $\Phi_{\mathbf{m}}(U^{-1})$ (proposition 6).

Recall the notation introduced in proposition 1. For any function $Q : V \to \mathbb{R}$, we write

$$T_p(Q)(\sigma) = \int_V Q(u)\gamma_{p,\sigma}(du). \tag{20}$$

It follows from proposition 1 that if Q belongs to $P(V)_k^K$, then $T_p(Q)$ applied to σ in Ω, an open subset of V, is also a homogeneous polynomial function of degree k. This polynomial function can be extended in a unique way to a homogeneous polynomial of degree k on V and therefore the mapping T_p defined by $Q \mapsto T_p(Q)$ is an endomorphism of $P(V)_k^K$ (which has already been considered by Kushner & Meisner, 1980). The expression of $T_p(\Phi_{\mathbf{m}})(\sigma)$ is particularly simple since, as stated in proposition 5 below, the $\Phi_{\mathbf{m}}$'s are in fact the eigenvectors of T_p.

To describe the eigenvalues of T_p we have to recall the definition of the Gamma function of V: for $s_j > (j-1)f$,

$$\Gamma_\Omega(s_1,\ldots,s_r) = \pi^{dr(r-1)/4}\prod_{j=1}^{r}\Gamma(s_j-(j-1)f). \tag{21}$$

If p is a positive real number and if $\mathbf{s} = (s_1,\ldots, s_r)$ we write $\Gamma_\Omega(p + \mathbf{s})$ for $\Gamma_\Omega(p + s_1,\ldots, p + s_r)$ and $\Gamma_\Omega(p)$ for $\Gamma_\Omega(p,\ldots, p)$ (this symbol was used in (4)). If $p > (r-1)f$ and $\mathbf{m} \in M_{k,r}$ we define the generalized Pochhammer symbol as $(p)_{\mathbf{m}} = \Gamma_\Omega(p + \mathbf{m})/\Gamma_\Omega(p)$ (see FK, p. 230). If p belongs to the singular part $\{f, 2f,\ldots, (r-1)f\}$ of the Gyndikin set, the generalized Pochhammer symbol is

$$(lf)_{\mathbf{m}} = \prod_{j=1}^{l}\frac{\Gamma(m_j+(l-j+1)f)}{\Gamma((l-j+1)f)}$$

for $l = 1, 2,\ldots, r-1$ (see Letac *et al.*, 2001).

At this point, we should mention that $(p)_{\mathbf{m}}$ depends on $\mathbf{m}$ and p, but does not depend on the rank r, in the following sense. Given a sequence $\mathbf{m} = (m_1,\ldots, m_c, 0, 0, \ldots)$ of integers such that $m_1 \geq m_2 \geq \cdots \geq m_c > 0$ and $m_1 + \cdots + m_c = k$, then $\mathbf{m}$ can be considered as a member of $M_{k,r}$ for any $r \geq c$ and the value of $(p)_{\mathbf{m}}$ will not change with r. Let us momentarily denote $(p)_{\mathbf{m}}$ by $((p)_{\mathbf{m}})_r$. We have to check that $((p)_{\mathbf{m}})_r = ((p)_{\mathbf{m}})_c$. This is easily done by considering separately the case $p = f, 2f,\ldots, (c-1)f$, the case $p = cf,\ldots, (r-1)f$, and the case $p > (r-1)f$. Our remark is thus proved.

The following proposition is most important as it gives us the eigenvectors and eigenvalues of T_p. For $p > (r-1)f$, its proof is essentially in FK (p. 226). For the singular case, see Letac *et al.* (2001, theorem 3.6). When p/f is an integer, it is also proved by beautiful combinatorial methods in Hanlon *et al.* (1992, formula 5.3, p. 166).

Proposition 5

Let U be $\gamma_{p,\sigma}$ distributed on V, where p is in the Gyndikin set (1). *For $\mathbf{m} \in M_{k,r}$ we have*

$$T_p(\Phi_{\mathbf{m}})(\sigma) = \mathbb{E}(\Phi_{\mathbf{m}}(U)) = (p)_{\mathbf{m}}\Phi_{\mathbf{m}}(\sigma), \tag{22}$$

where $(p)_{\mathbf{m}}$ is the generalized Pochhammer symbol.

Thus $\Phi_{\mathbf{m}}$ is an eigenvector of T_p associated with the eigenvalue $(p)_{\mathbf{m}}$. We have therefore solved, at least in theory, the problem of the computation of the expectation of the $\mathbb{E}(r_{(i)}(U))$'s. The only thing left to do is to write the $r_{(i)}$'s as a linear combination of the $\Phi_{\mathbf{m}}$'s or, equivalently, according to proposition 3, to write $p_{(i)}$ as defined in (19) as a linear combination of the $J_{\mathbf{m}}^{(1/f)}$'s. Following the notation of our Maple program (see section 6.2), we denote by $\mathtt{p}$ the basis of the $p_{(i)}$'s in $\mathcal{S}_{k,r}$ and by $\mathtt{J}$ the basis of the $J_{\mathbf{m}}^{(1/f)}$'s. With some abuse of language we also identify $\mathtt{p}$ and $\mathtt{J}$ to the corresponding basis $(r_{(i)}, (i) \in I_{k,r})$ and $(\Phi_{\mathbf{m}}, \mathbf{m} \in M_{k,r})$. Let $A_{k,r}$ be the passage matrix from the basis $\mathtt{p}$ to the basis $\mathtt{J}$, i.e. the matrix representative of the identity endomorphism of $\mathcal{S}_{k,r}$ from the basis $\mathtt{J}$ to the basis $\mathtt{p}$. This matrix $A_{k,r} = (P_{(i),\mathbf{m}}(a))$ is a matrix of polynomials with respect to one variable $a = 1/f$ (see Stanley, 1989, p. 97) whose rows are indexed by $I_{k,r}$ and columns are indexed by $M_{k,r}$. The inverse matrix $A_{k,r}^{-1} = (Q_{\mathbf{m},(i)}(a))$ is the passage matrix from the basis $\mathtt{J}$ to the basis $\mathtt{p}$. We also consider the square diagonal matrix $D_{k,r}(p)$ whose rows and columns are indexed by $M_{k,r}$ and with diagonal element corresponding to $\mathbf{m}$, the Pochhammer symbol $(p)_{\mathbf{m}}$. With these notations proposition 5 implies that the matrix representative of T_p in the basis J is

$$[T_p]^{\mathrm{J}}_{\mathrm{J}} = D_{k,r}(p),$$

and the representative matrix of T_p in the basis p is

$$[T_p]^{\mathrm{p}}_{\mathrm{p}} = A_{k,r}D_{k,r}(p)A^{-1}_{k,r}. \tag{23}$$

If we write $[r_{(i)}(u)]_{I_{k,r}}$ for the column vector made with the polynomials $r_{(i)}(u)$; $(i) \in I_{k,r}$, (23) is equivalent to

$$[\mathbb{E}(r_{(i)}(U))]_{I_{k,r}} = [A_{k,r}D_{k,r}(p)A^{-1}_{k,r}]^{\mathrm{t}}[r_{(i)}(\sigma)]_{I_{k,r}}, \tag{24}$$

where M^{t} is the transposed matrix of the matrix M.

It is now time to investigate the moments of U^{-1} when U has a Wishart distribution. This is considered in the literature as a difficult problem (see Haff, 1982; von Rosen, 1988; Lauritzen, 1996, p. 259; Maiwald & Kraus, 2000). However the expectation of $E(\Phi_{\mathbf{m}}(U^{-1}))$ and therefore of $E(r_{(i)}(U^{-1}))$ appears to be relatively easy. For $q > k - f$ and $\mathbf{m} \in M_{k,r}$ we denote

$$(q)^*_{\mathbf{m}} = \frac{\prod_{j=1}^r \Gamma(-m_j + jf + q)}{\prod_{j=1}^r \Gamma(jf + q)}. \tag{25}$$

Note that $q \mapsto (q)^*_{\mathbf{m}}$ is a rational function whose coefficients depend only on $\mathbf{m}$. A remark similar to the one made before proposition 5 is in order: for fixed q and $\mathbf{m}$, the number $(q)^*_{\mathbf{m}}$ does not depend on r. We have the following result.

Proposition 6

Let U be $\gamma_{p,\sigma}$ distributed on V, where $p > (r-1)f + k$. For $\mathbf{m} \in M_{k,r}$ we have

$$\mathbb{E}(\Phi_{\mathbf{m}}(U^{-1})) = ((p - rf)^*)_{\mathbf{m}}\Phi_{\mathbf{m}}(\sigma^{-1}) \tag{26}$$

where $(q)^*_{\mathbf{m}}$ is defined by (25).

Proof. To show that (26) is true, we need the following proposition, whose proof is given in the appendix.

Proposition 7

Let $\mathbf{m} = (m_1,\ldots, m_r)$ with $m_j \in \mathbb{R}$ and $p > f(r-1)$ such that $p - m_j > f(r-j)$ for $j = 1,\ldots, r$. Let $\mathbf{m}^* = (m_r,\ldots, m_1)$ and for $s = (s_{lk})_{1\le l,k\le r}$ in Ω, and for $j = 1,\ldots, r$, let $\Delta_j(s)$ be the determinant of the (j, j) Hermitian matrix $(s_{lk})_{1\le l,k\le j}$. We write

$$\Delta_{\mathbf{m}}(s) = (\Delta_1(s))^{m_1-m_2}(\Delta_2(s))^{m_2-m_3}\cdots(\Delta_r(s))^{m_r}.$$

Then, if the random variable S has the Wishart distribution $\gamma_{p,\sigma}$, we have

$$\mathbb{E}(\Delta_{\mathbf{m}}(S^{-1})) = \frac{\Gamma_\Omega(p - \mathbf{m}^*)}{\Gamma_\Omega(p)}\Delta_{\mathbf{m}}(\sigma^{-1}).$$

Proposition 7 enables us to complete the proof of proposition 6. Let du be the Haar measure on the compact group $\mathbb{O}_d(r)$. Then

$$\mathbb{E}(\Phi_{\mathbf{m}}(U^{-1})) \overset{(1)}{=} \mathbb{E}\left(\int_{\mathbb{O}_d(r)} \Delta_{\mathbf{m}}((uUu^*)^{-1})du\right) \overset{(2)}{=} \mathbb{E}\left(\int_{\mathbb{O}_d(r)} \Delta_{\mathbf{m}}(uU^{-1}u^*)du\right)$$
$$\overset{(3)}{=} \int_{\mathbb{O}_d(r)} \mathbb{E}(\Delta_{\mathbf{m}}(uU^{-1}u^*))du \overset{(4)}{=} \frac{\Gamma_\Omega(p-\mathbf{m}^*)}{\Gamma_\Omega(p)} \int_{\mathbb{O}_d(r)} \Delta_{\mathbf{m}}(u\sigma^{-1}u^*)du$$
$$\overset{(5)}{=} \frac{\Gamma_\Omega(p-\mathbf{m}^*)}{\Gamma_\Omega(p)} \Phi_{\mathbf{m}}(\sigma^{-1}).$$

In this sequence of equalities, (1) and (5) are due to the property of spherical functions as given in FK (p. 304, theorem XIV.3.1), (2) is clear, (3) follows from Fubini and (4) from proposition 7. The proof of proposition 6 is therefore complete.

Even the case $k = 1$ provides a non-trivial example of application of (26) as it yields $\mathbb{E}(U^{-1}) = \sigma^{-1}/(p-1-(r-1)f)$ if $p > (r-1)f + 1$, as $(q)_1^* = 1/(q+f-1)$ and $\Phi_1(u) = \text{tr } u$. If the polynomial Q is in $P(V)_k^K$, we introduce the symbol $T_p^*(Q)$ as follows

$$T_p^*(Q)(\sigma^{-1}) = \int_V Q(u^{-1})\gamma_{p,\sigma}(du) = \mathbb{E}(Q(U^{-1})). \tag{27}$$

Proposition 6 shows that T_p^*, like T_p, is an endomorphism of $P(V)_k^K$ and that the eigenvectors are the $\Phi_{\mathbf{m}}$'s again, with eigenvalues $((p-rf)^*)_{\mathbf{m}}$. Denoting by $D_{k,r}^*(q)$ the diagonal matrix whose entry corresponding to $\mathbf{m}$ is $(q)_{\mathbf{m}}^*$, we obtain the matrix representative of T_p^* in the basis $\mathfrak{p}$:

$$[T_p^*]_{\mathfrak{p}}^{\mathfrak{p}} = A_{k,r} D_{k,r}^*(p-rf) A_{k,r}^{-1}. \tag{28}$$

This is equivalent to

$$[\mathbb{E}(r_{(i)}(U^{-1}))]_{I_{k,r}} = [A_{k,r} D_{k,r}^*(p-rf) A_{k,r}^{-1}]^t [r_{(i)}(\sigma^{-1})]_{I_{k,r}}. \tag{29}$$

5. The expectations of $r_{(i)}(U^{\pm 1})$ and of $U^{\pm k}$

5.1. The expectation of $r_{(i)}(U^{\pm 1})$

In (24) and (29) above, we have already given the expected values of $r_{(i)}(U)$ and $r_{(i)}(U^{-1})$ for $k \le r$. We want to obtain these expectations for all k. We observe that for $k \le r$ we have the following four equalities:

$$I_{k,r} = I_{k,k},$$
$$D_{k,r}(p) = D_{k,k}(p),$$
$$D_{k,r}^*(q) = D_{k,k}^*(q),$$
$$A_{k,r} = A_{k,k}.$$

The first equality is obvious from the definition of $I_{k,r}$ given in proposition 4, the second and the third are consequences of the remarks made just before proposition 5 and proposition 6, respectively, and the fourth is a consequence of the properties of symmetric polynomials. A detailed proof of this last statement would imitate the proof of proposition 2 below, and so we skip it. We now denote for simplicity

$$D_{k,k}(p) = D_k(p), \quad D_{k,k}^*(q) = D_k^*(q), \quad A_{k,k} = A_k, \quad B_k = A_k^{\mathrm{t}}. \tag{30}$$

Recall that I_k has already been defined before (17). Clearly $I_{k,k} = I_k$.

The next theorem extends (24) and (29) to be valid for any k. If $r < k$ the polynomials $r_{(i)}$ are not linearly independent, thus formulae (24) and (29) do not give $\mathbb{E}(r_{(i)}(U))$ and $\mathbb{E}(r_{(i)}(U^{-1}))$ for $(i) \in I_k \backslash I_{k,r}$ in general. But we are going to see that the same results do hold true for any k. We imitate the notation used in (24) by considering the column vector $[r_{(i)}(\sigma)]_{I_k}$. Note that from now on, we use the transposed matrix $B_k = A_k^{\mathrm{t}}$ rather than A_k, as it is B_k rather than A_k which will appear in the numerical calculations of section 5.

Theorem 2

Let U follow the $\gamma_{p,\sigma}$ distribution. Then with the notation given in (30) *we have*

$$\text{for } p \in \Lambda, \quad [\mathbb{E}(r_{(i)}(U))]_{I_k} = B_k^{-1} D_k(p) B_k [r_{(i)}(\sigma)]_{I_k} \tag{31}$$

$$\text{for } \; p \geq k + (r-1)f, \quad [\mathbb{E}(r_{(i)}(U^{-1}))]_{I_k} = B_k^{-1} D_k^*(p - rf) B_k [r_{(i)}(\sigma^{-1})]_{I_k}. \tag{32}$$

Proof. For $r \geq k$ there is nothing to prove as the desired results are identical to (24) and (29). When $r = k$, (31) and (32) are identities in $P(V)_k^K$ that can be translated into identities on $\mathcal{S}_{k,k}$ by the map $Q \mapsto S_Q$ as defined in proposition 3. Thus we obtain identities between homogeneous symmetric polynomials of degree k with respect to $\lambda_1, \ldots, \lambda_k$. Let us now consider the case $r < k$ for some V with Peirce constant d and rank r. If we set $\lambda_{r+1} = \ldots = \lambda_k = 0$ in (31) and (32), we obtain identities in $\mathcal{S}_{k,r}$ and if we translate these identities through the isomorphism $Q \mapsto S_Q$ into identities in $P(V)_k^K$, we see immediately that they are nothing but (24) and (29) for V and our result is therefore proved.

The actual computation of B_k, B_k^{-1}, $B_k^{-1} D_k(p) B_k$ and $B_k^{-1} D_k^*(p - rf) B_k$ is detailed in section 6.

5.2. Extended results

If we apply proposition 1 to $Q(U) = r_{(i)}(U)$ with $g(U) = h^{1/2} U h^{1/2}$ for $h \in V$, we immediately obtain the following extension of theorem 2.

Theorem 3

Let U follow the $\gamma_{p,\sigma}$ distribution. Then for $h \in \Omega$ and with the notation given in (30) *we have*

$$\text{for } p \in \Lambda, \; [\mathbb{E}(r_{(i)}(h^{1/2} U h^{1/2})]_{I_k} = B_k^{-1} D_k(p) B_k [r_{(i)}(h^{1/2} \sigma h^{1/2})]_{I_k} \tag{33}$$

and for $p \geq k + (r - 1)f$,

$$[\mathbb{E}(r_{(i)}(h^{1/2} U^{-1} h^{1/2}))]_{I_k} = B_k^{-1} D_k^*(p - rf) B_k [r_{(i)}(h^{1/2} \sigma^{-1} h^{1/2})]_{I_k}. \tag{34}$$

We thus obtain the expected values of

$$r_{(i)}(h^{1/2} U h^{1/2}) = \prod_j (\mathrm{tr}(Uh)^j)^{i_j}, \quad r_{(i)}(h^{-1/2} U^{-1} h^{-1/2}) = \prod_j (\mathrm{tr}(Uh)^{-j})^{i_j}.$$

It is then clear that we can also compute the expectation of any $Q(h^{1/2} U h^{1/2})$ for Q in $\mathbb{P}(V)_k^K$. It will be useful to view $Q(h^{1/2} U h^{1/2})$ as both a polynomial in U and in h, which we will denote as

$\tilde{Q}(U, h)$. From proposition XIV.1.1 of FK, also given in the appendix as proposition 8, we know that the correspondence between Q and $\tilde{Q}$ is an isomorphism. The polynomial $\tilde{Q}$ is used in the next section to obtain such results as $\mathbb{E}(U^{\pm k})$.

5.3. Previous results

Results (33) when $k = 2, 3$ and (34) for $k = 2$ have already been obtained by Letac & Massam (1998) in a paper devoted to the characterization of the Wishart distribution. Indeed, for

$$M(p) = \begin{bmatrix} p^2 & p \\ pf & p - pf + p^2 \end{bmatrix}$$

Letac & Massam (1998, proposition 3.2, p. 584) have shown that

$$\begin{bmatrix} \mathbb{E}(U \otimes U) \\ \mathbb{E}(\mathbb{P}(U)) \end{bmatrix} = M(p) \begin{bmatrix} \sigma \otimes \sigma \\ \mathbb{P}(\sigma) \end{bmatrix}. \tag{35}$$

Here, $u \otimes u$ is the symmetric endomorphism of V defined by $x \mapsto u\langle u, x\rangle$ and $\mathbb{P}(u)$ is the symmetric endomorphism of V defined by $x \mapsto uxu$. Due to the relation between symmetric endomorphisms and quadratic forms on a Euclidean space and proposition 8, this is in turn equivalent to

$$\begin{bmatrix} \mathbb{E}(r_{(2)}(U)) \\ \mathbb{E}(r_{(01)}(U)) \end{bmatrix} = M(p) \begin{bmatrix} r_{(2)}(\sigma) \\ r_{(01)}(\sigma) \end{bmatrix}. \tag{36}$$

For the case $k = 3$, Letac & Massam (1998) also proved

$$\begin{bmatrix} \mathbb{E}(r_{(3)}(U)) \\ \mathbb{E}(r_{(11)}(U)) \\ \mathbb{E}(r_{(001)}(U)) \end{bmatrix} = \begin{bmatrix} p^3 & 3p^2 & 2p \\ p^2 f & p^3 + p^2(1-f) + 2pf & 2p^2 + 2p(1-f) \\ pf^2 & 3p^2 f + 3pf(1-f) & p^3 + 3p^2(1-f) + p(2 - 3f + 2f^2) \end{bmatrix} \times \begin{bmatrix} r_{(3)}(\sigma) \\ r_{(11)}(\sigma) \\ r_{(001)}(\sigma) \end{bmatrix}. \tag{37}$$

Letac & Massam (2000, theorem 6.1) gave the following second moments of U^{-1}. Writing $q = p - rf$, they obtain

$$\begin{bmatrix} \mathbb{E}(r_{(2)}(U^{-1})) \\ \mathbb{E}(r_{(01)}(U^{-1})) \end{bmatrix} = \frac{1}{(q+f-1)(q+f-2)(q+2f-1)} \begin{bmatrix} q+2f-2 & 1 \\ f & q+f-1 \end{bmatrix} \times \begin{bmatrix} r_{(2)}(\sigma^{-1})) \\ r_{(01)}(\sigma^{-1}) \end{bmatrix}. \tag{38}$$

The method of proof for these two latter results is too complicated to be extended to higher k. Let us also mention here that a second method of computation of the $\mathbb{E}(r_{(i)}(U))$ could be extracted from Hanlon *et al.* (1992). This elegant paper provides a general formula for $\mathbb{E}(r_{(i)}(U))$. In the present paper, we have used a third method: our tools, as we have seen, are the two different bases of $P(V)_k^K$ and the eigenvectors of T_p. Results (36), (37) and (38) can be obtained using theorem 2.

5.4. *The expectation of $U^{\pm k}$ and other animals*

For each polynomial $\widetilde{r_{(i)}}(u,h) = r_{(i)}(h^{1/2}uh^{1/2})$, we can derive by a process of polarization and differentiation a quantity that we are going to denote as

$$L_{r_{(i)}}(u) = r_{(i)}(u)\sum_{j=1}^{k} ji_j \frac{u^j}{\operatorname{tr}(u^j)}. \quad (39)$$

We can think of $L_{r_{(i)}}(u)$ as 'nearly a derivative' of $r_{(i)}(u)$ with respect to u. Note that $L_{r_{(i)}}(u)$ is not a number, but an element of V. We will call it the *lifting* of $r_{(i)}(u)$. Its expectation is given in the following theorem, which allows us to compute the expectations of random variables like U^k, U^{-k} and more generally $L_{r_{(i)}}(U)$, the 'other animals' of the title. The appendix considers in general the lifting L_Q of $Q \in P(V)_k^K$: see propositions 10 and 11.

Theorem 4

Let U follow the $\gamma_{p,\sigma}$ distribution. Let $[L_{r_{(i)}}]_{I_k}$ denote the column vector of the $L_{r_{(i)}}$'s. Then

$$\text{for } p \in \Lambda, \quad [\mathbb{E}(L_{r_{(i)}}(U))]_{I_k} = B_k^{-1}D_k(p)B_k[L_{r_{(i)}}(\sigma)]_{I_k} \quad (40)$$

$$\text{for } p \geq k + (r-1)f, \quad [\mathbb{E}(L_{r_{(i)}}(U^{-1}))]_{I_k} = B_k^{-1}D_k^*(p - rf)B_k[L_{r_{(i)}}(\sigma^{-1})]_{I_k}. \quad (41)$$

In particular the last line of (40) gives $k\mathbb{E}(U^k)$ and the last line of (41) gives $k\mathbb{E}(U^{-k})$.

Proof. From propositions 2 and 12 we have

$$[T_p(L_{r_{(i)}})]_{I_k} = [L_{T_p(r_{(i)})}]_{I_k} = L_{B_k^{-1}D_k(p)B_k[r_{(i)}]_{I_k}} = B_k^{-1}D_k(p)B_k[L_{r_{(i)}}]_{I_k}.$$

This proves (40). The proof of (41) is similar. From proposition 11 it follows that for $(i) = (0,\ldots, 0, 1)$, $L_{r_{(i)}(u)} = ku^k$ and this proves the last part of the theorem.

For instance, when $k = 1$, we obtain $\mathbb{E}(U^{-1}) = \sigma^{-1}/(p - (r-1)f)$ (see Muirhead, 1982, p. 103, with a misprint). Formulae (36) and (38) combined with proposition 11 and theorem 4 give

$$\mathbb{E}(U\operatorname{tr} U) = p^2\sigma \operatorname{tr}\sigma + p\sigma^2,$$

$$\mathbb{E}(U^2) = pf\sigma\operatorname{tr}\sigma + (p - pf + p^2)\sigma^2,$$

$$\mathbb{E}(U^{-1}\operatorname{tr} U^{-1}) = \frac{1}{(q+f-1)(q+f-2)(q+2f-1)}((q+2f-2)\sigma^{-1}\operatorname{tr}\sigma^{-1} + \sigma^{-2}),$$

$$\mathbb{E}(U^{-2}) = \frac{1}{(q+f-1)(q+f-2)(q+2f-1)}(f\sigma^{-1}\operatorname{tr}\sigma^{-1} + (q+f-1)\sigma^{-2})$$

(recall that $q = p - rf$ as usual). Similarly, formula (37) gives

$$\mathbb{E}(U^3) = pf^2\sigma(\operatorname{tr}\sigma)^2 + (p^2f + pf(1-f))(\sigma\operatorname{tr}(\sigma^2) + 2\sigma^2\operatorname{tr}\sigma) + (p^3 + 3p^2(1-f) + p(2-3f+2f^2))\sigma^3.$$

Using the traditional notations $p = n/2$ and $\sigma = 2\Sigma$, for $f = 1/2$ (real case), this becomes the formula given in the introduction.

Formulae parallel to the ones obtained above, for $\mathbb{E}(U^{-3})$, $\mathbb{E}(U^4)$ and $\mathbb{E}(U^{-4})$ are given below in (44), (45), (46) and (47). Note that, as $L_{r_{(1,1)}}(u) = u \operatorname{tr} u^2 + 2u^2 \operatorname{tr} u$, formula (37) leads to

$$\mathbb{E}(U\,\mathrm{tr}\,U^2 + 2U^2\,\mathrm{tr}\,U)$$
$$= 3p^2 f\sigma(\mathrm{tr}\,\sigma)^2 + (p^3 + p^2(1-f) + 2pf)(\sigma\mathrm{tr}(\sigma^2) + 2\sigma^2\mathrm{tr}\,\sigma) + 3(2p^2 + 2p(1-f))\sigma^3,$$

but the methods of the present paper cannot yield the individual computation of $\mathbb{E}(U\ \mathrm{tr}\ U^2)$.

Note that as the multilinear forms on V^k defined by $(h_1,\ldots, h_k) \mapsto r_\pi(u)(h_1,\ldots, h_k)$ are not necessarily symmetric, there is no hope of recovering $\mathbb{E}(r_\pi(U)(h_1,\ldots, h_k))$ from the knowledge of the $h \mapsto \mathbb{E}(r_\pi(U)(h,\ldots, h))$ by a polarization process.

6. Explicit computations

6.1. *The algorithm*

The expressions given in (31) and (32) can actually be computed explicitly, for any k. Here, we use Maple in order to take advantage of the Maple package SF developed by Stembridge (1998). The package SF does a number of things. The function that we are going to be most interested in is the function 'top', which allows the passage from the basis J to the basis p.

According to the theory developed in section 4, in order to obtain the matrix A_k for a given k, it is sufficient to obtain the coefficient of $J_{\mathbf{m}}^{(1/f)}$, for all $\mathbf{m} \in M_k$ in the basis p. These coefficients will form the columns of the matrix A_k, or equivalently, the rows of the matrix $B_k = A_k^t$. The matrices B_k and B_k^{-1} are therefore obtained immediately.

In order to compute all the expectations in (31) and (32), we need the diagonal matrices D_k and D_k^*. The diagonal element of D_k corresponding to the line indexed by $\mathbf{m}$ is

$$(p)_{\mathbf{m}} = \frac{\Gamma_\Omega(p+\mathbf{m})}{\Gamma_\Omega(p)} = \prod_{j=1}^{l_{\mathbf{m}}}\prod_{s=1}^{m_j}(p+s-1-(j-1)f), \tag{42}$$

where $l_{\mathbf{m}}$ is the number of non-zero m_j's in $\mathbf{m}$. Similarly, for $q = p - rf$ and $\mathbf{m}^* = (-m_k,\ldots, -m_1)$, according to (25), the diagonal element of D_k^* corresponding to row $\mathbf{m}$ is equal to

$$(q)_{\mathbf{m}}^* = \left[\prod_{j=k-l_{\mathbf{m}}+1}^{k}\ \prod_{s=1}^{m_{k-j+1}}(q+(k-j+1)f-s)\right]^{-1}. \tag{43}$$

We can write an elementary program in Maple to obtain each diagonal element (42) and (43).

Once the matrices $B_k(f)$, $B_k^{-1}(f)$, $D_k(p)$, $D_k^*(q)$ have been obtained in this fashion, the expected values are calculated following the formulae in (31) and (32).

The algorithm to compute the matrix $M(p) = B_k^{-1}D_k(p)B_k$ representative of T_p (in the basis p) and the matrix $M^*(q) = B_k^{-1}D_k^*(q)B_k$ representative of T_p^* for each $k \in \mathbb{N}$ is therefore very simple:

Step 1. Using the function 'top' in SF, build the matrix B_k,

Step 2. Using formulae (42) and (43), build the matrices $D_k(p)$ and $D_k^*(q)$,

Step 3. Compute $M(p) = B_k^{-1}D_k(p)B_k$ and $M^*(q) = B_k^{-1}D_k^*(q)B_k$.

In the rest of this section, we are going to give the Maple commands as well as the results given by Maple for (31) and (32) in the case $k = 2$. Similar programs can be written for any k. The results obtained allow us to obtain $\mathbb{E}(U^{\pm k})$ for any k. For example, using theorem 4 and the last line of $M_3^*(q)$, as given in section 6.2, our algorithm yields immediately

$$d_3\mathbf{E}(U^{-3}) = 2f^2\sigma^{-1}(\mathrm{tr}\,\sigma^{-1})^2 + f(q+f-1)(\sigma^{-1}\mathrm{tr}\,\sigma^{-2} + 2\sigma^{-2}\mathrm{tr}\,\sigma^{-1}) + (q+f-1)^2\sigma^{-3}, \tag{44}$$

where $d_3 = (q-1+3f)(q-1+2f)(q-1+f)(q-2+f)(q-3+f)$.

In the case $d = 2f = 1$ of the real Wishart distributions, with the classical notations $p = n/2$, $\sigma = 2\Sigma$ this gives

$$C\mathbf{E}(U^{-3}) = 2\Sigma^{-1}(\mathrm{tr}\Sigma^{-1})^2 + c(\Sigma^{-1}\mathrm{tr}\Sigma^{-2} + 2\Sigma^{-2}\mathrm{tr}\Sigma^{-1}) + c^2\Sigma^{-3}, \tag{45}$$

where $c = n - r - 1$ and $C = c(c + 1)(c + 2)(c - 2)(c - 4)$.

Similarly for $k = 4$, we obtain $\mathbf{E}(U^4)$ and $\mathbf{E}(U^{-4})$, respectively, as follows:

$$\begin{aligned}\mathbf{E}(U^4) = {} & [pf^3]\sigma(\mathrm{tr}\,\sigma)^3 \\ & + [3p^2f^2 + 3(-f+1)f^2p](\sigma\,\mathrm{tr}\,\sigma\,\mathrm{tr}\,\sigma^2 + \sigma^2(\mathrm{tr}\,\sigma)^2) \\ & + [p^3f + (-3f+3)fp^2 + (2f^2 - 3f + 2)fp](\sigma\,\mathrm{tr}\,\sigma^3 + 3\sigma^3\mathrm{tr}\,\sigma) \\ & + [2p^3f + (-5f+5)fp^2 + (3f^2 - 5f + 3)fp]\sigma^2\,\mathrm{tr}\,\sigma^2 \\ & + [p^4 + (-6f+6)p^3 + (-17f + 11 + 11f^2)p^2 + (-6f^3 - 11f + 11f^2 + 6)p]\sigma^4\end{aligned} \tag{46}$$

$$\begin{aligned}d_4\mathbf{E}(U^{-4}) = {} & [5f^3q + 11(-1+f)f^3]\sigma^{-1}(\mathrm{tr}\,\sigma^{-1})^3 \\ & + [5f^2q^2 + 16(-1+f)f^2q + 11(-1+f)^2f^2](\sigma^{-1}\mathrm{tr}\,\sigma^{-1}\mathrm{tr}(\sigma^{-2}) + (\sigma^{-2})(\mathrm{tr}\,\sigma^{-1})^2) \\ & + [fq^3 + 4f(-1+f)q^2 + (3f(-1+f)^2 + f(2f^2 - 3f + 2))q \\ & \quad + f(-1+f)(2f^2 - 3f + 2)](\sigma^{-1}\mathrm{tr}(\sigma^{-1})^3 + 3(\sigma^{-1})^3\mathrm{tr}(\sigma^{-1})) \\ & + [2fq^3 + 9(f-1)fq^2 + (7(-1+f)^2 + 5f^2 - 13f + 5)fq \\ & + (-1+f)(5f^2 - 13f + 5)f]\sigma^{-2}\,\mathrm{tr}(\sigma^{-2}) \\ & + [q^4 + 5(-1+f)q^3 + (4(-1+f)^2 + 5f^2 + 5 - 9f)q^2 \\ & + ((-1+f)(5f^2 + 5 - 9f) + 2f^3 - 5f^2 + 5f - 2)q \\ & + (-1+f)(2f^3 - 5f^2 + 5f - 2)]\sigma^{-4}\end{aligned} \tag{47}$$

where $d_4 = (q - 1 + 4f)(q - 1 + 3f)(q - 2 + 2f)(q - 1 + 2f)(q - 4 + f)(q - 3 + f)(q - 2 + f)(q - 1 + f)$.

6.2. *Example: the case $k = 2$*

We first call the SF library and introduce the $J_{\mathbf{m}}^{1/f}$ polynomials as a basis.

```
> restart:with(SF):
> add_basis(J,  proc(mu)  zee(mu,1/f)  end,  proc(mu)
  hooks(mu,1/f)
> end);
```

We then build the matrix B_2.

```
> l1:=top(J[2]);
```

$$l1 := p1^2 + \frac{p2}{f}$$

```
> with(linalg):
> l12:=vector([coeff(l1,p1^2),coeff(l1,p2)]);
```

$$l12 := \left[1, \frac{1}{f}\right]$$

```
> l2:=top(J[1,1]);
```

$$l2 := -p2 + p1^2$$

```
> l22:=vector([coeff(l2,p1^2),coeff(l2,p2)]);
```

$$l22 := ([1, -1])$$

The matrix B_2 is therefore equal to

```
> B2:=matrix(2,2,[l12,l22]);
```

$$B2 := \begin{bmatrix} 1 & \frac{1}{f} \\ 1 & -1 \end{bmatrix}$$

We now obtain the matrix B_2^{-1}.

```
> IB2:=inverse(B2);
```

$$IB2 := \begin{bmatrix} \frac{f}{f+1} & \frac{1}{f+1} \\ \frac{f}{f+1} & -\frac{f}{f+1} \end{bmatrix}$$

The diagonal matrix $D_2(p)$ is easily obtained:

```
> D2:=diag(d12(p),d22(p,f));
```

$$\mathrm{D2} := \begin{bmatrix} p(p+1) & 0 \\ 0 & p(p-f) \end{bmatrix}$$

The matrix $M_2(p) = B_2^{-1}D_2(p)B_2$:

```
> M2:=simplify(multiply(IB2,D2,B2));
```

$$M2 := \begin{bmatrix} p^2 & p \\ fp & (-f+p+1)p \end{bmatrix}$$

We now write the entries of $M_2(p)$ as polynomials in decreasing powers of p.

```
> MM2:=array(1..2,1..2):for i to 2 do for j to 2 do
> MM2[i,j]:=collect(M2[i,j],[p^2,p]) od od :print (MM2);
```

$$\begin{bmatrix} p^2 & p \\ fp & p^2 + (-f+1)p \end{bmatrix},$$

which is exactly the matrix $M_2(p)$ obtained in (36). For the computation of the moments of the inverse Wishart, we now need to know the elements of $D_2^{-1}(q)$. The diagonal elements of this matrix are equal to

$$\left[\prod_{j=1}^{j=r}\prod_{i=1}^{i=m_j}(q + jf - i)\right]^{-1}$$

where $q = p - rf$.

The matrix $D_2^{-1}(q)$, denoted here `SD2` is:

```
> SD2:=diag(ds12(q,f),ds22(q,f));
```

$$SD2 := \begin{bmatrix} \frac{1}{(q+f-1)(q+f-2)} & 0 \\ 0 & \frac{1}{(q+f-1)(q+2f-1)} \end{bmatrix}$$

The matrix $M_2^*(q)$, denoted here `SM2` is therefore equal to

```
> SM2:=simplify(multiply(IB2,SD2,B2));
```

$$SM2 := \begin{bmatrix} \frac{2f+q-2}{(q+2f-1)(q+f-2)(q+f-1)} & \frac{1}{(q+f-1)(q+f-2)(q+2f-1)} \\ \frac{f}{(q+2f-1)(q+f-2)(q+f-1)} & \frac{1}{(q+2f-1)(q+f-2)} \end{bmatrix}$$

which is, of course, the matrix obtained in (38).

7. Conclusion and bibliographical comments

Our paper stands in a long series of works on the moments of the Wishart distribution. Our main contribution is twofold. First, in order to obtain the algebraic expression of the invariant moments, we use tools such as the basis of spherical polynomials and the method of 'lifting' never used before in such a context. Secondly, we give an easy to implement algorithm to compute these different moments as illustrated in section 6 with some examples. We would also like to add that a more complete version of this paper including the case of the quaternionic Wishart distribution, that is the case when the Peirce constant is $d = 4$, is available on the website of the authors (see Letac & Massam, 2003).

Let us now make some bibliographical comments. First and second moments of the Wishart and its inverse can be found in the literature: see e.g. Muirhead (1982) and Eaton (1983). Haff (1982) gives the mean, covariance of U^{-1} and also $\mathbb{E}(U^2)$ and $\mathbb{E}(U^{-2})$. Wong & Liu (1995) have considered moments for generalized Wishart distribution, including non-central Wishart distributions. Their methods are different from those used by previous authors in the sense that they use differential forms and permutations instead of matrix derivatives and commutation matrices. In that sense, their methods are closer to ours than to any other method used so far.

Hanlon *et al.* (1992) give some results which in particular cases can be reformulated with the notations of the present paper. For A and B given elements of V, A in the cone Ω and $Z = (Z_{i,j})_{1\le i,j\le r}$ a Gaussian random matrix, Hanlon *et al.* (1992) prove the existence of, and compute the numbers $c((i), (j), (l))$, such that for all (i) in I_k we have

$$\mathbb{E}(r_{(i)}(A^{1/2}ZBZ^*A^{1/2})) = \sum_{(j),(l)\in I_k} c((i),(j),(l))r_{(i)}(A)r_{(i)}(B). \tag{48}$$

When $B = e$ then $ZBZ^* = ZZ^* = (\sum_{k=1}^r Z_{ik}\bar{Z}_{jk})_{1\le i,j\le r}$ is Wishart distributed with shape parameter $p = rf$ and scale parameter e/f. Therefore, from proposition 1, the distribution of $U = A^{1/2}ZZ^*A^{1/2}$ is $\gamma_{rf,A/f}$. As $r_{(i)}(e) = r^{\sum_j i_j}$, (48) implies that

$$T_{rf}(r_{(i)}(\sigma)) = \mathbf{E}(r_{(i)}(U)) = \sum_{(j)\in I_k} c((i),(j),\cdot)r^{\sum_j i_j}r_{(i)}(\sigma)$$

where $c((i), (j), \cdot) = \sum_{(l)\in I_k} c((i), (j), (l))$. This is a particular case of our formula (31) for $p = rf$ and $d = 1, 2, 4$.

Another group of related results can be found in Graczyk *et al.* (2003) where only the complex case $d = 2$ is considered. This paper extends the results of Maiwald & Strauss (2000). For the r_π's defined here in theorem 1, the above paper by Graczyk *et al.* gives explicitly the expectation of $r_\pi(Uh_1,\ldots, Uh_k)$ as a linear combination of the $r_{\pi'}(\sigma h_1,\ldots, \sigma h_k)$ for $\pi' \in \mathcal{S}_k$ and it also gives a similar result for the expectation of $r_\pi(U^{-1}h_1,\ldots, U^{-1}h_k)$. The theory of these non-invariant moments is not well understood yet for other values of d. Letac & Massam (2001) give an analogue to theorem 1 for an arbitrary Jordan algebra V by replacing r_π by another closely related k linear form R_π on V. Interesting information about the complex case can be found in section 5 of Haagerup & Thorbjørnsen (1998).

Interestingly enough, all results found in the literature on the moments of the inverse Wishart use Stokes formula, including Maiwald & Kraus (2000) and Graczyk *et al.* (2003), while the calculations of the present paper involving the inverse Wishart (in particular proposition 6) do not.

Acknowledgements

We are most grateful to Ms Kathi Gerber from Research Computing Support at the University of Virginia for her help with Maple, particularly for installing the SF library. We also thank two anonymous referees and the editor for their useful comments. The second author is partially supported by NSERC grant A8947.

References

Anderson, T. (1984). *An introduction to multivariate statistical analysis*. Wiley, New York.

Andersson, S. (1975). Invariant linear models. *Ann. Statist*. **3**, 132–154.

Carter, E. M. (1975). Characterization and testing problems in the complex Wishart distribution. Thesis, University of Toronto, Toronto.

Casalis, M. & Letac, G. (1994). Characterisation of the Jørgensen set in generalized linear models. *Test* **3**, 145–162.

Das Gupta, S. (1968). Some aspects of discriminant function coefficients. *Sankhya Ser. A* **30**, 387–400.

Dawid, A. P. (1981). Some matrix-variate distribution theory: notational considerations and a Bayesian application. *Biometrika* **68**, 265–274.

Eaton, M. L. (1983). *Multivariate statistics. A vector space approach*. John Wiley and Sons, New York.

Faraut, J. & Koranyi, A. (1994). *Analysis on symmetric cones*. Oxford University Press, New York.

Goodman, N. R. (1963). Statistical analysis based on a certain multivariate complex Gaussian distribution. *Ann. Math. Statist*. **34**, 152–176.

Graczyk, P., Letac, G. & Massam, H. (2003). The moments of the complex Wishart distribution and the symmetric group. *Ann. Statist*. **41**, 287–309.

Haff, L. R. (1982). Identities for the inverse Wishart distribution with computational results in linear and quadratic discrimination. *Sankhya Ser. A* **44**, 245–258.

Hanlon, P., Stanley, R. & Stembridge, J. (1992). Some combinatorial aspects of the spectra of normally distributed random matrices. *Contemp. Math*. **138**, 151–174.

Haagerup, U. & Thorbjørnsen, S. (1998). Random matrices with complex Gaussian entries. Preprint 1–49.

James, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist*. **35**, 475–501.

Jensen, S. T. (1988). Covariances hypotheses which are linear in both the covariance and in the inverse covariance. *Ann. Statist*. **16**, 302–322.

Kushner, H. B. & Meisner, M. (1980). Eigenvalues of expected value operators in the Wishart distribution. *Ann. Statist*. **8**, 977–988.

Krishnamoorty, K. & Mathew, T. (1999). Comparison of approximation methods for computing tolerance factors for a multivariate normal population. *Technometrics*, **41**, 234–239.

Lassalle, M. (1992). Matrice de Hua et polynômes de Jack. *Contemp. Math*. **138**, 223–237.

Lauritzen, S. (1996). *Graphical models*. Oxford University Press, Oxford.

Letac, G. & Massam, H. (1998). Quadratic and inverse regressions for Wishart distributions. *Ann. Statist*. **26**, 573–595.

Letac, G. & Massam, H. (2000). Representations of the Wishart distributions. *Contemp. Math*. **261**, 121–142.

Letac, G. & Massam, H. (2001). Les moments des lois de Wishart, les arbres binaires et les produits triples. *Comptes R. Acad. Paris Série I* **333** (Série I), 377–382.

Letac, G. & Massam, H. (2003). All invariant moments of the Wishart distribution. http://www.lsp.ups-tlse.fr/Fp/Letac/index.html or http://www.math.yorku.ca/Who/Faculty/~massamh.

Letac, G., Massam, H. & Richards, D. (2001). An expectation formula for the multivariate Dirichlet distribution. *J. Multivariate. Anal*. **77**, 117–137.

Macdonald, I. G. (1995). *Symmetric functions and Hall polynomials*, 2nd edn. Clarendon Press, Oxford.

Massam, H. (1994). An exact decomposition theorem and a unified view of some related distributions for a class of exponential transformation models on symmetric cones. *Ann. Statist.* **22**, 369–394.
Maiwald, D. & Kraus, D. (2000). Calculation of moments of complex Wishart and complex Wishart distributed matrices. *IEE Proc. Radar Sonar Navigation* **147**, 162–168.
Muirhead, R. J. (1982). *Aspects of multivariate statistical theory*. Wiley, New York.
Shaman, P. (1980). The inverted complex Wishart distribution and its application to spectral estimation. *J. Multivariate Anal.* **10**, 51–59.
Stanley, R. P. (1989). Some combinatorial properties of Jack symmetric functions. *Adv. Math.* **77**, 76–115.
Stembridge, J. R. (1998). A Maple package for symmetric functions. Web site: http://www.math.lsa.umich.edu/jrs.
Takemura, A. (1984). *Zonal polynomials*. Institute of Mathematical Statistics, Hayward, CA.
von Rosen, D. (1988). Moments for the inverted Wishart distribution. *Scand. J. Statist.* **15**, 97–109.
von Rosen, D. (1991). Moments of maximum likelihood estimators in the growth curve model. *Statistics* **22**, 111–131.
von Rosen, D. (1997). Distribution and density approximations of a maximum likelihood estimator in the growth curve model. *Statistics* **29**, 37–59.
Wong C. S. & Liu, D. (1995). Moments of generalized Wishart distributions. *J. Multivariate Anal.* **52**, 281–294.

Received May 2001, in final form August 2003

Hélène Massam, Department of Mathematics and Statistics, York University, 4700 Keele Street, Toronto, Ontario, M3J 1P3, Canada.
E-mail: massamh@yorku.ca

Appendix

A1: Proof of proposition 1

It is well known that if U follows the Wishart $\gamma_{p,\sigma}$ distribution, then $g(U)$ follows the Wishart $\gamma_{p,g(\sigma)}$ distribution and we shall not reprove it. Then, (8) simply follows by making the change of variable $Y = g(U)$. Applying it to g in K shows that $T_p(Q)$ is K-invariant. Now, from

$$\int_V Q(u)e^{\langle\theta,u\rangle}\mu_p(du) = Q\left(\frac{\partial}{\partial\theta}\right)\int_V e^{\langle\theta,u\rangle}\mu_p(du)$$

it follows easily that if Q is a homogeneous polynomial of degree k, then so is $T_p(Q)$.

A2: Proof of proposition 7

For $s = (s_{lk})_{1\le l,k\le r}$ in Ω, denote by $\Delta_j^*(s)$ the determinant of $(s_{lk})_{r-j+1\le l,k\le r}$ and

$$\Delta_{\mathbf{m}}^*(s) = \Delta_1^*(s)^{m_1-m_2}\Delta_2^*(s)^{m_2-m_3}\cdots\Delta_r^*(s)^{m_r}.$$

From FK VII.1.5 (ii), we know that

$$\Delta_{\mathbf{m}}(s^{-1}) = \Delta_{-\mathbf{m}^*}^*(s). \tag{49}$$

We also introduce the (r, r) matrix $u = (u_{ij})_{1\le i,j\le r}$ defined by $u_{i,r+1-i} = 1$ for all $i = 1,\ldots, r$ and $u_{ij} = 0$ otherwise. Thus $usu = (s_{r+1-i,r+1-j})$, $\mathrm{tr}(\sigma^{-1}usu) = \mathrm{tr}((u\sigma u)^{-1}s)$ and

$$\Delta_{\mathbf{m}}^*(s) = \Delta_{\mathbf{m}}(usu). \tag{50}$$

Therefore, as $\Delta_r^*(s) = \Delta_r(s) = \det s$, we have

$$
\begin{aligned}
&\mathbb{E}(\Delta_{\mathbf{m}}(S^{-1}))\\
&\overset{(1)}{=}\int_\Omega \Delta_1^*(s)^{-m_r+m_{r-1}}\Delta_2^*(s)^{-m_{r-1}+m_{r-2}}\cdots\Delta_r^*(s)^{-m_1}\Delta_r^*(s)^{p-r}e^{-\mathrm{tr}(\sigma^{-1}s)}\frac{ds}{(\det\sigma)^p\Gamma_\Omega(p)}\\
&\overset{(2)}{=}\int_\Omega \Delta_1^*(s)^{p-m_r-(p-m_{r-1})}\Delta_2^*(s)^{p-m_{r-1}-(p-m_{r-2})}\cdots\Delta_r^*(s)^{p-m_1}\Delta_r^*(s)^{-r}e^{-\mathrm{tr}(\sigma^{-1}s)}\frac{ds}{(\det\sigma)^p\Gamma_\Omega(p)}\\
&\overset{(3)}{=}\int_\Omega \Delta_{p-m^*}^*(s)(\det s)^{-r}\frac{e^{-\mathrm{tr}(\sigma^{-1}s)}ds}{(\det\sigma)^p\Gamma_\Omega(p)}\\
&\overset{(4)}{=}\int_\Omega \Delta_{p-m^*}(s')(\det s')^{-r}e^{-\mathrm{tr}((u\sigma u)^{-1}s')}\frac{ds'}{(\det\sigma)^p\Gamma_\Omega(p)}\\
&\overset{(5)}{=}\Delta_{p-m^*}(u\sigma u)\frac{\Gamma(p-m^*)}{(\det\sigma)^p\Gamma_\Omega(p)}\\
&\overset{(6)}{=}\Delta_{p-m^*}^*(\sigma)\frac{\Gamma(p-m^*)}{(\det\sigma)^p\Gamma_\Omega(p)}\\
&\overset{(7)}{=}(\det\sigma)^p\Delta_{-m^*}^*(\sigma)\frac{\Gamma(p-m^*)}{(\det\sigma)^p\Gamma_\Omega(p)}\\
&\overset{(8)}{=}\frac{\Gamma(p-m^*)}{\Gamma_\Omega(p)}\Delta_{\mathbf{m}}(\sigma^{-1}).
\end{aligned}
$$

Equalities (1), (2) and (7) are due to definitions, equalities (3) and (8) are due to (49), equality (4) is obtained by the change of variables $s' = usu$ which has Jacobian 1, (5) is true by FK VII.1.2 while equality (6) is due to (50). The proof of the proposition is thus complete.

A3: Lifting

Now, let k be a non-negative integer and let $P(V)_k^K$ be the space of homogeneous polynomials $Q{:}V \mapsto \mathbb{R}$ of degree k which are K-invariant, i.e. which satisfy $Q(k(u)) = Q(u)$ for all $k \in K$ and all $u \in V$. As explained in section 4.2, these invariant polynomials Q will serve to create new polynomials $\tilde{Q} : V \times V \mapsto \mathbb{R}$. The following proposition gives the correspondence between Q and $\tilde{Q}$. As this is essentially shown in FK (p. 291), we do not give a proof. We write G for the automorphism group of Ω.

Proposition 8
Let Q be in $P(V)_k^K$. There exists a unique polynomial $\tilde{Q} : V \times V \mapsto \mathbb{R}$ such that for all h and u in Ω, one has

$$Q(h^{1/2}uh^{1/2}) = Q(u^{1/2}hu^{1/2}) = \tilde{Q}(u,h). \tag{51}$$

Furthermore, $\tilde{Q}$ is homogeneous of degree $2k$ and for all $g \in G$, $\tilde{Q}(g(u),h) = \tilde{Q}(u, g^*(h))$, where g^* is the adjoint of g.

Conversely, if $Q_1{:}V \times V \mapsto \mathbb{R}$ is a homogeneous polynomial of degree $2k$ such that, for all $g \in G$

$$Q_1(g(u),h) = Q_1(u,g^*(h)), \tag{52}$$

then, there exists a unique $Q \in P(V)_k^K$ such that $Q_1 = \tilde{Q}$, namely $Q(h) = Q_1(e,h) = Q_1(h,e)$.

To compute $\tilde{Q}$ given $Q(u) = \mathrm{tr}(u^k)$, we need to introduce the triple product $\{x, y, z\}$ of x, y and z in V, $\{x, y, z\} = \frac{1}{2}(xyz + zyx)$, which satisfies the following lemma. The proof is easy and not given here.

Lemma 1

If a is an invertible matrix of dimension r with coefficients in K_d, denote $g(x) = axa^*$ and $g^*(x) = a^*xa$ for $x \in V$. Then

$$g^*\{x, g(y), z\} = \{g^*(x), y, g^*(z)\}. \tag{53}$$

Proposition 9

Let u and h be in V. Define the sequence $(q_k)_{k\geq 1}$ in V by $q_1 = h$ and $q_k = \{h, u, q_{k-1}\}$. Then, for $Q(u) = \operatorname{tr} u^k$, we have $\tilde{Q}(u, h) = \langle u, q_k \rangle$.

Proof. We write $q_k = q_k(u, h)$ and we show by induction on k that, for all $g \in G$, one has

$$g^*(q_k(g(u), h)) = q_k(u, g^*(h)). \tag{54}$$

The result is obvious for $k = 1$. Assuming it is true for $k - 1$, we obtain (54) immediately by applying (53) to q_k. It follows from (54) that $Q_1(u,h) = \operatorname{tr}(u\, q_k(u,h))$ satisfies (52). Clearly Q_1 is a homogeneous polynomial of degree $2k$. Moreover, as $\{x, e, z\} = \frac{1}{2}(xz + zx)$, we have $q_k(e, h) = h^k$. Thus if $Q(h) = \operatorname{tr} h^k$, it follows from proposition 8 that $Q_1 = \tilde{Q}$.

With the notations of proposition 8 we have

$$\widetilde{r_{(i)}}(u, h) = \prod_{j=1}^{k} (\operatorname{tr}(uh)^j)^{i_j} \tag{55}$$

that is $\widetilde{r_{(i)}}(u, h)$ is the value of $r_\pi(u)(h,\ldots, h)$ when π has portrait (i). We mention a more compact form of theorem 1 for all $h_1,\ldots, h_k$ equal.

Theorem 5

If U has the Wishart distribution $\gamma_{p,\sigma}$ on V, then for all $h \in V$, one has

$$\mathbb{E}(\langle U, h\rangle^k) = \sum_{(i)\in I_k} a_i\, p^{i_1+\cdots+i_k} \widetilde{r_{(i)}}(\sigma, h) \tag{56}$$

where $\widetilde{r_{(i)}}$ is defined by (55) and $a_{(i)} = k!/(i_1!\cdots i_k! 1^{i_1}\cdots k^{i_k})$.

Proof. This is an immediate consequence of theorem 1, of (55) and of the fact that $a_{(i)}$ is the number of permutations π in $\mathcal{S}_k$ with portrait (i).

Let us now turn to the definition of the *lifting* $L_{r_{(i)}}(u)$ of $r_{(i)}(u)$. Given a real homogeneous polynomial Q of degree k on V, we construct its *polarized* form as the unique symmetric k-linear form $F_Q(h_1,\ldots, h_k)$ on V^k such that $Q(h) = F_Q(h,\ldots, h)$. This process is familiar for the quadratic forms and the associated symmetric bilinear forms. For instance, if $Q(h) = \operatorname{tr}(h^k)$ we have

$$F_Q(h_1, \ldots, h_k) = \frac{1}{k!} \sum_{\pi\in\mathcal{S}_k} \operatorname{tr}(h_{\pi(1)} \cdots h_{\pi(k-1)} h_{\pi(k)}). \tag{57}$$

We use the following proposition.

Proposition 10

If Q is a homogeneous polynomial on V of degree k. Then for all h_0 and h in V one has

$$F_Q(h_0, \ldots, h_0, h) = \frac{1}{k} \langle Q'(h_0), h \rangle.$$

Proof. Among the many ways to express F_Q, we choose the following: $F_Q(h_1,\ldots, h_k)$ is the coefficient of $x_1 \cdots x_k$ in the polynomial on $\mathbb{R}^k$ defined by

$$(x_1, \ldots, x_k) \mapsto \frac{1}{k!} Q(x_1 h_1 + \cdots + x_k h_k).$$

Suppose now that $h_1 = \cdots = h_{k-1} = h_0$ and $h_k = h$. Using the homogeneity of Q and its Taylor's expansion, we write

$$\begin{aligned}
&\frac{1}{k!} Q((x_1 + \cdots + x_{k-1})h_0 + x_k h) \\
&= \frac{1}{k!} (x_1 + \cdots + x_{k-1})^k Q(h_0 + \frac{x_k}{x_1 + \cdots + x_{k-1}} h) \\
&= \frac{1}{k!} (x_1 + \cdots + x_{k-1})^k [Q(h_0) + \frac{x_k}{x_1 + \cdots + x_{k-1}} \langle Q'(h_0), h \rangle + \cdots].
\end{aligned}$$

The coefficient of $x_1 \cdots x_k$ is therefore the coefficient of $x_1 \cdots x_k$ in

$$\frac{1}{k!} (x_1 + \cdots + x_{k-1})^{k-1} x_k \langle Q'(h_0), h \rangle,$$

and this gives the desired result.

When furthermore Q is a K-invariant polynomial, using the polarized form F_Q, we are going to define the *lifting* of Q. To do so, we fix u in V and consider the polynomial $h \mapsto \tilde{Q}(u, h)$ as defined in proposition 8. The lifting $L_Q(u)$ of Q is the element of V defined by

$$\langle L_Q(u), h \rangle = k F_{\tilde{Q}(u,\cdot)}(e, \ldots, e, h). \tag{58}$$

For instance, we are going to see that for $Q(u) = \text{tr}(u^k)$, then $L_Q(u) = ku^k$ and more generally we are going to calculate $L_{r_{(i)}}$.

Proposition 11

For $(i) \in I_k$ *we have*

$$L_{r_{(i)}}(u) = r_{(i)}(u) \sum_{j=1}^{k} j i_j \frac{u^j}{\text{tr}(u^j)},$$

Proof. We show first the result for $(i) = (0,\ldots, 0, 1)$. From proposition 10, this is equivalent to showing that $\partial \tilde{Q}(u, e)/\partial h = ku^k$ for $Q(h) = \text{tr}(h^k)$. To see this, we observe first that the differential of the map $x \mapsto x^k$ from V to V is

$$s \mapsto s x^{k-1} + x s x^{k-2} + x^2 s x^{k-3} + \cdots + x^{k-1} s.$$

Thus the differential of $x \mapsto \text{tr}(x^k)$ is $s \mapsto k\langle s, x^{k-1} \rangle$. Taking $h = e$ in this last formula we obtain, as desired, $\langle \frac{\partial}{\partial h} \tilde{Q}(u, e), s \rangle = k \langle u^k, s \rangle$. From this, using the chain rule, we compute the differential of $h \mapsto \tilde{Q}(u, h) = \text{tr}\, u^{1/2} h u^{1/2}$ and obtain the differential at h as

$$s \mapsto k \langle u^{1/2} s u^{1/2}, u^{1/2} h u^{1/2} \rangle.$$

Letting $h = e$ in this last formula gives $\langle \partial \tilde{Q}(u, e)/\partial h, s \rangle = k \langle u^k, s \rangle$ as desired.

Having proved the result of the proposition in this particular case leads us immediately to the general case, since from proposition 10 we have for any Q_1, Q_2 in $P(V)_k^K$

$$L_{Q_1Q_2} = Q_1L_{Q_2} + Q_2L_{Q_1}.$$

As $r_{(i)}$ is a product of functions of type $Q(h) = \mathrm{tr}(h^k)$, the general result is obtained.

In the next proposition we use in an obvious way the notations $T_p(F)$ and $T_p^*(F)$ for $\mathbb{E}(F(U))$ and $\mathbb{E}(F(U^{-1}))$, respectively, when F is not necessarily a real function of u but takes its values in a finite-dimensional real linear space.

Proposition 12

Let U follow the $\gamma_{p,\sigma}$ distribution and let Q be in $P(V)_k^K$. Then with the notation of (7) *and* (27), *respectively, we have*

$$\text{for } p \in \Lambda,\ T_p(L_Q) = L_{T_p(Q)} \tag{59}$$

$$\text{for } p \geq k + (r-1)f,\ T_p^*(L_Q) = L_{T_p^*(Q)}. \tag{60}$$

Proof. This is an immediate consequence of the linearity of $Q \mapsto L_Q$.