Journal of Multivariate Analysis **77**, 117–137 (2001)
doi:10.1006/jmva.2000.1928, available online at http://www.idealibrary.com on IDEAL®

An Expectation Formula for the Multivariate Dirichlet Distribution

Gérard Letac

Université Paul Sabatier, Toulouse, France

and

Hélène Massam and Donald Richards[1]

University of Virginia

Received April 27, 1999; published online February 5, 2001

$\mathbb{R}^q_+$ with parameter $(p_1, ..., p_q) \in \mathbb{R}^q_+$. For $f_1, ..., f_q > 0$, it is well-known that $\mathbb{E}\,(f_1 X_1 + \cdots + f_q X_q)^{-(p_1 + \cdots + p_q)} = f_1^{-p_1} \cdots f_q^{-p_q}$. In this paper, we generalize this expectation formula to the singular and non-singular multivariate Dirichlet distributions as follows. Let Ω_r denote the cone of all $r \times r$ positive-definite real symmetric matrices. For $x \in \Omega_r$ and $1 \leqslant j \leqslant r$, let $\det_j x$ denote the jth principal minor of x. For $\mathbf{s} = (s_1, ..., s_r) \in \mathbb{R}^r$, the *generalized power function* of $x \in \Omega_r$ is the function $\Delta_{\mathbf{s}}(x) = (\det_1 x)^{s_1 - s_2} (\det_2 x)^{s_2 - s_3} \cdots (\det_{r-1} x)^{s_{r-1} - s_r} (\det_r x)^{s_r}$; further, for any $t \in \mathbb{R}$, we denote by $\mathbf{s} + t$ the vector $(s_1 + t, ..., s_r + t)$. Suppose $X_1, ..., X_q \in \Omega_r$ are random matrices such that $(X_1, ..., X_q)$ follows a multivariate Dirichlet distribution with parameters $p_1, ..., p_q$. Then we evaluate the expectation $\mathbb{E}\,[\Delta_{\mathbf{s}_1}(X_1) \cdots \Delta_{\mathbf{s}_q}(X_q)\, \Delta_{\mathbf{s}_1 + \cdots + \mathbf{s}_q + p}((a + f_1 X_1 + \cdots + f_q X_q)^{-1})]$, where $a \in \Omega_r$, $p = p_1 + \cdots + p_q$, $f_1, ..., f_q > 0$, and $\mathbf{s}_1, ..., \mathbf{s}_q$ each belong to an appropriate subset of $\mathbb{R}^r_+$. The result obtained is parallel to that given above for the univariate case, and remains valid even if some of the X_j's are singular. Our derivation utilizes the framework of symmetric cones, so that our results are valid for multivariate Dirichlet distributions on all symmetric cones. © 2001 Academic Press

AMS 1991 subject classifications: 60E05; 62H05; 62E10.

Key words and phrases: Dirichlet distribution; gamma distribution; Gaussian hypergeometric function; generalized power function; Jordan algebra; Laplace transform; Lauricella function; multivariate beta distribution; multivariate gamma function; Riesz measure; symmetric cone; Wishart distribution.

1. INTRODUCTION

Let r be a positive integer and V_r denote the space of real symmetric $r \times r$ matrices; thus, V_r is a vector space of dimension $n = r(r+1)/2$. Denote by

[1] Research supported in part by the National Science Foundation under Grant DMS-9703705.

0047-259X/01 $35.00

Ω_r the cone of positive-definite elements of V_r, and denote by e_r the $r \times r$ identity matrix. Let $q \geqslant 2$ be an integer, and let $\mathbf{p} = (p_1, ..., p_q) \in \mathbb{R}^q$ be such that $p_i > (r-1)/2$ for all $i = 1, ..., q$.

Let $X_1, ..., X_q$ be random matrices in V_r such that the probability distribution of $(X_1, ..., X_q)$ is concentrated on the set

$$T_q = \{(x_1, ..., x_q) \in V_r^q : x_1, ..., x_q \in \Omega_r \text{ and } x_1 + \cdots + x_q = e_r\}.$$

Then $(X_1, ..., X_q)$ is said to follow the multivariate Dirichlet distribution $D_{\mathbf{p}}$ on V_r^q if the probability density function of $(X_1, ..., X_{q-1})$ exists and is given by

$$K_{\mathbf{p}} (\det x_1)^{p_1 - (n/r)} \cdots (\det x_{q-1})^{p_{q-1} - (n/r)} (\det(e_r - x_1 - \cdots - x_{q-1}))^{p_q - (n/r)},$$

where $x_1, ..., x_{q-1}, e_r - x_1 - \cdots - x_{q-1} \in \Omega_r$ and $K_{\mathbf{p}}$ is the normalizing constant.

For any matrix $x \in V_r$ and $1 \leqslant j \leqslant r$, we denote by $\det_j x$ the jth *principal minor* of x. Further, for $\mathbf{s} = (s_1, ..., s_r) \in \mathbb{R}^r$, the *generalized power function of* $x \in \Omega_r$ is defined to be

$$\Delta_{\mathbf{s}}(x) = (\det_1 x)^{s_1 - s_2} (\det_2 x)^{s_2 - s_3} \cdots (\det_{r-1} x)^{s_{r-1} - s_r} (\det_r x)^{s_r}. \tag{1.1}$$

For the special case in which x is the diagonal matrix $\operatorname{diag}(\lambda_1, ..., \lambda_r)$, it is not difficult to see that $\Delta_{\mathbf{s}}(x)$ has the simple form $\lambda_1^{s_1} \cdots \lambda_r^{s_r}$.

We shall adopt the convention that if $\mathbf{s} = (s_1, ..., s_r) \in \mathbb{R}^r$ and $t \in \mathbb{R}$ then $\mathbf{s} + t \equiv (s_1 + t, ..., s_r + t)$. Let $I_r = \{(z_1, ..., z_r) \in \mathbb{R}^r : z_j > (j-1)/2, j = 1, ..., r\}$. For $\mathbf{z} = (z_1, ..., z_r) \in I_r$, the *multivariate gamma function* Γ_r is defined as

$$\Gamma_r(\mathbf{z}) = (2\pi)^{(n-r)/2} \prod_{j=1}^{r} \Gamma(z_j - (j-1)/2).$$

For the special case in which $z_1 = \cdots = z_r$, we shall denote $\Gamma_r(\mathbf{z})$ by $\Gamma_r(z_1)$; in this case, $\Gamma_r(z_1)$ is the well-known multivariate gamma function which arises naturally in classical multivariate statistical analysis (cf. Muirhead (1982, p. 61, section 2.1.2)).

Suppose that $(X_1, ..., X_q)$ follows the Dirichlet distribution $D_{\mathbf{p}}$, $\mathbf{s}_1, ..., \mathbf{s}_q \in \mathbb{R}_+^r$, $a \in \Omega_r$, and $f_1, ..., f_q > 0$. Let $\mathbf{s} = \mathbf{s}_1 + \cdots + \mathbf{s}_q$ and $p = p_1 + \cdots + p_q$. In this paper, under appropriate conditions on $\mathbf{s}_1, ..., \mathbf{s}_q$, and for $p_1, ..., p_q > (r-1)/2$, we establish the expectation formula

$$\begin{aligned} &\mathbb{E}\, [\Delta_{\mathbf{s}_1}(X_1) \cdots \Delta_{\mathbf{s}_q}(X_q)\, \Delta_{\mathbf{s}+p}((a + f_1 X_1 + \cdots + f_q X_q)^{-1})] \\ &\quad = C_{\mathbf{s}, \mathbf{p}} \prod_{i=1}^{q} \Delta_{\mathbf{s}_i + p_i}((a + f_i e_r)^{-1}), \end{aligned} \tag{1.2}$$

where

$$C_{\mathbf{s},\mathbf{p}} = \frac{\Gamma_r(p)}{\Gamma_r(\mathbf{s}+p)} \prod_{i=1}^{q} \frac{\Gamma_r(\mathbf{s}_i+p_i)}{\Gamma_r(p_i)}. \tag{1.3}$$

In fact, we will prove in the main result (Theorem 4.1) that (1.2) is valid under conditions on $p_1, ..., p_q$ more general than those given above.

Some particular cases of (1.2) can be found in the literature. Suppose $r=1$, in which case the one-dimensional random variables $X_1, ..., X_q$ follow a classical Dirichlet distribution. If we set $\mathbf{s}_1 = \cdots = \mathbf{s}_q = 0$ and let $a \to 0+$, then (1.2) becomes

$$\mathbb{E}\,(f_1 X_1 + \cdots + f_q X_q)^{-(p_1 + \cdots + p_q)} = f_1^{-p_1} \cdots f_q^{-p_q}. \tag{1.4}$$

The formula (1.4) is known to form the basis of a characterization of the Dirichlet distribution. Mauldon (1959), in a remarkable article which appears to have been widely overlooked since its appearance, was first to utilize (1.4) as the basis for a characterization of the Dirichlet distributions and to study a more general class of distributions. Other applications of (1.4), and references thereof, are given by Karlin, Micchelli and Rinott (1986), Chamayou and Letac (1994), Letac and Scarsini (1998) and Gupta and Richards (2000).

In the general case, $r \geqslant 1$, suppose we set $a = \operatorname{diag}(\lambda_1, ..., \lambda_r)$ and $\mathbf{s}_i = (s_{i1}, ..., s_{ir})$ for $i = 1, ..., q$. Then (1.2) reduces to

$$\begin{aligned}
&\mathbb{E}\,[\Delta_{\mathbf{s}_1}(X_1) \cdots \Delta_{\mathbf{s}_q}(X_q)\, \Delta_{\mathbf{s}_1 + \cdots + \mathbf{s}_q + p_1 + \cdots + p_q}((a + f_1 X_1 + \cdots + f_q X_q)^{-1})] \\
&\quad = \Delta_{\mathbf{s}_1 + p_1}((a + f_1\, e)^{-1}) \cdots \Delta_{\mathbf{s}_q + p_q}((a + f_q e)^{-1}) \\
&\quad = \prod_{i=1}^{q} \prod_{j=1}^{r} (\lambda_i + f_j)^{-s_{ij} - p_i},
\end{aligned} \tag{1.5}$$

which is close in form to (1.4). If, further, we set $s_{ij} = 0$ and let $\lambda_j \to 0+$ for all i and j then we obtain from (1.5) the result

$$\mathbb{E}\,\det(f_1 X_1 + \cdots + f_q X_q)^{-(p_1 + \cdots + p_q)} = f_1^{-rp_1} \cdots f_q^{-rp_q}. \tag{1.6}$$

This expectation formula for the random matrices $X_1, ..., X_q$ was proved by Letac and Massam (1998b) and it generalizes (1.4).

For the case in which $q=2$, Letac and Massam (1998b) also noted that (1.6) can be extended to the situation in which the random matrices X_1, X_2 may be singular. This result can be proved by applying a detailed analysis of the eigenvalues of a beta-distributed random matrix. Indeed, the argument requires an application of a celebrated theorem which was proved separately in 1939 by five different authors: see Muirhead (1982, Theorem 3.3.4, p. 112), and Anderson (1996). For $q > 2$, the extension of (1.6) to the case in which the Dirichlet random variables are singular was left open; the method of proof for $q = 2$ relies on the computation of Jacobians and densities and appears to be difficult to extend to the case in which $q > 2$.

We remarked earlier that the aim of this paper is to establish the expectation formula (1.2). We consider our work as a step toward a study of characterizations of the multivariate Dirichlet distribution, analogous to the characterization given in the univariate case by Mauldon (1959). We shall establish the identity (1.2), in both the singular and non-singular cases, with methods different from those used by Letac and Massam (1998b). Our derivation also utilizes the framework of symmetric cones, so that the final results are valid for multivariate Dirichlet distributions on cones more general than Ω_r.

As motivation for our proof of (1.2) in the general case, we will now give a proof of (1.4). Our proof of (1.2) will be based upon a generalization of this idea. Let us denote the classical gamma distribution on $\mathbb{R}$, with shape parameter $p > 0$ and scale parameter 1, by

$$\gamma_p(dy) = \frac{1}{\Gamma(p)}\, y^{p-1} \exp(-y)\, \mathbf{1}_{(0,\,\infty)}(y)\, dy,$$

where, for a given set A, $\mathbf{1}_A$ denotes the corresponding indicator function.

Let $Y_1, ..., Y_q$ be a sequence of independent random variables with respective distributions $\gamma_{p_1}, ..., \gamma_{p_q}$. For $t_1, ..., t_q > -1$, the moment-generating function of $Y_1, ..., Y_q$ is well-known to be

$$M(t_1, ..., t_q) := \mathbb{E} \exp(-(t_1 Y_1 + \cdots + t_q Y_q)) = (1+t_1)^{-p_1} \cdots (1+t_q)^{-p_q}. \tag{1.7}$$

Set $S = Y_1 + \cdots + Y_q$ and

$$(X_1, ..., X_q) = \left(\frac{Y_1}{S}, ..., \frac{Y_q}{S}\right). \tag{1.8}$$

It is well-known that the random variables S and $(X_1, ..., X_q)$ are mutually independent. Moreover, $(X_1, ..., X_q)$ has a Dirichlet distribution with parameters $(p_1, ..., p_q)$, and S has a gamma distribution γ_p with $p = p_1 + \cdots + p_q$. Thus, another way to evaluate the moment-generating function M is to introduce the decomposition $Y = S(X_1, ..., X_q)$, evaluate a conditional expectation with respect to $(X_1, ..., X_q)$, and then apply the independence of X and S. Then we obtain

$$M(t_1, ..., t_q) = \mathbb{E}_{X_1, ..., X_q} \mathbb{E}_S [\exp(-(t_1 X_1 + \cdots + t_q X_q) S) \mid X_1, ..., X_q]$$
$$= \mathbb{E}(1 + t_1 X_1 + \cdots + t_q X_q)^{-p}. \tag{1.9}$$

Replacing each t_j by $f_j - 1$, noting that $X_1 + \cdots + X_q = 1$, and comparing the expressions (1.7) and (1.9) for $M(t_1, ..., t_q)$, we then obtain the result (1.4).

In Section 2, we provide some preliminary material on symmetric cones. This material, which is abstracted from Faraut and Korányi (1994) (henceforth abbreviated F–K) has been presented here so as to make the paper self-contained. Thus, we provide simple guidelines to enable our readers to translate standard symmetric cone notation into traditional matrix notation. In Section 3, we assemble some basic facts about the Wishart and Dirichlet distributions associated with the symmetric cones. Further, we establish a crucial auxiliary result, Theorem 3.6, needed for the proof of the main theorem. In Section 4, we state and prove formula (1.2), in complete generality, i.e., in both the non-singular and singular cases within the context of symmetric cones. We also mention that the fundamental basis on which our proof rests is that the cone Ω_r of $r \times r$ positive-definite symmetric matrices, and more generally the symmetric cone associated with a simple Jordan algebra, has the structure of a Gelfand pair; however we will not elaborate on this point here. Finally, in Section 5, we provide a connection between our results and the theory of multiple hypergeometric functions.

2. PRELIMINARIES ON SYMMETRIC CONES

We shall begin with a brief review of the structure of symmetric cones, providing those properties that are needed for the results that follow. For a presentation in more detail we refer to F-K.

The prototypical example of a symmetric cone is Ω_r, the cone of real, positive-definite, symmetric $r \times r$ matrices. The cone Ω_r is *irreducible* because it cannot be decomposed into a direct product of non-zero symmetric cones. We shall refer to this example as the *classical case*. The Euclidean Jordan algebra associated with Ω_r is V_r, the linear space of $r \times r$ real symmetric matrices; the algebra V_r also is *simple* because it contains no proper ideals. The space V_r is endowed with an inner product $\langle \cdot, \cdot \rangle$, given by $\langle x, y\rangle := \mathrm{tr}(xy)$ for all $x, y \in V_r$. Moreover, V_r is a commutative algebra under the *Jordan product* $x \circ y := \frac{1}{2}(x \cdot y + y \cdot x)$, where $x \cdot y$ denotes the standard matrix product. The Jordan product also satisfies a fundamental identity, called the *Jordan identity* (cf. F-K (p. 24)); however, we will not make explicit use of that identity in this paper.

There are five types of irreducible symmetric cones. The cones of self-adjoint matrices with entries in $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{H}$ are the most commonly encountered examples, each of which arises naturally in multivariate statistical analysis (cf. Andersson (1975), Andersson *et al.* (1983), Massam (1994), Casalis and Letac (1996)). The remaining two types are the Lorentz cone (cf. Jensen, 1988)) and a cone of 3×3 matrices with entries in the octonions or Cayley numbers (denoted by $\mathbb{O}$). We will generally adopt the notations for symmetric cones as prescribed by F-K.

Turning to the general context, let V be a simple Euclidean Jordan algebra. In particular, V is a commutative algebra over $\mathbb{R}$; and V also is a finite-dimensional Euclidean space, of dimension n, say. The space V is equipped with an inner product, which we denote by $\langle x, y\rangle$ (rather than the standard notation $(x \mid y)$, as used by F-K). The product of two elements $x, y \in V$ is denoted by $x \circ y$; thus the product may be viewed as a map from $V \times V$ into V such that $(x, y) \mapsto x \circ y$. In particular, we denote $x \circ x$, the square of $x \in V$, by x^2.

Let Ω be the *interior* of the set $\{x^2 : x \in V\}$, the cone of squares in V; then the space Ω is an *irreducible symmetric cone*. Moreover, any irreducible symmetric cone is isomorphic to a cone of this type (cf. F-K (p. 49, Theorem III.3.1)), so that the classification of the irreducible symmetric cones reduces to the classification of the simple Euclidean Jordan algebras. Therefore, to work within the framework of symmetric cones is equivalent to working within the framework of Jordan algebras.

Let $GL(V)$ denote the general linear group of invertible linear transformations on V, and denote by $O(V)$ the subgroup of $GL(V)$ containing all orthogonal linear transformations. We let G be the connected component of the subgroup of $GL(V)$ which preserves Ω; then, G contains the identity element in $GL(V)$. Further, we denote by $K = O(V) \cap G$ the orthogonal subgroup of G.

For each $x \in V$ we define the *regular representation* $L(x)\colon V \to V$ by $L(x)(y) = x \circ y$, $y \in V$. For each $x \in V$, the *trace* of x is $\mathrm{tr}(x) := \langle x, e\rangle$ and the inner product on V is given by $\langle x, y\rangle := \mathrm{tr}(x \circ y)$. Following F-K (p. 29), for any $x \in V$, we also denote by $\det(x)$ the *determinant* of x, which may be defined explicitly in terms of the coefficients of the characteristic polynomial of the linear transformation $L(x)$.

The map $\mathbb{P}(x)\colon V \to V$, defined by

$$\mathbb{P}(x) = 2L(x)^2 - L(x^2) \tag{2.1}$$

is called the *quadratic representation* of V (cf. F-K, p. 32) because it satisfies the identity

$$\mathbb{P}(\mathbb{P}(x)\,y) = \mathbb{P}(x)\,\mathbb{P}(y)\,\mathbb{P}(x).$$

Moreover, it can be deduced from the definition of $\mathbb{P}$ that

$$\operatorname{tr}(\mathbb{P}(y)\,x^2) = \operatorname{tr}(\mathbb{P}(x)\,y^2) = \langle x^2, y^2 \rangle \tag{2.2}$$

for all x and y in V.

An element $c \in V$ is *idempotent* if $c^2 = v$. A scalar β is an *eigenvalue* of $c \in V$ if there exists a nonzero $x \in V$ such that $c \circ x = \beta x$. If c is idempotent then it can be shown that its eigenvalues must be equal to 1, 1/2 or 0 (cf. F-K, p. 62). An idempotent c is *primitive* if it is nonzero and is not expressible as the sum of two nonzero idempotents. Two idempotents c_1 and c_2 are *orthogonal* if $c_1 \circ c_2 = 0$. A maximal system of orthogonal primitive idempotents is called a *Jordan frame*. It may be shown that any Jordan frame has the same number, r, of elements; and r is called the *rank* of Ω. If $\{c_1, ..., c_r\}$ is a Jordan frame, then $c_1 + \cdots + c_r = e$, the identity element in V. By constructing a Jordan frame we are simply choosing a basis for the vector space V.

Let us choose and fix a Jordan frame $\{c_1, ..., c_r\}$ in V and define a collection of subspaces, $V_j = \{x \in V\colon c_j \circ x = x\}$ and $V_{ij} = \{x \in V\colon c_i \circ x = \frac{1}{2}x$ and $c_j \circ x = \frac{1}{2}x\}$, $i, j = 1, ..., r$. Each V_j, for $j = 1, ..., r$, is a one-dimensional subalgebra. Further, the subspaces V_{ij}, for $i, j = 1, ..., r$ with $i \neq j$, all have a common dimension, called the *Peirce constant*, denoted by d. The constant d is independent of the choice of Jordan frame; for convenience, we denote $d/2$ by d'. It may be shown that n, d and r are related by the formula

$$n = r + d'r(r-1).$$

In the classical case of V_r, the space of $r \times r$ real symmetric matrices, all of these concepts are familiar. In this case, $d = 1$ and $n = r(r+1)/2$; the identity element $e \in V$ is e_r, the usual $r \times r$ identity matrix. Further, Ω is the cone of $r \times r$ positive-definite symmetric matrices, and its closure $\bar{\Omega}$ is the cone of positive $r \times r$ positive semi-definite symmetric matrices. The trace and determinant functions on V reduce to the classical trace and determinant functions, respectively, on V_r; and a natural Jordan frame for V_r is obtained by choosing c_j as the $r \times r$ matrix whose (j, j)-th entry is 1 and all other entries are zero. Next, for each $x \in V_r$, the linear map $\mathbb{P}(x)\colon V_r \to V_r$, defined in (2.1), is given by $\mathbb{P}(x)(y) = x \cdot y \cdot x$, $y \in V_r$, and (2.2) corresponds to the simple formula,

$$\mathrm{tr}(y \cdot x \cdot x \cdot y) = \mathrm{tr}(x \cdot x \cdot y \cdot y).$$

For any $r \times r$ real matrix a, denote by a^* the transpose of a. Then G, the connected component of the identity in $GL(V_r)$, is the group of linear maps $g_a : V_r \to V_r$ such that $g_a(x) = a \cdot x \cdot a^*$, $x \in V_r$, where a is invertible. If a is an orthogonal $r \times r$ matrix then $g_a \in K$; and conversely, every element of K is of this form.

In the case of the algebras of complex and quaternionic matrices of order r, the values of d are 2 and 4, respectively. In the case of the Lorentz algebra of dimension n, $r = 2$ and $d = n - 2$. In the case of the Albert algebra, which corresponds to the cone of 3×3 matrices with entries in the octonions, $n = 27$, $r = 3$ and $d = 8$. We refer the reader to a summary listing (cf. F-K, page 97) of all five types of Jordan algebras, their associated symmetric cones, ranks and Peirce constants.

Returning to the general context, define the set

$$I_\Omega = \{(z_1, ..., z_r) \in \mathbb{R}^r : z_j > (j-1)\, d',\ j = 1, ..., r\}.$$

Then the *multivariate gamma function* for the cone Ω, denoted Γ_Ω, is defined on the domain I_Ω by

$$\Gamma_\Omega(\mathbf{z}) = (2\pi)^{(n-r)/2} \prod_{j=1}^{r} \Gamma(z_j - (j-1)\, d'),$$

where $\mathbf{z} = (z_1, ..., z_r) \in I_\Omega$. For cases in which $z_1 = \cdots = z_r$, we denote $\Gamma_\Omega(\mathbf{z})$ by $\Gamma_\Omega(z_1)$. As before, we retain the convention that, for any $\mathbf{s} = (s_1, ..., s_r) \in \mathbb{R}^r$ and $t \in \mathbb{R}$, $\mathbf{s} + t \equiv (s_1 + t, ..., s_r + t)$.

3. THE WISHART AND DIRICHLET DISTRIBUTIONS

3.1. *The Wishart Distributions on* **V**. The *Gindikin set* is defined to be the set of real numbers

$$\Lambda = \{d', 2d', ..., d'(r-1)\} \cup (d'(r-1), \infty) \tag{3.1}$$

(cf. F-K, p. 137). For $p \in \Lambda$ it is known (cf. F-K, p. 123) that there exists a positive measure μ_p on $\bar{\Omega}$ such that the Laplace transform of μ_p exists for $-\theta \in \Omega$ and equals

$$\int_{\bar{\Omega}} \exp(\langle \theta, x \rangle)\, \mu_p(dx) = (\det\ -\theta)^{-p}. \tag{3.2}$$

For $p > d'(r-1)$, the measure μ_p has a density with respect to the Lebesgue measure dx on the cone Ω. Then we have

$$\mu_p(dx) = \frac{1}{\Gamma_\Omega(p)} (\det\ x)^{p-(n/r)} \mathbf{1}_\Omega(x)\, dx,$$

a result which follows directly from F-K (p. 123, Theorem VII.1.1). If $p = jd'$ where j is an integer, $1 \leqslant j \leqslant r-1$, then μ_p is a singular measure which is concentrated on the set of elements of rank j of $\bar{\Omega}$.

For $p \in \Lambda$ and $\sigma \in \Omega$, the Laplace transform (3.2) implies that the measure $\gamma_{p,\sigma}$ on $\bar{\Omega}$, defined by

$$\gamma_{p,\sigma}(dx) = (\det \sigma)^{-p} \exp(-\langle x, \sigma^{-1} \rangle)\, \mu_p(dx),$$

is a probability distribution. This distribution is called the *Wishart distribution on V*, with shape parameter p and scale parameter σ. The Laplace transform of $\gamma_{p,\sigma}$ exists for $\sigma^{-1} - \theta \in \Omega$ and is equal to

$$\int_{\bar{\Omega}} \exp(\langle \theta, x \rangle)\, \gamma_{p,\sigma}(dx) = [\det (e - \mathbb{P}(\sigma^{1/2})(\theta))]^{-p}. \tag{3.3}$$

Further details of the Wishart distribution on Ω may be obtained from Artzner and Fourt (1974), Fourt (1974), Massam (1994), Gupta and Richards (1995), Casalis and Letac (1996) and Letac and Massam (1998a).

In the classical case of Ω_r, the Wishart distribution is usually denoted by $W_r(m, \Sigma)$. The correspondence between the two notations is given by $W_r(m, \Sigma) \equiv \gamma_{p,\sigma}$ with $m = 2p$ and $\Sigma = \sigma/2$. Then $\langle x, \sigma^{-1} \rangle = \operatorname{tr}(x\Sigma^{-1})/2$, and the Laplace transform (3.3) reduces to the familiar formula $[\det (e_r - 2\Sigma\theta)]^{-m/2}$, $\frac{1}{2}\Sigma^{-1} - \theta \in \Omega_r$ (cf. Muirhead (1982, p. 87)). For purposes of applications, interest has focused traditionally on integer values of m; however, in this article there is no need for restriction to integer values of m so that p is free to take any value in Λ.

3.2. *The Dirichlet distributions on* **V**. In order to study the singular case, it is necessary to present a nontraditional description of the Dirichlet distributions. We remark that, even in the classical matrix case, there exist singular versions of the Dirichlet distributions, and we shall pay special attention to them throughout the present paper. We proceed toward the definition of the Dirichlet distributions as follows.

Let q be an integer, $q \geqslant 2$. Let $p_1, ..., p_q \in \Lambda$, the Gindikin set defined in (3.1), be such that $p := p_1 + \cdots + p_q > d'(r-1)$. For $\sigma \in \Omega$, let $Y_1, ..., Y_q$ be mutually independent random variables in V with Wishart distributions $\gamma_{p_1,\sigma}, ..., \gamma_{p_q,\sigma}$, respectively, and set $S = Y_1 + \cdots + Y_q$. Then, using the

Laplace transform (3.3), we see immediately that S has the Wishart distribution $\gamma_{p,\sigma}$. Moreover, since $p > d'(r-1)$ then, with probability one, the distribution of S is concentrated on Ω and is invertible.

We define the random variable $(X_1, ..., X_q)$ taking values in V^q by

$$(X_1, ..., X_q) = (\mathbb{P}(S^{-1/2})(Y_1), ..., \mathbb{P}(S^{-1/2})(Y_q)). \tag{3.4}$$

It is not difficult to see that (3.4) is a natural generalization to the symmetric cone setting of the univariate transformation (1.8). The distribution of $(X_1, ..., X_q)$ in (3.4) is called the *Dirichlet distribution on V with parameter* $\mathbf{p} = (p_1, ..., p_q)$, and is denoted by $D_{\mathbf{p}}$. This distribution was studied in the absolutely continuous (or nonsingular) case by Artzner and Fourt (1974) and by Massam (1994, Theorems 4.1 and 4.2). In the general case, the distribution was studied by Casalis and Letac (1996, p. 774). In those papers, it is proved that, analogous to the one-dimensional setting, the random variables $(X_1, ..., X_q)$ and S are independent and the distribution of $(X_1, ..., X_q)$ does not depend on the parameter σ; this result can be established by application of Basu's theorem in the nonsingular case. In the general case, however, it is simpler to prove directly the independence of $(X_1, ..., X_q)$ and S, as is done by Casalis and Letac (1996, Theorem 3.1), than to prove that the conditions of Basu's theorem are satisfied.

Next, we define the set

$$\overline{T_q} := \{(x_1, ..., x_q) \in V^q : x_i \in \bar{\Omega}, x_1 + \cdots + x_q = e\}. \tag{3.5}$$

Since $\mathbb{P}(S^{-1/2})(S) = e$ then, clearly, $(X_1, ..., X_q) \in \overline{T_q}$.

Let $B_\Omega(\mathbf{p})$ be the beta function for the cone Ω, defined by

$$B_\Omega(\mathbf{p}) = \frac{\Gamma_\Omega(p_1) \cdots \Gamma_\Omega(p_q)}{\Gamma_\Omega(p_1 + \cdots + p_q)}.$$

Suppose that $p_i > d'(r-1)$, $i = 1, ..., q$. Then Massam (1994) proved that the image of $D_{\mathbf{p}}$ under the projection of $\overline{T_q}$ on V^{q-1}, defined by $(x_1, ..., x_q) \mapsto (x_1, ..., x_{q-1})$, has the density

$$\frac{1}{B_\Omega(\mathbf{p})} (\det x_1)^{p_1 - (n/r)} \cdots (\det x_q)^{p_q - (n/r)}, \tag{3.6}$$

where $x_q = e - x_1 - \cdots - x_{q-1}$. Simply put, if $p_i > d'(r-1)$ for all $i = 1, ..., q$, then the random variable $(X_1, ..., X_{q-1})$ has a density function which is given by (3.6).

In situations in which $p_i \leqslant d'(r-1)$, the corresponding X_i are singular, and then explicit expressions for the law $D_{\mathbf{p}}$ are complicated (cf. Uhlig (1994), Díaz-Garcia and Gutiérrez Jáimez (1997)) and so cannot be applied here. To treat both the singular and non-singular cases, we now introduce the generalized Pochhammer symbol (or rising factorial).

3.3. *The generalized Pochhammer symbol.* For $s \geqslant 0$ and $p > 0$, the classical Pochhammer symbol $(p)_s$ may be defined as

$$(p)_s = \frac{\Gamma(s+p)}{\Gamma(p)}.$$

Using the multivariate gamma function, we extend this definition as follows.

Let $\mathbf{s} = (s_1, ..., s_r) \in \mathbb{R}^r_+$ and let $p \in \Lambda$, the Gindikin set. For $p > d'(r-1)$, the *generalized Pochhammer symbol* is defined to be

$$(p)_{\mathbf{s}} := \frac{\Gamma_\Omega(\mathbf{s}+p)}{\Gamma_\Omega(p)} = \prod_{j=1}^{r} \frac{\Gamma(s_j + p - (j-1)\, d')}{\Gamma(p - (j-1)\, d')}.$$

As a special case, if $\mathbf{s} = (s_1, ..., s_{j_0 - 1}, 0, ..., 0)$ and $p = (j_0 - 1)\, d'$ for some $j_0 \in \{1, ..., r\}$, then

$$(p)_{\mathbf{s}} = \prod_{j=1}^{j_0 - 1} \frac{\Gamma(s_j + (j_0 - j)\, d')}{\Gamma((j_0 - j)\, d')}.$$

The generalized Pochhammer symbol is also defined in F-K (p. 129 and 230), for the case in which $s_1, ..., s_r$ is a non-increasing sequence of non-negative integers and p is a complex number. In our definition, the pair $(p, \mathbf{s})$ belongs to a different domain; however, our definition is compatible with the definition in F-K whenever the two domains coincide.

We shall use the Pochhammer symbol in several ways. For now, let us observe that it can be used to simplify an expression such as (1.2) by writing the normalizing constant $C_{\mathbf{s},\mathbf{p}}$ in (1.3) in a simpler manner as

$$C_{\mathbf{s},\mathbf{p}} = \frac{(p)_{\mathbf{s}_1} \cdots (p)_{\mathbf{s}_q}}{(p)_{\mathbf{s}}} \tag{3.7}$$

when $p_j > d'(r-1)$. However the main reason for introducing the Pochhammer symbol is to provide a meaning for (3.7) above even in case some p_j belong to the singular part, $\{d', 2d', ..., (r-1)\, d'\}$, of the Gindikin set Λ. For in that case, the constant $C_{\mathbf{s},\mathbf{p}}$ in (1.2) and (1.3) is undefined.

3.4. *The generalized power functions* $\mathbf{\Delta}_{\mathbf{s}}$. Following F-K (p. 122), we first choose a Jordan frame $\{c_1, ..., c_r\}$ of V. Clearly, for each $j = 1, ..., r$, $c_1 + \cdots + c_j$ is an idempotent in V. Now define

$$V(c_1 + \cdots + c_j, 1) := \{x \in V : (c_1 + \cdots + c_j) \circ x = x\},$$

$j=1, ..., r$; then $V(c_1+\cdots+c_j, 1)$ is a subalgebra of V of rank j. We denote by $\Delta_j(x)$ the determinant of x in the Jordan subalgebra $V(c_1+\cdots+c_j, 1)$, as defined in F-K (p. 122). For $\mathbf{s}=(s_1, ..., s_r)\in\mathbb{R}^r$, we define the *generalized power function*

$$\Delta_{\mathbf{s}}(x)=(\Delta_1(x))^{s_1-s_2}(\Delta_2(x))^{s_2-s_3}\cdots(\Delta_{r-1}(x))^{s_{r-1}-s_r}(\Delta_r(x))^{s_r}.$$

In case $\Delta_j(x)=0$, we adopt the usual convention that $(\Delta_j(x))^0=1$.

Remark 3.5. In the classical case of the matrix cone over $\mathbb{R}$, the formula which represents $\Delta_{\mathbf{s}}(x)$ as a product of powers of principal minors is given in (1.1). In the case of the matrix cones over $\mathbb{C}$, $\mathbb{H}$ or $\mathbb{O}$, expressions for the determinant, principal minors, and generalized power function $\Delta_{\mathbf{s}}$ are well-known and are similar to (1.1). The Lorentz cone, however, is less familiar to readers, so we shall provide an explicit description of the principal minors for that cone.

For $n\geqslant 2$, the Lorentz cone is the space

$$L_n=\{(x_1, ..., x_n)\in\mathbb{R}^n: x_1^2-x_2^2-\cdots-x_n^2>0, x_1>0\}.$$

For $x=(x_1, ..., x_n)\in L_n$, the first principal minor on L_n is the function $\det_1(x):=x_1^2$; the second principal minor, or determinant, is the function $\det_2(x):=x_1^2-x_2^2-\cdots-x_n^2$.

For $\mathbf{s}=(s_1, s_2)\in\mathbb{R}^2$ and $x=(x_1, ..., x_n)\in L_n$, it now follows that the generalized power function is $\Delta_{\mathbf{s}}(x)=x_1^{2(s_1-s_2)}(x_1^2-x_2^2-\cdots-x_n^2)^{s_2}$.

We are now in position to formulate a crucial result.

THEOREM 3.6. *Let $p\in\Lambda$ and μ_p be the measure defined by* (3.2). *For $\mathbf{s}\in\mathbb{R}^r$ and $\sigma\in\Omega$, the Laplace transform*

$$\int_{\bar{\Omega}}\exp(-\langle x, \sigma^{-1}\rangle)\,\Delta_{\mathbf{s}}(x)\,\mu_p(dx)$$

converges in the following cases:

(1) $p>d'(r-1)$ *and* $s_j\geqslant 0$ *for all* $j=1, 2, ..., r$;

(2) $p=d'(j_0-1)$ *for some* $j_0\in\{2, 3, ..., r\}$, $s_j\geqslant 0$ *for all* $j=1, 2, ..., j_0-1$ *and* $s_j=0$ *for all* $j=j_0, ..., r$.

In either case, we have

$$\int_{\bar{\Omega}}\exp(-\langle x, \sigma^{-1}\rangle)\,\Delta_{\mathbf{s}}(x)\,\mu_p(dx)=(p)_{\mathbf{s}}\,\Delta_{\mathbf{s}+p}(\sigma). \tag{3.8}$$

Proof. In the non-singular case (1), which corresponds to the continuous part of the Gindikin set Λ, the result (3.8) is a reformulation of Prop. VII.1.2 in F-K (p. 124).

In the singular case (2), the details of the proof are rather delicate, and rely on results obtained by Lajmi (1998) (cf. Hassairi and Lajmi (1999, Theorem 2.2)). As in F-K (p. 138), for $u \geqslant 0$, set

$$\varepsilon(u) = \begin{cases} 0 & \text{if} \quad u = 0, \\ 1 & \text{if} \quad u > 0. \end{cases}$$

Further, let Ξ denote the image of the function $S: \mathbb{R}_+^r \to \mathbb{R}_+^r$, with

$$\mathbf{u} = (u_1, ..., u_r) \in \mathbb{R}_+^r \mapsto S(\mathbf{u}) = (s_1, ..., s_r) \in \mathbb{R}^r, \tag{3.9}$$

where

$$s_k = \begin{cases} u_1, & \text{if} \quad k = 1, \\ u_k + d'(\varepsilon(u_1) + \cdots + \varepsilon(u_{k-1})), & \text{if} \quad 2 \leqslant k \leqslant r. \end{cases}$$

For $\mathbf{s} \in \Xi$, Lajmi (1998, Chapt. 1), establishes the existence of the positive *Riesz measure* $R_{\mathbf{s}}$, which is defined as the measure on $\bar{\Omega}$ whose Laplace transform is given by

$$\int_{\bar{\Omega}} \exp(-\langle x, \sigma^{-1} \rangle) \, R_{\mathbf{s}}(dx) = \Delta_{\mathbf{s}}(\sigma)$$

for σ in Ω.

Furthermore, Lajmi defines a set of *minimal* elements in Ξ such that any $\mathbf{s}$ in Ξ can be written uniquely as a sum $\sum_{j=1}^{k} \mathbf{s}^{(j)}$ of minimal elements. By the convolution property of the Laplace transform, we then obtain the Riesz measure $R_{\mathbf{s}}$ as a convolution,

$$R_{\mathbf{s}} = R_{\mathbf{s}^{(1)}} * \cdots * R_{\mathbf{s}^{(k)}}.$$

Therefore the problem of the existence of $R_{\mathbf{s}}$ is thus reduced to the case in which $\mathbf{s}$ itself is minimal. For such a minimal $\mathbf{s}$, Lajmi (1998) applies the results of Lassalle (1987) to establish the existence of $R_{\mathbf{s}}$.

We shall give here a sketch of Lajmi's proof only for a special minimal element of interest to us, namely $\mathbf{s} + p$, where $p = d'(j_0 - 1)$ for some $j_0 = 2, 3, ..., r$, $s_j \geqslant 0$ for all $j = 1, 2, ..., j_0 - 1$ and $s_j = 0$ for all $j = j_0, ..., r$. With S denoting the map defined in (3.9), the corresponding $\mathbf{u}$ such that $\mathbf{s} + p = S(\mathbf{u})$ is

$$u_j = \begin{cases} s_j + p - (j-1)\, d', & j = 1, ..., j_0 - 1 \\ 0, & j = j_0, ..., r. \end{cases}$$

In particular, we risk no confusion by considering $\mathbf{u}$ as an element of $\mathbb{R}^{j_0 - 1}$.

Let $c = c_1 + \cdots + c_{j_0 - 1}$, where $\{c_1, ..., c_r\}$ is the fixed Jordan frame of V. Let $V(c, 1)$ and $V(c, 1/2)$ be the eigenspaces of c corresponding to the eigenvalues 1 and 1/2, respectively. We also denote by $\det_c$, $\Delta^c_{\mathbf{u}}(x)$ and Ω_c the determinant, generalized power function and cone, respectively, associated with the Jordan algebra $V(c, 1)$. We now consider the measure $\gamma_{\mathbf{u}}$ on $\Omega_c \times V(c, 1/2)$ defined by

$$\gamma_{\mathbf{u}}(dx, dv) = \frac{\Delta^c_{\mathbf{u}}(x)(\det_c x)^{-1 - d'(j_0 - 2)}}{\Gamma_{\Omega_c}(\mathbf{u})(2\pi)^{d'(j_0 - 1)(r + 1 - j_0)}}. \tag{3.10}$$

We also consider the map α from $\Omega_c \times V(c, 1/2)$ to V defined by

$$(x, v) \mapsto x + 2\sqrt{x} \circ v + (e - c) \circ v^2, \tag{3.11}$$

where $\sqrt{x}$ is the unique element in Ω_c whose square is x. Then Lajmi (1998, Théorème 1.4) proves that $R_{\mathbf{s}+p}$ equals $\alpha(\gamma_{\mathbf{u}})$, the image of $\gamma_{\mathbf{u}}$ under α. Note that this result holds also for the case in which $\mathbf{s} = \mathbf{0}$, and then we denote the corresponding $\mathbf{u}$ by $\mathbf{u}_0$; in this case R_p is the familiar measure μ_p which generates the Wishart distributions with p as the shape parameter. A lengthy but straightforward computation shows that

$$\gamma_{\mathbf{u}}(dx, dv) = (p)_{\mathbf{s}}\, \Delta_{\mathbf{s}}(\alpha(x, v))\, \gamma_{\mathbf{u}_0}(dx, dv). \tag{3.12}$$

Applying the map α in (3.11) to both sides of (3.12) we obtain

$$R_{\mathbf{s}+p} = (p)_{\mathbf{s}}\, \Delta_{\mathbf{s}} R_p,$$

which is the desired result.

We close this section by remarking that the hypotheses on the s_i's in Theorem 3.6 are sufficient. These hypotheses can be made less restrictive, however the corresponding formulation of the theorem would then be far more complicated. A similar remark holds also for Theorem 4.1 below.

4. THE EXPECTATION FORMULA

In the following we will choose and fix a Jordan frame, and denote by $\Delta_{\mathbf{s}}$ the generalized power function associated with that frame.

We are now in position to state and prove the main result.

THEOREM 4.1. *Let* $p_1, ..., p_q \in \Lambda$ *where* $p := p_1 + \cdots + p_q > d'(r-1)$. *For* $i = 1, ..., q$, *let* $\mathbf{s}_i = (s_{i1}, ..., s_{ir}) \in \mathbb{R}^r$ *satisfying the following restrictions*:

(1) *For any i such that* $p_i > d'(r-1)$, *then* $s_{i,j} \geqslant 0$ *for all* $j = 1, 2, ..., r$.

(2) *For any i such that* $p_i = d'(j_0 - 1)$ *for some* j_0 *where* $2 \leqslant j_0 \leqslant r$, *then* $s_{i,j} \geqslant 0$ *for all* $j = 1, 2, ..., j_0 - 1$ *and* $s_{i,j} = 0$ *for all* $j = j_0, ..., r$.

Let $\mathbf{p} = (p_1, ..., p_q)$, $\mathbf{s} = \mathbf{s}_1 + \cdots + \mathbf{s}_q$, *and* $(X_1, ..., X_q)$ *be a random variable on* V^q *with the Dirichlet distribution* $D_{\mathbf{p}}$, *as defined by* (3.4). *Then, for* $a \in \Omega$ *and* $f_1, ..., f_q \in \mathbb{R}_+$, *we have*

$$\mathbb{E}\,[\Delta_{\mathbf{s}_1}(X_1) \cdots \Delta_{\mathbf{s}_q}(X_q)\, \Delta_{\mathbf{s}+p}((a + f_1 X_1 + \cdots + f_q X_q)^{-1})]$$

$$= \frac{1}{(p)_{\mathbf{s}}} \prod_{i=1}^{q} (p_i)_{\mathbf{s}_i}\, \Delta_{\mathbf{s}_i + p_i}((a + f_i e)^{-1}). \tag{4.1}$$

Proof. Let $Y_1, ..., Y_q$ be independent Wishart random variables with a common scale parameter $\sigma = a^{-1}$ and respective shape parameters $p_1, ..., p_q$. Then the random variable $S = Y_1 + \cdots + Y_q$ has a Wishart distribution with shape parameter σ and scale parameter p. Since $p > d'(r-1)$ then, almost surely, the inverse S^{-1} exists. Moreover, without loss of generality, we may write

$$(X_1, ..., X_q) = (\mathbb{P}(S^{-1/2})(Y_1), ..., \mathbb{P}(S^{-1/2})(Y_q)).$$

The basic idea of the proof, as explained in the introduction, is to evaluate by two different methods the moment-generating function

$$M_{Y_1, ..., Y_q}(f_1, ..., f_q) := \mathbb{E}\left[\exp\left(-\langle e, f_1 Y_1 + \cdots + f_q Y_q\rangle\right) \prod_{i=1}^{q} \Delta_{\mathbf{s}_i}(Y_i)\right]. \tag{4.2}$$

The first method of evaluation simply utilizes the mutual independence of $Y_1, ..., Y_q$ and Theorem 3.6. Thus, in (3.8) we replace p by p_i, $\mathbf{s}$ by $\mathbf{s}_i$, and σ^{-1} by $\sigma^{-1} + f_i e$. Noting that $\sigma^{-1} + f_i e \in \Omega$, we obtain for each $i = 1, ..., q$,

$$\mathbb{E}\,[\exp(-\langle e, f_i Y_i\rangle)\, \Delta_{\mathbf{s}_i}(Y_i)] = (\det \sigma)^{-p_i} (p_i)_{s_i}\, \Delta_{\mathbf{s}_i + p_i}((\sigma^{-1} + f_i e)^{-1}).$$

The product of these expressions for all $i = 1, ..., q$ leads to the result

$$M_{Y_1, ..., Y_q}(f_1, ..., f_q) = (\det \sigma)^{-p} \prod_{i=1}^{q} (p_i)_{s_i}\, \Delta_{\mathbf{s}_i + p_i}((\sigma^{-1} + f_i e)^{-1}). \tag{4.3}$$

The second method of evaluating (4.2) is more subtle. We shall use the fact that $(X_1, ..., X_q)$ and S are mutually independent, and then we write

$Y_i = \mathbb{P}(S^{1/2})(X_i)$, $i = 1, ..., q$, and $\mathbb{P}(S^{1/2}) = t_S k_S$, where t_S is in the triangular group T associated with the fixed Jordan frame (see F-K, Chap. VI), and $k_S \in K$, the orthogonal group. (In the classical case, t_s is the mapping $x \mapsto t_s(x) = t \cdot x \cdot t'$ where t is the lower triangular matrix such that, in the basis corresponding to the chosen Jordan frame, $s = t \cdot t'$; and the mapping $k_s = t_s^{-1} \mathbb{P}(s^{1/2})$, is $x \mapsto t^{-1} \cdot s^{1/2} \cdot x \cdot s^{1/2} \cdot (t^{-1})')$.

On applying the identity (2.2) to $x = S^{1/2}$ and $y = (f_1 X_1 + \cdots + f_q X_q)^{1/2}$, we obtain

$$\begin{aligned}\langle e, f_1 Y_1 + \cdots + f_q Y_q \rangle &= \operatorname{tr}(f_1 Y_1 + \cdots + f_q Y_q)\\ &= \operatorname{tr}(\mathbb{P}(S^{1/2})(f_1 X_1 + \cdots + f_q X_q))\\ &= \langle f_1 X_1 + \cdots + f_q X_q, S \rangle.\end{aligned}$$

Substituting this result in (4.2) and applying the well-known property of conditional expectations, $\mathbb{E}(\cdot) = \mathbb{E}(\mathbb{E}[\,\cdot \mid S])$, we obtain

$$\begin{aligned}&M_{Y_1, ..., Y_q}(f_1, ..., f_q)\\ &\quad = \mathbb{E}\,\mathbb{E}\left[\exp(-\langle f_1 X_1 + \cdots + f_q X_q, S \rangle) \prod_{i=1}^{q} \Delta_{\mathbf{s}_i}(t_S k_S(X_i)) \,\middle|\, S\right]. \qquad (4.4)\end{aligned}$$

Next, we use the fact that the law $\mathscr{L}(X_1, ..., X_q)$ of $(X_1, ..., X_q)$ is invariant under the action of K on V^q; specifically, for all k in K we have

$$\mathscr{L}(k(X_1), ..., k(X_q)) = \mathscr{L}(X_1, ..., X_q). \qquad (4.5)$$

This invariance property has been proved by Casalis and Letac (1996, Theorem 3.1(i)); however we can give here a direct proof, as follows. Suppose $s \in \Omega$ and $k \in K$ then, by decomposing s using a suitable Jordan frame, we find that $(k(s))^{-1/2} = k(s^{-1/2})$. Thus

$$\mathbb{P}((k(s))^{-1/2}) = k\mathbb{P}(s^{-1/2})\, k^* = k\mathbb{P}(s^{-1/2})\, k^{-1}.$$

Hence for all s in Ω and y in V, we have $k\mathbb{P}((k(s))^{-1/2})(y) = \mathbb{P}((k(s))^{-1/2})(k(y))$. Thus

$$\begin{aligned}(k(X_1), ..., k(X_q)) &= (k(\mathbb{P}(S^{-1/2})(Y_1)), ..., k(\mathbb{P}(S^{-1/2})(Y_q)))\\ &= (\mathbb{P}((k(S))^{-1/2})(k(Y_1)), ..., \mathbb{P}((k(S))^{-1/2})(k(Y_1))).\end{aligned}$$

Now $k(Y_1), ..., k(Y_q)$, and therefore $k(S)$ also, have Wishart distributions with scale parameter $k(\sigma)$. Since the distribution of $X_1, ..., X_q$ does not depend on σ then the proof of (4.5) is complete.

Applying the crucial property of independence of $(X_1, ..., X_q)$ and S, and the invariance property (4.5), it is clear that we can omit k_S in (4.4). Then (4.4) becomes

$$M_{Y_1, ..., Y_q}(f_1, ..., f_q) = \mathbb{E}\,\mathbb{E}\left[\exp(-\langle f_1 X_1 + \cdots + f_q X_q, S\rangle) \prod_{i=1}^{q} \Delta_{\mathbf{s}_i}(t_S(X_i)) \mid S\right]. \tag{4.6}$$

Next, we apply the identity

$$\Delta_{\mathbf{s}_i}(t_S(X_i)) = \Delta_{\mathbf{s}_i}(t_S(e))\, \Delta_{\mathbf{s}_i}(X_i),$$

$i = 1, ..., q$ (cf. F-K, Proposition VI.3.10, p. 114), and the fact that $t_S(e) = S$ to obtain

$$\prod_{i=1}^{q} \Delta_{\mathbf{s}_i}(t_S(e)) = \Delta_{\mathbf{s}}(t_S(e)) = \Delta_{\mathbf{s}}(S).$$

We substitute these equalities in (4.6) and ignore the conditioning with respect to S since S and X are independent. Then (4.6) becomes

$$M_{Y_1, ..., Y_q}(f_1, ..., f_q) = \mathbb{E}\left[\exp(-\langle f_1 X_1 + \cdots + f_q X_q, S\rangle)\, \Delta_{\mathbf{s}}(S) \prod_{i=1}^{q} \Delta_{\mathbf{s}_i}(X_i)\right]. \tag{4.7}$$

We now rewrite (4.7) by conditioning with respect to $X_1, ..., X_q$ and, in the conditional expectation $\mathbb{E}(\cdot \mid X_1, ..., X_q)$, we factorize out all terms which depend on $X_1, ..., X_q$ only. This leads to the result

$$M_{Y_1, ..., Y_q}(f_1, ..., f_q) = \mathbb{E} \prod_{j=1}^{q} \Delta_{\mathbf{s}_j}(X_j)\, \mathbb{E}[\Delta_{\mathbf{s}}(S) \exp(-\langle f_1 X_1 + \cdots + f_q X_q, S\rangle) \mid X_1, ..., X_q]. \tag{4.8}$$

In the final stage in the proof, we need to evaluate the conditional expectation, $\mathbb{E}[\cdot \mid X_1, ..., X_q]$, appearing in (4.8). We again shall use the facts that the random variables $(X_1, ..., X_q)$ and S are independently distributed and that S also has a Wishart distribution $\gamma_{p,\sigma}$. Replacing σ^{-1} in (3.8) by $\sigma^{-1}+f_1X_1+\cdots+f_qX_q$, which itself belongs to Ω, then we obtain

$$\begin{aligned}\mathbb{E}\,[\Delta_{\mathbf{s}}(S)\exp(-\langle f_1X_1+\cdots+f_qX_q, S\rangle)\mid X_1, ..., X_q]\\ =(\det\sigma)^{-p}\,(p)_{\mathbf{s}}\,\Delta_{\mathbf{s}+p}((\sigma^{-1}+f_1X_1+\cdots+f_qX_q)^{-1}),\end{aligned}$$

and (4.8) reduces to

$$\begin{aligned}&M_{Y_1, ..., Y_q}(f_1, ..., f_q)\\ &\quad=(\det\sigma)^{-p}\,(p)_{\mathbf{s}}\,\mathbb{E}\,\Delta_{\mathbf{s}+p}((\sigma^{-1}+f_1X_1+\cdots+f_qX_q)^{-1})\prod_{j=1}^{q}\Delta_{\mathbf{s}_j}(X_j)).\end{aligned} \tag{4.9}$$

By comparing the expressions for $M_{Y_1, ..., Y_q}(f_1, ..., f_q)$ in (4.9) and in (4.3), and noting that $a=\sigma^{-1}$, we obtain the desired result. Then the proof is complete.

To conclude this section, we note that by using the explicit descriptions of principal minors provided in Remark 3.5, the right-hand side of (4.1) can be written down in a straightforward manner for each of the five types of symmetric cones.

5. CONCLUDING REMARKS

A referee kindly noted the following connection between our results and the theory of Lauricella functions. Formula (1.4), which deals with the univariate Dirichlet distribution, has also been proved in the literature as a particular case of a Lauricella function. Indeed, for $a, p_i>0$ and $0<f_i<2$, $i=1, ..., q$, Carlson (1963) defined the multiple hypergeometric function

$$\begin{aligned}&R(a; p_1, ..., p_q; f_1, ..., f_q)\\ &\quad:=\sum_{m_1=0}^{\infty}\cdots\sum_{m_q=0}^{\infty}\frac{(a)_{m_1+\cdots+m_q}}{(p_1+\cdots+p_q)_{m_1+\cdots+m_q}}\prod_{i=1}^{q}\frac{(p_i)_{m_i}}{m_i!}(1-f_i)^{m_i}.\end{aligned}$$

It follows from a formula given by Carlson (1963, Eq. (7.10)) that, with the notation of (1.4) in our paper,

$$R(a; p_1, ..., p_q; f_1, ..., f_q)=\mathbb{E}(f_1X_1+\cdots+f_qX_q)^{-a}.$$

This clearly yields (1.4) for the case in which $a = p_1 + \cdots + p_q$. Formula (2.3.5) of Exton (1976) provides the same result.

For the case in which $q = 2$ and $p_1, p_2 \geqslant d'(r-1)$, there is also a connection between formula (1.6) and the Gaussian hypergeometric functions of matrix argument (cf. Muirhead, 1982). This connection can be extended to any symmetric cone Ω, in which case the expectation formula appears in terms of the Gaussian hypergeometric functions on Ω, denoted by ${}_2F_1$ (cf. F-K, p. 329). For $q = 2$ and $p_1, p_2 > d'(r-1)$, we have $(X_1, X_2) \in \overline{T_2}$, the set defined in (3.5); then, by (3.6), the marginal density of X_1 is

$$\frac{1}{B_\Omega(p_1, p_2)} (\det x_1)^{p_1 - (n/r)} \det(e - x_1)^{p_2 - (n/r)}, \tag{5.10}$$

$x_1, e - x_1 \in \Omega$. Since $X_2 = e - X_1$ it follows that

$$f_1 X_1 + f_2 X_2 = f_2 e + (f_1 - f_2)\, X_1 = f_2(e - f_0 X_1),$$

where $f_0 = (f_2 - f_1)/f_2$. By (5.10), for any $p \in \mathbb{R}$, $|f_0| < 1$, and $p_1, p_2 > d'(r-1)$,

$$\begin{aligned}\mathbb{E} \det(e - f_0 X_1)^{-p} &= \frac{1}{B_\Omega(p_1, p_2)} \int_{x_1,\, e - x_1 \in \Omega} (\det x_1)^{p_1 - (n/r)} \det(e - x_1)^{p_2 - (n/r)} \\ &\quad \times \det(e - f_0 x_1)^{-p}\, dx_1 \\ &= {}_2F_1(p, p_1; p_1 + p_2; f_0 e),\end{aligned}$$

where the last equality follows from F-K, (p. 330, Proposition XV.3.2). Assuming also that $f_1, f_2 > 0$, we then deduce that

$$\mathbb{E} \det(f_1 X_1 + f_2 X_2)^{-p} = f_2^{-rp}\, {}_2F_1(p, p_1; p_1 + p_2; f_0 e). \tag{5.11}$$

Since ${}_2F_1(p_1 + p_2, p_1; p_1 + p_2; x) = \det(e - x)^{-p_1}$ (cf. F-K, p. 330, Proposition XV.3.4(i)), then it follows that

$$\begin{aligned}\mathbb{E} \det(f_1 X_1 + f_2 X_2)^{-(p_1 + p_2)} &= f_2^{-r(p_1 + p_2)} \det((1 - f_0)\, e)^{-p_1} \\ &= f_2^{-r(p_1 + p_2)}\, (1 - f_0)^{-rp_1} \\ &\equiv f_1^{-rp_1} f_2^{-rp_2}.\end{aligned} \tag{5.12}$$

The above derivation proceeded under the additional assumption that $|f_0| < 1$; however, this assumption can be removed by elementary analyticity considerations, so that (5.12) holds for all $f_1, f_2 > 0$.

The formula (5.11) also leads to simple evaluations for values of p other than p_1+p_2. For instance, suppose that $p=p_1+p_2+1$; then by a second application of F-K (p. 330, Proposition XV.3.4(i) to (5.11)), we obtain

$$\begin{aligned}\mathbb{E}\det(f_1X_1+f_2X_2)^{-(p_1+p_2+1)} &= f_2^{-r(p_1+p_2+1)}\det((1-f_0)\,e)^{-p_1} \\ &\quad\times {}_2F_1(-1,p_2;p_1+p_2;-f_0(1-f_0)^{-1}e).\end{aligned}$$

This latter ${}_2F_1$, by virtue of the parameter -1, reduces to a finite series, each term of which can be calculated explicitly from F-K (p. 329). After a lengthy, but elementary, calculation we obtain

$$\begin{aligned}&\mathbb{E}\,[\det(f_1X_1+f_2X_2)^{-(p_1+p_2+1)}] \\ &= f_1^{-rp_1}f_2^{-r(p_2+1)}\sum_{k=0}^{r}\frac{\left(-\frac{1}{d'}-k+1\right)_k\left(\frac{p_1}{d'}-k+1\right)_k\left(\frac{1}{d'}+r-k\right)_k}{\left(\frac{p_1+p_2}{d'}-k+1\right)_k\left(\frac{n}{d'r}-k+1\right)_k\left(\frac{1}{d'}\right)_k} \\ &\quad\times\binom{r}{k}\left(1-\frac{f_2}{f_1}\right)^k.\end{aligned}$$

Thus, even for $q=2$, a small change in the exponent of $\det(f_1X_1+f_2X_2)$ leads to an expectation formula which cannot be reduced to closed form. This result underscores the remarkable nature of the closed expression (4.1), our main result.

Finally, we remark that for $q=2$, some special cases of the evaluation formula (4.1) can be given an interpretation in terms of a class of Gaussian hypergeometric functions treated by Gindikin (1964), Section 4.

ACKNOWLEDGMENTS

The authors are grateful to the referees for helpful and detailed comments on the initial version of this article.

REFERENCES

T. W. Anderson, R. A. Fisher and multivariate analysis, *Statist. Sci.* **11** (1996), 20–34.

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

S. Andersson, H. K. Brons, and S. T. Jensen, Distribution of eigenvalues in multivariate statistical analysis, *Ann. Statist.* **11** (1983), 392–415.

P. Artzner and G. Fourt, Lois gamma et bêta sur un cône convexe homogène. Applications en analyse multivariée, *C.R. Acad. Sci. Paris Sér. A* **278** (1974), 293–295.

B. C. Carlson, Lauricella's hypergeometric function F_D, *J. Math. Anal. Appl.* **73** (1963), 452–470.

M. Casalis and G. Letac, The Lukacs–Olkin–Rubin characterization of the Wishart distributions on the symmetric cones, *Ann. Statist.* **24** (1996), 763–784.

J.-F. Chamayou and G. Letac, Transient random walk on stochastic matrices with Dirichlet distribution, *Ann. Probab.* **22** (1994), 424–430.

J. A. Diáz-García and R. Gutiérrez Jáimez, Proof of the conjectures of H. Uhlig on the singular multivariate beta and the Jacobian of a certain matrix transformation, *Ann. Statist.* **25** (1997), 2018–2023.

H. Exton, "Multiple Hypergeometric Functions and Applications," Ellis Horwood, New York, 1976.

J. Faraut and A. Korányi, "Analysis on Symmetric Cones," Oxford University Press, New York, 1994.

G. Fourt, Lois B sur les cones homogènes, *Ann. Sci. Univ. Clermont* **51** (1974), 22–30.

S. G. Gindikin, Analysis in homogeneous domains, *Uspekhi Mat. Nauk* **19**, 4 (1964), 3–92. [English transl.: *Russian Math. Surveys* **19** (1964), 1–90].

R. D. Gupta and D. St. P. Richards, Multivariate Liouville distributions, IV, *J. Multivariate Anal.* **54** (1995), 1–17.

R. D. Gupta and D. St. P. Richards, The history of the Dirichlet and Liouville distributions, *J. Statist. Plann. Inference*, preprint, 1999, University of Virginia.

A. Hassaïri and S. Lajmi, Riesz exponential families on symmetric cones, preprint, Université de Sfax, 1999.

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

S. Karlin, C. A. Micchelli, and Y. Rinott, Multivariate splines: a probabilistic perspective, *J. Multivariate Anal.* **20** (1986), 69–90.

S. Lajmi, "Les Familles Exponentielles de Riesz sur les Cônes Symétriques," Thèse, Université de Sfax, 1998.

M. Lassalle, Algèbre de Jordan et ensemble de Wallach, *Invent. Math.* **89** (1987), 375–393.

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

G. Letac and H. Massam, A formula on multivariate Dirichlet distributions, *Statist. Probab. Lett.* (1998b), 247–254.

G. Letac and M. Scarsini, Random nested tetrahedra, *Adv. Appl. Prob.* **30** (1998), 619–627.

H. Massam, 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** (1994), 369–394.

J. G. Mauldon, A generalization of the beta-distribution, *Ann. Math. Statist.* **30** (1959), 509–520.

R. J. Muirhead, "Aspects of Multivariate Statistical Theory," Wiley, New York, 1982.

H. Uhlig, On singular Wishart and singular multivariate beta distributions, *Ann. Statist.* **22** (1994), 395–405.