The Annals of Statistics
1996, Vol. 24, No. 2, 763–786

THE LUKACS–OLKIN–RUBIN CHARACTERIZATION OF WISHART DISTRIBUTIONS ON SYMMETRIC CONES

BY M. CASALIS AND G. LETAC

Université Paul Sabatier

We characterize the Wishart distributions on a symmetric cone C. If $C = (0, +\infty)$, this has been done by Lukacs in 1955. If C is the cone of positive definite symmetric matrices, this has been done by Olkin and Rubin in 1962. We both shorten and extend the Olkin–Rubin proof (sometimes obscure) by using three modern ideas: (i) try to avoid artificial coordinates in differential geometry; (ii) the variance function of a natural exponential family F characterizes F; (iii) symmetric matrices are a particular example of a Euclidean simple Jordan algebra.

1. Introduction. The last decade of this century does not seem to be fond of characterizations of distributions in statistics: too many trivial theorems about the characterizations of the uniform or exponential laws have appeared in the literature, and even elegant theorems are in disrepute because they have not yet proved useful in applied statistics. However, they give insight into the laws of nature and they may reveal quite beautiful mathematics.

One of these elegant theorems is due to Lukacs (1955). We state it as follows.

THEOREM 1.1. *Let $\sigma > 0$ and $p > 0$. The gamma distribution on $\mathbb{R}$ with scale parameter σ and shape parameter p is*

$$\gamma_{p,\sigma}(dx) = \exp(-x\sigma^{-1})x^{p-1}\sigma^{-p}(\Gamma(p))^{-1}\mathbb{1}_{(0,+\infty)}(x)\,dx. \tag{1.1}$$

Let U and V be two independent non-Dirac and nonnegative random variables such that $U+V$ is a.s. positive, and define $Z = U/(U+V)$. Then $U+V$ and Z are independent if and only if there exists $\sigma > 0$, $p > 0$ and $q > 0$ such that $\mathscr{L}(U) = \gamma_{p,\sigma}$ and $\mathscr{L}(V) = \gamma_{q,\sigma}$.

We shall give a proof of Theorem 1.1 in Section 2: intrinsically, it is the Lukacs proof, but here Laplace transforms replace characteristic functions; an essential idea to permit the use of natural exponential families.

Theorem 1.1 has been nicely generalized by Olkin and Rubin (1962). We shall state their result in Theorem 1.2 in a slightly different but more general

Received November 1994; revised May 1995.
[1]Research funded by NATO Research Grant CRG 92 1347.
AMS 1991 *subject classifications*. Primary 62H05; secondary 60E10.
Key words and phrases. Beta and Dirichlet distributions, natural exponential families, Jordan algebras.

form. Their paper suffers from several obscurities, and we comment on them in Section 7.

Let us first introduce some notation, including the Wishart distributions on symmetric matrices. For the sake of convenience and future generalizations, we change the traditional notation for Wishart distributions slightly [as described in e.g., Muirhead (1982) and Seber (1984)], just as it is easier to work with gamma distributions instead of χ^2 if we do not have practical purposes.

For a fixed integer $r \geq 1$, $M \supset E \supset \overline{E}_+ \supset E_+$ denote, respectively, the (r, r) real matrices, the symmetric matrices, the positive matrices and the positive definite matrices. If σ is in E_+ and if p belongs to the set

$$\left\{\frac{1}{2}, 1, \dots, \frac{r-1}{2}\right\} \cup \left(\frac{r-1}{2}, +\infty\right), \tag{1.2}$$

then the Wishart distribution $\gamma_{p,\sigma}$ on $\overline{E}_+$ with scale parameter σ and shape parameter p is defined by its Laplace transform as follows: for θ in E_+,

$$\int_{\overline{E}_+} \exp(-\operatorname{Trace} \theta x)\gamma_{p,\sigma}(dx) = \left(\det(I_r + \theta\sigma)\right)^{-p}, \tag{1.3}$$

where I_r is the identity matrix. [For explanations about the gaps of (1.2) consult Gindikin (1975); a self contained proof is in Casalis and Letac (1994).]

We are not going to give too many details here about $\gamma_{p,\sigma}$; see the references above. Let us just recall that if $\mathscr{L}(U) = \gamma_{p,\sigma}$, then U^{-1} exists a.s. if $p > (r-1)/2$, and U^{-1} exists with probability 0 if $p \leq (r-1)/2$. Here is now our version of the Olkin–Rubin theorem [Olkin and Rubin (1962)].

THEOREM 1.2. *With the previous notation, let w: $E_+ \to M$ be a measurable function such that, for all b in E_+, one has $w(b)w(b)^t = b$. Let U and V be two independent non-Dirac random variables of $\overline{E}_+$ such that $U + V$ is in E_+ almost surely, and define*

$$Z = (w(U+V))^{-1}U\left((w(U+V))^t\right)^{-1}.$$

Then we have equivalence between the following:

(a) *$U + V$ and Z are independent, U and V are not concentrated on the same one-dimensional space and, for any orthogonal matrix Γ, $\mathscr{L}(Z) = \mathscr{L}(\Gamma Z \Gamma^t)$;*

(b) *there exist σ in E_+, p and q in* (1.2), *such that $p + q > (r-1)/2$, and $\mathscr{L}(U) = \gamma_{p,\sigma}$ and $\mathscr{L}(V) = \gamma_{q,\sigma}$.*

Olkin and Rubin seem to assume that U and V are in E_+. They do not prove (b) $\Rightarrow$ (a); they call the proof straightforward: in this case one can actually use the fact that U and V have densities; this enables us to compute the joint density of $(Z, U+V)$. Details appear in a later paper by Olkin and Rubin (1964). If U and V are not necessarily invertible, this method is no longer available, and we shall give a proof of (b) $\Rightarrow$ (a) in Section 4 which uses only the fact that $\mathscr{L}(U)$ and $\mathscr{L}(V)$ belong to natural exponential families which are generated by measures on E_+ which are quasi-invariant by transformations $E \to E$ of the form $u \mapsto wuw^t$.

Anyway, the hard part is (a) ⇒ (b). We became interested in Theorem 1.2 because of the fact that many properties of gamma and Wishart distributions have recently been extended to Wishart distributions on symmetric cones [Casalis (1990, 1991), Letac (1994) Massam (1994), Casalis and Letac (1994), Massam and Neher (1996) and Letac and Massam (1995)]. See also Artzner and Fourt (1974). However, understanding the Olkin–Rubin proof of Theorem 1.2 appears to be a strenuous task [we gave up after their identity (23)]. In fact, we have realized that considering their beautiful ideas before their identity (17), but using modern tools like:

(i) differentiation without coordinates,
(ii) variance functions of natural exponential families and
(iii) Euclidean simple Jordan algebras,

leads to a neater proof, ready for a generalization to a general symmetric cone.

We are aware that concepts (i) and (ii) belong to the realm of statisticians, and not (iii) per se (although there are some timid appearances, in the statistical literature, of Wishart distributions on Hermitian complex or quaternionic matrices). For this reason we organize this paper as follows: Section 2 contains a reminder of the variance function of a natural exponential family, as well a proof of Theorem 1.1. Section 3 recalls the definitions about Euclidean Jordan algebras, symmetric cones and Wishart distributions on them, and states the extension of Theorem 1.2 in this context (Theorems 3.1 and 3.2). Section 4 proves the (b) ⇒ (a) part of Theorem 1.2 by proving Theorem 3.1. Sections 5 and 6 are really the heart of the paper and prove the (a) ⇒ (b) part of Theorem 1.2 through Theorem 3.2. Things are explained in such a way that the reader who is interested only in Theorem 1.2 and in classical Wishart distributions can follow the proofs of Sections 4, 5 and 6. The concluding Section 7 offers some comments on Olkin and Rubin (1962).

2. The variance function of a natural exponential family. Since coordinates are a special burden in Olkin and Rubin (1962), we try to work without them as much as possible.

Let E be a finite-dimensional real linear space, and let E^* be its dual. We let

$$(\theta, x) \mapsto \langle \theta, x \rangle, \qquad E^* \times E \to \mathbb{R}$$

denote the canonical bilinear map on $E^* \times E$. The set $\mathscr{M}(E)$ is the set of positive (possibly unbounded) measures μ on E such that the following hold:

1. μ is not concentrated on any affine hyperplane;
2. if $L_\mu(\theta) = \int_E \exp\langle \theta, x\rangle \mu(dx)$ $(\leq +\infty)$ is the Laplace transform of μ, then the interior $\Theta(\mu)$ of the convex set
$$\{\theta \in E^*;\, L_\mu(\theta) < +\infty\}$$
is not empty.

If μ is in $\mathscr{M}(E)$ and if θ is in $\Theta(\mu)$, we write $k_\mu(\theta) = \log L_\mu(\theta)$ and

$$P(\theta, \mu)(dx) = \exp\bigl(\langle \theta, x\rangle - k_\mu(\theta)\bigr)\mu(dx).$$

The set $F = F(\mu) = \{P(\theta, \mu);\ \theta \in \Theta(\mu)\}$ is called the natural exponential family (NEF) generated by μ. The differential $k'_\mu(\theta)$ of k_μ: $\Theta(\mu) \to \mathbb{R}$, evaluated in θ, is a linear form on E^*, that is, an element of E, which is related to $P(\theta, \mu)$ by

$$k'_\mu(\theta) = \int_E x P(\theta, \mu)(dx). \tag{2.1}$$

It is a standard exercise to see that k_μ is strictly convex and real analytic on $\Theta(\mu)$. Thus $\theta \mapsto k'_\mu(\theta)$ is one-to-one. The set $k'_\mu(\Theta(\mu)) = M_F$ of the images is called, because of (2.1), the domain of the means of the NEF $F(\mu)$. Since k'_μ is a bijection, let ψ_μ denote its inverse function $M_F \to \Theta(\mu)$, and define

$$P(m, F) = P\bigl(\psi_\mu(m), \mu\bigr). \tag{2.2}$$

It is easily seen that if $F(\mu) = F(\mu')$, that is, if there exists (θ_0, c) in $E^* \times \mathbb{R}$ such that

$$\mu'(dx) = \exp(\langle \theta_0, x\rangle + c)\mu(dx),$$

then neither M_F nor (2.2) change when replacing μ by μ'.

Finally, for m in M_F, we define the symmetric linear operator $V_F(m)$: $E^* \to E$ as the covariance operator of the probability $P(m, F)$ on E. Denoting by $L_S(E^*, E)$ the space of the symmetric operators from E^* to E, the map

$$M_F \to L_S(E^*, E), \qquad m \mapsto V_F(m) \tag{2.3}$$

is called the variance function of the NEF F; it has the important property that it characterizes F in the following sense: if F and F' are NEF's on E such that V_F and $V_{F'}$ are equal on a nonvoid open set I contained in $M_F \cap M_{F'}$, then $F = F'$. This fact is easily deduced from the following formula: if $F = F(\mu)$, then, for all θ in $\Theta(\mu)$,

$$k''_\mu(\theta) = V_F\bigl(k'_\mu(\theta)\bigr). \tag{2.4}$$

Note that, in (2.4), $k''_\mu(\theta)$ is in $L_S(E^*, E)$.

Of course, if U is a connected open set of E, not all analytic maps V from U to the set of positive definite elements of $L_S(E^*, E)$ are such that there exists a NEF F on E with $U \subset M_F$ and $V = V_F$ on U. In particular, for $\dim E > 1$, variance functions have to satisfy the symmetry condition (2.5). It is similar to the fact that a smooth vector field φ: $U \to E^*$ is the differential of some function f: $U \to \mathbb{R}$ only if $\varphi'(m)$ is symmetric, that is, belongs to $L_S(E, E^*)$. This symmetry condition (2.5) will be the crux of our proof of Theorem 1.2 and Theorem 3.2 [see after (6.18)].

PROPOSITION 2.1. *Let F be a NEF on E. Then, for all m in M_F and for all α and β in E^*, we have*

$$\bigl(V'_F(m)(V_F(m)(\alpha))\bigr)(\beta) = \bigl(V'_F(m)(V_F(m)(\beta))\bigr)(\alpha). \tag{2.5}$$

PROOF. From the definition of $\psi_\mu\colon M_F \to \Theta(\mu)$, (2.4) gives

$$k''_\mu\big(\psi_\mu(m)\big) = V_F(m),$$

which implies that, since $k'_\mu(\psi_\mu(m)) = m$, for all h in E,

$$\psi'_\mu(m)(h) = \big(V_F(m)\big)^{-1}(h). \tag{2.6}$$

Differentiating (2.6) with respect to m gives, for all h and k in E,

$$\big(\psi''_\mu(m)\big)(h,k) = -\big(V_F(m)\big)^{-1}\big(V'_F(m)(k)\big)\big(V_F(m)\big)^{-1}(h). \tag{2.7}$$

From the symmetry property of Hessians, that is, from

$$\psi''_\mu(m)(h,k) = \psi''_\mu(m)(k,h),$$

we get that (2.7) is symmetric in (h,k). Writing $\alpha = (V_F(m))^{-1}(h)$ and $\beta = (V_F(m))^{-1}(k)$ now gives (2.5). □

To complete this presentation of NEF's, let us define the Jorgensen set of a μ in $\mathscr{M}(E)$ and of a NEF. For μ in $\mathscr{M}(E)$, the Jorgensen set is

$$\Lambda(\mu) = \Big\{p > 0; \exists\, \mu_p \text{ in} \mathscr{M}(E) \text{ with } \Theta(\mu) = \Theta(\mu_p) \text{ and } \big(L_\mu(\theta)\big)^p = L_{\mu_p}(\theta)\Big\}.$$

It is trivial to check that if $F = F(\mu) = F(\mu')$, then $\Lambda(\mu) = \Lambda(\mu')$ and $F(\mu_p) = F(\mu'_p)$. Thus we are allowed to talk about the Jorgensen set $\Lambda(F)$ of F and to write $F_p = F(\mu_p)$ if $p \in \Lambda(F)$.

Our best examples of NEF's for the present paper are made with the Wishart distributions on the space E of (r,r) real symmetric matrices. We still denote by E_+ the cone of positive definite matrices. Let us fix p in (1.2), with $p \neq 0$. Then one can find a μ_p in $\mathscr{M}(E)$ such that

$$F_p = F(\mu_p) = \{\gamma_{p,\sigma};\, \sigma \in E_+\},$$

where $\gamma_{p,\sigma}$ is defined by (1.3). In this case $M_{F_p} = E_+$. To describe V_F it is convenient to introduce the bilinear form on E,

$$(a,b) \mapsto \operatorname{Trace}(ab).$$

Writing $\varphi_a(b) = \operatorname{Trace}(ab)$, then $a \mapsto \varphi_a$ is an isomorphism between E and E^*. Identifying E and E^* through it, the variance function of F_p is the element of $L_S(E,E) = L_S(E)$ defined by

$$\theta \mapsto \frac{1}{p} m\theta m = V_{F_p}(m)(\theta). \tag{2.8}$$

Details about this example can be found in Letac (1989). Note that if $r = 1$, $M_F = (0,\infty)$ and $V_F(m) = m^2/p$. Note also that the Jorgensen set of F_1 is equal to (1.2). We shall comment on this nontrivial fact, called Gindikin's theorem, in Sections 3 and 7.

Having done this, we now recall the proof of the necessary part of Theorem 1.1 in such a way that the concepts of NEF become apparent.

PROOF OF THEOREM 1.1 ($\Rightarrow$ First part). [If the law μ of a random variable X, with values in a linear space E, happens to belong to $\mathscr{M}(E)$, we adopt the obvious notations $L_X, k_X, \Theta(X), \dots$ instead of $L_\mu, k_\mu, \Theta(\mu), \dots$.]

Note that U and $V \geq 0$ imply that $\Theta(U)$ and $\Theta(V)$ contain $(-\infty, 0)$, and that $0 \leq Z \leq 1$ implies that $\Theta(Z) = \mathbb{R}$. Observe that, for (θ, ξ) in $(-\infty, 0) \times \mathbb{R}$, we have

$$\mathbb{E}(\exp\{\theta(U+V) + \xi Z\}) = \exp(k_U(\theta) + k_V(\theta) + k_Z(\xi)). \tag{2.9}$$

(i) Write $c_1 = k_Z'(0) = \mathbb{E}(Z)$. Applying $\partial^2/\partial\theta\,\partial\xi$ to both sides of (2.9) and setting $\xi = 0$ gives, since $(U+V)Z = U$,

$$k_U' L_U L_V = (k_U' + k_V') c_1 L_U L_V.$$

Writing $\chi = k_{U+V}$ for simplicity, we get

$$k_U = c_1 \chi \quad \text{and} \quad k_V = (1 - c_1)\chi, \tag{2.10}$$

that is, c_1 and $1 - c_1$ are in the Jorgensen set of $U + V$, and $(\mathscr{L}(U+V))_{c_1} = \mathscr{L}(U)$, $(\mathscr{L}(U+V))_{1-c_1} = \mathscr{L}(V)$.

(ii) Write $c_2 = \mathbb{E}(Z^2) = k_Z''(0) + c_1^2$. Applying $\partial^4/\partial\theta^2\,\partial\xi^2$ to both sides of (2.9) and letting $\xi = 0$ gives, since $(U+V)^2 Z^2 = U^2$,

$$\left(k_U'' + (k_U')^2\right) L_U L_V = \left(\chi'' + (\chi')^2\right) c_2 L_U L_V.$$

Using (2.10), we get, for $\theta < 0$,

$$\chi''(\theta) = \frac{c_2 - c_1^2}{c_2 - c_1} (\chi'(\theta))^2 \tag{2.11}$$

($c_2 - c_1 = 0$ would imply that U and V both have Dirac distributions; our hypothesis excludes this; the same remark holds for $c_2 - c_1^2 = 0$).

Taking $\lambda = (c_2 - c_1)/(c_2 - c_1^2)$ and comparing (2.11) with (2.4), we see that the variance function of the NEF $F = F(U+V)$ is such that $V_F(m) = m^2/\lambda$ for $m > 0$, that is, $U + V$ is gamma distributed with shape parameter λ. $\square$

Of course, in the above proof, NEF's are not really useful, since (2.11) can be easily integrated. However, in higher dimensions, (2.11) is replaced by a complicated system of differential equations and, as we shall see in Section 6, NEF's become an essential tool.

3. Wishart distributions on irreducible symmetric cones. We first recall the essential definitions and facts about these objects and about the Euclidean Jordan algebras.

A Euclidean Jordan algebra is a Euclidean space E (with scalar product denoted $\langle a, b\rangle$) equipped with a bilinear application

$$(a, b) \mapsto a \cdot b, \qquad E \times E \to E$$

and a neutral element e in E such that, for all a, b, c in E, one has

$$
(3.1)\qquad
\begin{aligned}
&\text{(i)} && a \cdot b = b \cdot a,\\
&\text{(ii)} && a \cdot ((a \cdot a) \cdot b) = (a \cdot a) \cdot (a \cdot b),\\
&\text{(iii)} && \langle a, b \cdot c \rangle = \langle a \cdot b, c \rangle,\\
&\text{(iv)} && a \cdot e = a.
\end{aligned}
$$

If E is the Cartesian product of two Euclidean Jordan algebras E_1 and E_2 with positive dimension, E is nonsimple. Otherwise, E is said to be simple. There are essentially (up to linear isomorphism) only five kinds of Euclidean simple Jordan algebras. If K denotes either the real numbers $\mathbb{R}$, the complex ones $\mathbb{C}$, the quaternions $\mathbb{H}$ or the octonions $\mathbb{O}$, write $S_r(K)$ for the space of (r, r) Hermitian matrices valued in K, endowed with the Euclidean structure $\langle a, b \rangle = \mathrm{Trace}(a\bar{b})$ and with the Jordan product

$$a \cdot b = \tfrac{1}{2}(ab + ba),$$

where ab is the ordinary product of matrices. Then $S_r(\mathbb{R})$, $r \geq 1$, $S_r(\mathbb{C})$, $r \geq 2$, $S_r(\mathbb{H})$, $r \geq 2$, and the exceptional $S_3(\mathbb{O})$ gives the list of the four first kinds. The fifth kind is the Euclidean space $\mathbb{R}^n$, $n \geq 3$, with Jordan product

$$
(3.2)\qquad
\begin{aligned}
&(x_0, x_1, \dots, x_{n-1}) \cdot (y_0, y_1, \dots, y_{n-1})\\
&\quad = \left(\sum_{i=0}^{n-1} x_i y_i, x_0 y_1 + y_0 x_1, \dots, x_0 y_{n-1} + y_0 x_{n-1} \right).
\end{aligned}
$$

For our purposes $E = S_r(\mathbb{R})$ is the most important example.

Now, to each Euclidean simple Jordan algebra E, we attach the set of Jordan squares

$$\overline{E}_+ = \{ x \in E; \text{there exists } a \text{ in } E \text{ such that } x = a \cdot a \}.$$

Its interior is denoted E_+. If $E = S_r(\mathbb{R})$, $\overline{E}_+$ and E_+ are the familiar cones of positive and positive definite symmetric matrices. For (3.2), $\overline{E}_+$ is a closed cone of revolution. In general, E_+ is a symmetric cone, that is, a convex cone which is as follows:

(3.3)

(i) self dual, that is, $E_+ = \left\{ x \in E; \langle x, y \rangle > 0 \ \forall \ y \in \overline{E}_+ \setminus \{0\} \right\}$;

(ii) homogeneous, that is, the group of linear automorphisms of E which preserve E_+ acts transitively on E_+;

(iii) salient, that is, E_+ does not contain a line.

Furthermore it is irreducible, in the sense that it is not the Cartesian product of two convex cones. One can prove [see, e.g., Faraut and Koranyi (1994) Theorem 3.3.1, p. 49] that an open convex cone is symmetric and irreducible if and only if it is the E_+ of some Euclidean simple Jordan algebra.

Given a Euclidean simple Jordan algebra E, we denote by $G(E)$ the subgroup of the linear group $\mathrm{GL}(E)$ of linear automorphisms which preserves E_+, and we denote by G the connected component of $G(E)$ containing the identity. Recall that if $E = S_r(\mathbb{R})$ and $\mathrm{GL}(\mathbb{R}^r)$ is the group of invertible (r, r) matrices, elements of $G(E)$ are the maps $g\colon E \to E$ such that there exists a in $\mathrm{GL}(\mathbb{R}^r)$ with

$$g(x) = axa^t. \tag{3.4}$$

If E is given by (3.2), $G(E)$ is made with two of the four connected components of $\mathbb{O}(1, n-1)$.

We also write $K = G \cap \mathbb{O}(E)$, where $\mathbb{O}(E)$ is the orthogonal group of E. The elements k of K satisfy

$$k(x \cdot y) = kx \cdot ky \tag{3.5}$$

and are called Jordan automorphisms. In particular, $ke = e$ and this equality characterizes K,

$$K = \{g \in G;\ ge = e\}. \tag{3.6}$$

Now for x in E, one denotes by $L(x)$ the linear operator on E $y \mapsto x \cdot y$ and

$$P(x) = 2L(x)^2 - L(x \cdot x). \tag{3.7}$$

From (3.1), $L(x)$ and hence $P(x)$ are symmetric.

The map $P\colon E \mapsto L_S(E)\colon x \mapsto P(x)$ is called the *quadratic representation* of E. It satisfies the following properties:

1. $P(e) = \mathrm{id}_E$, $P(x)e = x \cdot x$;
2. $P(x)$ is invertible if and only if x is invertible in E, that is, there exists $y = x^{-1}$ in E such that $x \cdot y = e$; then $P(x)^{-1} = P(x^{-1})$;
3. if x is in E_+, $P(x)$ is positive definite and $P(x) \in G$;
4. for g in G, with transpose g^*, and for x in E,

$$P(gx) = gP(x)g^*. \tag{3.8}$$

This last equality is not obvious and relies on the following three facts:

$$kP(x)k^* = P(kx) \qquad \text{for all } (x, k) \text{ in } E \times K \text{ [from (3.5)]};$$

$$P(P(y)x) = P(y)P(x)P(y) \quad \text{for all } (x, y) \text{ in } E^2;$$

and, for g in G, there exists (y, k) in $E_+ \times K$ such that $g = P(y)k$ [see Faraut and Koranyi (1994), Proposition 2.3.3, p. 33 and Theorem 3.5.1, p. 55].

We now briefly introduce some useful decompositions in E. An element c of E is said to be a *primitive idempotent* if $c \cdot c = c \neq 0$ and if c is not the sum $t + u$ of two nonnull idempotents t and u such that $t \cdot u = 0$. A *complete system of primitive orthogonal idempotents* is a set $\{c_1, \dots, c_r\}$ such that

$$\sum_{i=1}^{r} c_i = e \quad \text{and} \quad c_i \cdot c_j = \delta_{ij} c_i \quad \text{for } 1 \le i, j \le k. \tag{3.9}$$

The size r of such a system is a constant called the *rank* of E. When E is $S_r(\mathbb{R})$, this is the set of projection matrices on the r lines generated by the vectors of an orthogonal basis. Hence the rank is r. When E is given by (3.2), then $r = 2$ and $c_1 = \frac{1}{2}(1, x_1, \dots, x_{n-1})$, $c_2 = \frac{1}{2}(1, -x_1, \dots, -x_{n-1})$, with $x_1^2 + \cdots + x_{n-1}^2 = 1$.

Any element x of a Euclidean simple Jordan algebra can be written $x = \sum_{i=1}^{r} \lambda_i c_i$ in a suitable complete system of primitive orthogonal idempotents. The real numbers $\lambda_1, \dots, \lambda_r$ are the *eigenvalues* of x, and this decomposition is called the *spectral decomposition* of x. One then defines the *trace* of x by $\operatorname{Trace}(x) = \sum_{i=1}^{r} \lambda_i$ and its *determinant* by $\det x = \prod_{i=1}^{r} \lambda_i$. When x is invertible, clearly $x^{-1} = \sum_{i=1}^{r} \lambda_i^{-1} c_i$ and, $\overline{E}_+$ being the set of squares of E, x belongs to E_+ if and only if its eigenvalues are strictly positive. Its *square root* is then defined by $x^{1/2} = \sum_{i=1}^{r} \lambda_i^{1/2} c_i$ and satisfies $x^{1/2} \cdot x^{1/2} = x$. All these definitions are the usual ones in $S_r(\mathbb{R})$. For the cone of revolution, $\operatorname{Trace}(x) = 2x_0$, $\det x = x_0^2 - x_1^2 \cdots - x_{n-1}^2$.

If c is a primitive idempotent of E, the only possible eigenvalues of $L(c)$ are 0, $\frac{1}{2}$, and 1. We denote by $E(c, 0)$, $E(c, \frac{1}{2})$ and $E(c, 1)$ the corresponding eigenspaces. The decomposition

$$E = E(c, 0) \oplus E\left(c, \tfrac{1}{2}\right) \oplus E(c, 1)$$

is called the *Peirce decomposition of E with respect to c.*

Now, we fix a complete system of primitive orthogonal idempotents $\{c_i\}_{i=1}^{r}$ and for any (i, j) we write

$$E_{ii} = E(c_i, 1) = \mathbb{R} c_i,$$
$$E_{ij} = E\left(c_i, \tfrac{1}{2}\right) \cap E\left(c_j, \tfrac{1}{2}\right) \quad \text{if } i \neq j.$$

It can be proved that

$$E = \bigoplus_{i \le j} E_{ij} \tag{3.10}$$

[see Faraut and Koranyi (1994) Theorem 4.2.1, p. 68]. This is the *Peirce decomposition of E with respect to* $\{c_i\}_{i=1}^{r}$. Moreover the dimension of E_{ij} when $i \neq j$ is a constant d, so that (3.10) yields the relation

$$n = r + \frac{dr(r-1)}{2} \tag{3.11}$$

between the dimension n of E, its rank r and the integer d. Any x of E can then be written $x = \sum_{i=1}^{r} x_i c_i + \sum_{i<j} x_{ij}$ with x_{ij} in E_{ij}. When E is $S_r(K)$, if $(e_1, \dots, e_r)$ is an orthonormal basis of $\mathbb{R}^r$, then $E_{ii} = \mathbb{R} e_i e_i^t$ and $E_{ij} = K(e_i e_j^t + e_j e_i^t)$ for $i < j$ and d is equal to $\dim_{|\mathbb{R}} K$. For the cone of revolution,

$$E_{11} = \mathbb{R}(1, x_1, \dots, x_{n-1}), \qquad E_{22} = \mathbb{R}(1, -x_1, \dots, x_{n-1})$$
$$\text{with } x_1^2 + \cdots + x_{n-1}^2 = 1,$$
$$E_{12} = \{(0, y_1, \dots, y_{n-1});\ x_1 y_1 + \cdots + x_{n-1} y_{n-1} = 0\} \quad \text{and} \quad d = n - 2.$$

Let us conclude with the following remark: the bilinear map $(x, y) \mapsto \operatorname{Trace}(x \cdot y)$ defines a scalar product satisfying (3.1) which is often chosen as canonical scalar product. [Actually this is the only one satisfying (3.1) up to a factor.] Observe then from (3.1) and (3.9) that, for x in E_+,

$$\langle x, x^{-1} \rangle = \operatorname{Trace}(e) = r. \tag{3.12}$$

We now define the Wishart distributions on $\overline{E}_+$, where E is a Euclidean simple Jordan algebra with structural constants n, r and d as defined in (3.11). Consider the set

$$\Lambda = \left\{0, \frac{d}{2}, d, \dots, \frac{d}{2}(r-1)\right\} \cup \left(\frac{d(r-1)}{2}, +\infty\right) \tag{3.13}$$

[if $E = S_r(\mathbb{R})$, this is (1.2)]. Let p be greater than 0. One can prove that there exists a μ_p in $\mathscr{M}(E)$ concentrated on $\overline{E}_+$ such that, for all θ in E_+, one has

$$\int_E \exp(-\operatorname{Trace}(\theta x))\, \mu_p(dx) = (\det \theta)^{-p} \tag{3.14}$$

if and only if p is in Λ [Gindikin (1975); a self-contained proof is in Casalis and Letac (1994)].

If $p > d(r-1)/2$, one has

$$\mu_p(dx) = C(p)(\det x)^{p-(n/r)} \mathbb{1}_{E_+}(x)\, dx,$$

where $C(p)$ is a constant depending on p and on the structural constants (d, r, n) of the Jordan algebra E. Here dx is the Lebesgue measure on E with the normalization naturally induced by the Euclidean structure of E.

If $p = dk/2$ with $k = 0, 1, \dots, r-1$, μ_p is a singular measure concentrated on $\overline{E}_+ \setminus E_+$. Observe from (3.14) that $\mu_{p+p'}$ is just the convolution $\mu_p * \mu_{p'}$. Note also that μ_p is a quasi-invariant measure for the group G: if $g \in G$ and $\theta \in E$, then

$$\det(g(\theta)) = (\det g)^{r/n} \det \theta \tag{3.15}$$

[see Faraut and Koranyi (1994), Proposition 3.4.3, p. 53]. [Note that in formula (3.15), $\det g$ is taken in the ordinary sense, since g is a linear transformation of E into itself, although $\det(g(\theta))$ and $\det \theta$ are taken in the sense of E.] From (3.14) and (3.15) we get that, for all p in Λ and all g in G, we have

$$g(\mu_p) = (\det g)^{-pr/n} \mu_p, \tag{3.16}$$

and this proves the quasi-invariance of μ_p by g.

For instance, if $E = S_r(\mathbb{R})$ and g is given by (3.4), then $\det g = (\det a)^{r+1}$ (a classical exercise). If $p > (r-1)/2$, then

$$\mu_p(dx) = C(p)(\det x)^{p-(r+1)/2} \mathbb{1}_{E_+}(x)\, dx,$$

but if $p \le (r-1)/2$, μ_p is concentrated on the boundary of the cone $\overline{E}_+$ of positive symmetric matrices, it is not easy to get an explicit form of the measure [see Uhlig (1994) for the case $S_r(\mathbb{R})$ and Casalis (1990) for the general case].

For p in Λ and for σ in E_+, we then define the Wishart distribution $\gamma_{p,\sigma}$ on E_+ by

$$\gamma_{p,\sigma}(dx) = \exp\left(-\operatorname{Trace}(x\sigma^{-1})\right)(\det \sigma)^{-p} \mu_p(dx). \tag{3.17}$$

The NEF $F_p = \{\gamma_{p,\sigma},\ \sigma \in E_+\}$ is characterized by its variance function defined on $M_F = E_+$ by

$$V_F(m) = \frac{1}{p} P(m),$$

where P is the quadratic representation of E. [Here again, as in (2.8), E and E^* are identified through the isomorphism $E \to E^*$: $a \mapsto \varphi_a$, where $\varphi_a(b) = \operatorname{Trace}(a \cdot b)$.] Recall here that if $E = S_r(\mathbb{R})$, then $P(m)(\theta) = m\theta m$.

We now come to the generalization of the Olkin–Rubin theorem to irreducible symmetric cones. Let us point out an important fact. One of the clever ideas of Olkin and Rubin (1962) is to find a right way to replace the ordinary division in real numbers which defines $Z = U/(U+V)$ in the Lukacs theorem. Although the product of symmetric matrices is not symmetric, they observe that several substitutes are available. Since $U + V$ is positive definite, one could write

(a) $$U + V = \sqrt{U+V}\sqrt{U+V},$$

(b) $$U + V = TT^t,$$

(c) $$U + V = T_1^t T_1,$$

where T and T_1 are unique lower-triangular matrices with positive diagonal. Then one can define possible Z as

(a) $$Z = (U+V)^{-1/2} U (U+V)^{-1/2},$$

(b) $$Z = T^{-1} U (T^{-1})^t,$$

(c) $$Z = {}^t T_1^{-1} U T_1.$$

Since an infinity of such algorithms are possible, we have restated Olkin and Rubin with the form given in Theorem 1.2. However, from (3.4), one can observe that, actually, considering a map $b \mapsto w(b)$ from E_+ to M, as we do in Theorem 1.2, and then considering maps

$$g(b)\colon u \mapsto (w(b))^{-1} u \big((w(b))^{-1}\big)^t$$

from E to itself is equivalent to considering a map $b \mapsto g(b)$ from E_+ to G. For this reason, we coin the following definition

DEFINITION. Let E be a Euclidean simple Jordan algebra. A *division algorithm* is a measurable map

$$E_+ \to G, \qquad b \mapsto g(b)$$

such that $g(b)(b) = e$ for all b in E_+.

On E, $g(b) = P(b^{-1/2})$ corresponds to the above algorithm (a). If $\{c_i\}_{i=1}^r$ is a complete system of primitive orthogonal idempotents, one can define a triangular subgroup T of G from the Peirce decomposition (3.10) [see Faraut and Koranyi (1994), p. 110, 111 and Theorem 6.3.6]. Thus the map $b \mapsto g(b) = t_b^{-1}$ gives the equivalent of the above algorithm (b).

We now state Theorem 1.2 for symmetric cones. For convenience we split its two parts (b) ⇒ (a) and (a) ⇒ (b) into two theorems. Actually, we expand a little bit on the first part (b) ⇒ (a): instead of dealing with a kind of Beta distribution of the first kind on symmetric cones, we deal with a kind of Dirichlet distribution, which slightly extends the one considered in Theorem 4.1 of Massam (1994). See Seber (1984) for a careful study of Dirichlet distributions on symmetric real matrices. Note also that Theorem 3.1 generalizes the first part of Theorem 7 of Uhlig (1994).

THEOREM 3.1. *Let E be a Euclidean simple Jordan algebra, σ in E_+, $p_0, p_1, \dots, p_m$ in Λ defined by* (3.13) *with $p = p_0 + p_1 + \cdots + p_m > d(r-1)/2$, and let $U_0, U_1, \dots, U_m$ be independent random variables valued in $\overline{E}_+$ with respective Wishart distributions $\gamma_{p_j, \sigma}$, $j = 0, \dots, m$. Write $S = U_0 + \cdots + U_m$ with distribution $\gamma_{p, \sigma}$. Let $b \mapsto g(b)$, $E_+ \to G$ be a division algorithm. Then the following hold*:

(i) *the distribution of*

$$Z = (Z_1, \dots, Z_m) = (g(S)(U_1), \dots, g(S)(U_m))$$

depends neither on the particular division algorithm nor on σ and is K-invariant, that is,

$$\mathscr{L}((k(Z_1), \dots, k(Z_m))) = \mathscr{L}((Z_1, \dots, Z_m))$$

for all k in K;

(ii) *Z and S are independent.*

Theorem 3.2 is the converse of Theorem 3.1 for $m = 1$.

THEOREM 3.2. *Let E be a Euclidean simple Jordan algebra with rank $r \geq 2$. Let U and V be independent non-Dirac random variables valued in $\overline{E}_+$ such that U and V are not concentrated on the same one-dimensional subspace and such that $U + V$ is almost surely in E_+. Let $b \mapsto g(b)$ be a division algorithm and consider*

$$Z = g(U + V)(U).$$

If the distribution of Z is K-invariant and if Z and $U + V$ are independent, then there exist p and q in Λ with $p + q > d(r-1)/2$, and σ in E_+ such that U and V have respective Wishart distributions $\gamma_{p,\sigma}$ and $\gamma_{q,\sigma}$.

4. Proof of Theorem 3.1. With the hypothesis of the Theorem 3.1, we first prove the following.

LEMMA 4.1. *Denote by $K_s(du_1, \dots, du_m)$ the conditional distribution $\mathscr{L}(U_1, \dots, U_m | S = s)$ and, for g in G, write $g^{(m)}$ for the action on E^m defined by $g^{(m)}(x_1, \dots, x_m) = (g(x_1), \dots, g(x_m))$. Then $g^{(m)} K_s = K_{g(s)}$ μ_p-almost everywhere.*

PROOF. Let $G\colon E \to \mathbb{R}$ and $F\colon E^m \to \mathbb{R}$ be any continuous functions with compact support. Then

$$\begin{aligned} I &= \int_{E^{m+1}} G(s) F(x_1, \dots, x_m) g^{(m)}(K_s)(dx_1, \dots, dx_m) \mu_p(ds) \\ &= \int_{E^{m+1}} G(s)(F \circ g^{(m)})(u_1, \dots, u_m) K_s(du_1, \dots, du_m) \mu_p(ds) \\ &= \int_{E^{m+1}} G(u_0 + u_1 + \cdots + u_m)(F \circ g^{(m)})(u_1, \dots, u_m) \mu_{p_0}(du_0) \cdots \mu_{p_m}(du_m). \end{aligned}$$

Now write $x_j = g(u_j)$, $j = 0, \dots, m$. We get

$$\begin{aligned} I &= \int_{E^{m+1}} G\big(g^{-1}(x_0 + \cdots + x_m)\big) F(x_1, \dots, x_m)(g\mu_{p_0})(dx_0) \cdots (g\mu_{p_m})(dx_m) \\ &= (\det g)^{-pr/n} \int_{E^{m+1}} G\big(g^{-1}(x_0 + \cdots + x_m)\big) F(x_1, \dots, x_m) \mu_{p_0}(dx_0) \\ &\quad \cdots \mu_{p_m}(dx_m) \end{aligned}$$

from the quasi-invariance (3.16). Thus I is

$$(\det g)^{-pr/n} \int_{E^{m+1}} G\big(g^{-1}(s)\big) F(x_1, \dots, x_m) K_s(dx_1, \dots, dx_m) \mu_p(ds).$$

Denoting $s' = g^{-1}(s)$, I is again

$$\begin{aligned} &(\det g)^{-pr/n} \int_{E^{m+1}} G(s') F(x_1, \dots, x_m) K_{g(s')}(dx_1, \dots, dx_m)\big(g^{-1}\mu_p\big)(ds') \\ &\qquad = \int_{E^{m+1}} G(s') F(x_1, \dots, x_m) K_{g(s')}(dx_1, \dots, dx_m) \mu_p(ds'), \end{aligned}$$

where we have used (3.16) a second time. Comparing this last expression with the definition of I, the lemma is proved. □

We can now prove Theorem 3.1. With the notation of Lemma 4.1, we shall prove that $\mathscr{L}(Z) = K_e$: this implies (i), since K_e is defined independently of the division algorithm and since e (and thus K_e, by the lemma) is K-invariant.

Let us take (as in the proof of the lemma) arbitrary continuous functions with compact support $G\colon E \to \mathbb{R}$ and $F\colon E^m \to \mathbb{R}$. Then

$$\begin{aligned} J &= \mathbb{E}\big(G(S) F(Z_1, \dots, Z_m)\big) \\ &= \int_E G(s) \exp\big[-\operatorname{Trace}(s\sigma^{-1})\big](\det \sigma)^{-p} \mu_p(ds) \\ &\quad \times \int_{E^m} F\big(g(s)^{(m)}(u_1, \dots, u_m)\big) K_s(du_1, \dots, du_m). \end{aligned}$$

Writing $z_j = g(s)(u_j)$, $j = 1, \dots, m$, in the last integral, using Lemma 4.1 and the fact that $g(s)(s) = e$ [since $s \mapsto g(s)$ is a division algorithm], we get

$$J = \int_E G(s)\exp\left[-\mathrm{Trace}(s\sigma^{-1})\right](\det\sigma)^{-p}\mu_p(ds) \times \int_{E^m} F(z_1, \dots, z_m)K_e(dz)$$
$$= \mathbb{E}(G(S))\mathbb{E}(F(Z_1, \dots, Z_m)).$$

This proves part (ii) and $\mathscr{L}(Z) = K_e$, thus part (i). □

The following extends Theorem 1 of Uhlig (1994).

COROLLARY 4.2. *With the hypothesis and notation of Theorem* 3.1, *let H be a random variable independent of Z with Wishart distribution $\gamma_{p,\sigma}$. Then, for $j = 1, \dots, n$, the $G_j = (g(H))^{-1}Z_j$ are independent with distribution $\gamma_{p_j,\sigma}$.*

PROOF. Since S is independent of Z, then $(G_1, \dots, G_n)$ and $((g(S))^{-1}Z_1, \dots, (g(S))^{-1}Z_n) = (U_1, \dots, U_n)$ are identically distributed. □

5. Proof of Theorem 3.2 (first part). As in the proof of Theorem 1.1, the proof here is split into two parts. The first one, also the easiest, is devoted to the analog of (2.10). To get it, we first fix θ in $-E_+$ and, for a in the space $L(E)$ of the linear endomorphisms of E and for ζ in E, we consider the two random variables

$$F(a) = \exp\left(\mathrm{Trace}\left(ag^{-1}(U+V)\right) + \langle\theta, U+V\rangle\right),$$
$$G(\zeta) = \exp\langle\zeta, Z\rangle.$$

The relation between the log Laplace transforms k_U and $\chi = k_{U+V}$ will be deduced from some expression linking $F(a)$, $F'(a)$, $G(\zeta)$, $G'(\zeta)$ together [relation (5.6)]. We first state the following lemma, which is essentially (5) in Olkin and Rubin (1962).

LEMMA 5.1. *If x is in E, then $k(x) = x$ for all k in K if and only if $x \in \mathbb{R}e$.*

PROOF. Write $x = \sum_{i=1}^r \lambda_i c_i$ the spectral decomposition of x in a complete system of primitive orthogonal idempotents $\{c_i\}_{i=1}^r$. Now, for all i in $\{2, \dots, r\}$, choose k_i in K inducing a permutation between $c_1, \dots, c_r$ such that $k_i c_1 = c_i$. Then, from the equality $k_i x = x$, we get $\lambda_1 = \lambda_i$ for any $i \geq 2$ and hence $x = \lambda_1 e$. The converse is trivial from (3.6). □

We now prove the existence of $E(G(\zeta))$ for all ζ in E an $E(F(a))$ for small a in $L(E)$.

One sees easily that $\|Z\|^2 \leq r$; for this, observe that if z and z' are in $\overline{E}_+$ with $z + z' = e$, then $z \cdot z + z \cdot z' = z$. Since $\langle z, z'\rangle = \mathrm{Trace}(z \cdot z') \geq 0$ from (3.3), this implies that

$$\|z\|^2 = \mathrm{Trace}(z \cdot z) \leq \mathrm{Trace}(z) \leq \mathrm{Trace}(e) = r,$$

and we apply this to $z = g(U+V)U$ and $z' = g(U+V)V$. Thus $\mathbb{E}(G(\zeta))$ exists for all ζ in E.

We now show that there exists a neighborhood of 0 in $L(E)$, depending on θ, such that $f(a) = \mathbb{E}(F(a))$ exists. To do so, we prove the following inequality. Defining $C = \max(1, \sqrt{d/2})$, we have for all x in E_+ and all a in $L(E)$,

$$\text{Trace}\big(ag^{-1}(x)\big) \leq C(\text{Trace}\, aa^*)^{1/2}\, \text{Trace}\, x. \tag{5.1}$$

[Recall that $x \mapsto g(x)$, $E_+ \to G$ is a division algorithm.] We equip $L(E)$ with the Euclidean structure

$$\langle a, a_1 \rangle_{L(E)} = \text{Trace}(aa_1^*).$$

Schwarz's inequality gives

$$\big(\text{Trace}\big(ag^{-1}(x)\big)\big)^2 \leq (\text{Trace}\, aa^*)\text{Trace}\big(g^{-1}(x)\big(g^{-1}(x)\big)^*\big).$$

Now we write $x = \sum_{i=1}^r \lambda_i c_i$, where $\{c_1, \dots, c_r\}$ is a complete system of primitive orthogonal idempotents and where $\lambda_i > 0$ (recall that x is in E_+). We get, from (3.8),

$$g^{-1}(x)g^{-1}(x)^* = g^{-1}(x)P(e)\big(g^{-1}(x)\big)^* = P\big(g^{-1}(x)(e)\big),$$

and this is simply $P(x)$, since $x \mapsto g(x)$ is a division algorithm. Now

$$\text{Trace}\, P(x) = \sum_{i=1}^r \lambda_i^2 + 2\frac{d}{2}\sum_{i<j}\lambda_i\lambda_j \leq C^2\left(\sum_{i=1}^r \lambda_i\right)^2,$$

as easily seen. Relation (5.1) is now proved. We take $\theta = -\sum_{i=1}^r \mu_i c_i$, with $\{c_1, \dots, c_r\}$ being a suitable complete system of primitive orthogonal idempotents with $0 < \mu_1 \leq \mu_2 \leq \cdots \leq \mu_r$. Thus, from (5.1) we get that if $\|a\| < C^{-1}\mu_1$, we have, for all x in E_+,

$$\text{Trace}\big(ag^{-1}(x)\big) + \langle \theta, x \rangle \leq (C\|a\| - \mu_1)\text{Trace}\,(x) < 0. \tag{5.2}$$

Replacing x by $U + V$, (5.2) shows that $\mathbb{E}(F(a))$ exists for $\|a\| < C^{-1}\mu_1$.

One beautiful idea of Olkin and Rubin is to consider the differential of $a \mapsto \mathbb{E}(F(a))$. Clearly, in $\{a \in L(E);\, \|a\| < C^{-1}\mu_1\}$, it is

$$\mathbb{E}(F'(a)) = \mathbb{E}\big(F(a)g^{-1}(U+V)\big). \tag{5.3}$$

[Here, we identify the space of linear maps from $L(E)$ to $\mathbb{R}$ to $L(E)$ itself by the bilinear map on $L(E) \times L(E)$: $(a, b) \mapsto \text{Trace}(ab)$.]

Similarly, the differential of $\zeta \mapsto \mathbb{E}(G(\zeta))$ is identified with an element of E, through the Euclidean structure of E, and is

$$\mathbb{E}(G'(\zeta)) = \mathbb{E}(G(\zeta)Z) \tag{5.4}$$

We now introduce the canonical bilinear map

$$\phi\colon L(E) \times E \to E, \qquad (f, h) \mapsto \phi(f, h) = f(h).$$

Using (5.3) and (5.4) and the fact that $U + V$ and Z are independent, and from the fact that

$$g^{-1}(U+V)(Z) = \phi\big(g^{-1}(U+V), Z\big) = U, \tag{5.5}$$

we get

$$\phi(\mathbb{E}(F'(a)),\mathbb{E}(G'(\zeta))) = \mathbb{E}(\phi(F'(a),G'(\zeta))) = \mathbb{E}(F(a)G(\zeta)U). \tag{5.6}$$

We now set $a = O_{L(E)}$ and $\zeta = O_E$ in (5.6). From (5.4), $\mathbb{E}(Z) = \mathbb{E}(G'(0))$ is an element of E which is K-invariant [since $\mathscr{L}(Z)$ is K-invariant by the hypothesis of Theorem 3.2]. Thus Lemma 5.1 implies that there exists c_1 in $\mathbb{R}$ such that $\mathbb{E}(G'(0)) = c_1 e$. Thus (5.6) gives

$$\begin{aligned}
&\mathbb{E}(\exp\langle\theta,U\rangle U)\mathbb{E}(\exp\langle\theta,V\rangle)\\
&\quad= c_1\phi(\mathbb{E}(F'(0)),e)\\
&\quad= c_1\mathbb{E}(\exp\langle\theta,U+V\rangle g^{-1}(U+V)(e)) \quad [\text{from } (5.3)]\\
&\quad= c_1\mathbb{E}(\exp\langle\theta,U+V\rangle(U+V))
\end{aligned}\tag{5.7}$$

[since $x \mapsto g(x)$ is a division algorithm].

If $L_U(\theta) = \mathbb{E}(\exp\langle\theta,U\rangle)$, $k_U = \log L_U$ (for θ in $-E_+$), and if L_V, k_V and $\chi = k_{U+V}$ are similarly defined, (5.7) yields

$$L_U k'_U L_V = c_1 L_U L_V \chi',$$

thus

$$k_U = c_1\chi \quad\text{and}\quad k_V = (1-c_1)\chi, \tag{5.8}$$

as in (2.10), since $k_U(0) = k_V(0) = \chi(0) = 0$.

6. Proof of Theorem 3.2 (second part). This second part is devoted to the analog of (2.11). The notations are those of §5. Using the second differential of $F(a)$ and $G(\zeta)$ as it is done on $\mathbb{R}$ would here give too complicated relations. On the other hand the simple computation of $\mathbb{E}(F(0)Q(U))$ for two special quadratic polynomials Q on E [cf. (6.9) and (6.13)] leads to the fundamental equality (6.17) equivalent to (2.11). From this point, as explained at the end of Section 2, we come back to the variance function of the NEF generated by $U + V$ to conclude. The following proposition is the version for symmetric cones of (6) in Olkin and Rubin (1962).

PROPOSITION 6.1. *Let f be a symmetric endomorphism of E such that $f = kfk^*$ for all k in K. Then there exists (λ, μ) in $\mathbb{R}^2$ such that*

$$f = \lambda\,\mathrm{id}_E + \mu e\otimes e,$$

where $e\otimes e$ denotes the endomorphism $x \mapsto e\langle x,e\rangle$.

PROOF. The result is trivial if $r = 1$, that is, $E = \mathbb{R}$. Thus we assume $r \geq 2$. Let E_0 be the orthogonal subspace of $\mathbb{R}e$ in E, and let x_0 be in $E_0 \setminus \{0\}$. Then $\{k(x_0);\ k\in K\}$ generates E_0.

If not, there exists x'_0 in $E_0\setminus\{0\}$ such that $\langle x'_0, k(x_0)\rangle = 0$ for all k in K. Let $\{c_i\}_{i=1}^r$ be a complete system of primitive orthogonal idempotents such that there exists $(\lambda'_i)_{i=1}^r$ in $\mathbb{R}^r$ with

$$x'_0 = \sum_{i=1}^r \lambda'_i c_i.$$

Since x_0' is in $E_0 \setminus \{0\}$ we have $\sum_{i=1}^r \lambda_i' = 0$ and there exist i and j such that $\lambda_i' \neq \lambda_j'$. Without loss of generality, we assume $\lambda_1' \neq \lambda_2'$. Since K acts transitively on the set of complete systems of primitive orthogonal idempotents, there exists k in K and $(\lambda_i)_{i=1}^r$ in $\mathbb{R}^r$ such that

$$k(x_0) = \sum_{i=1}^{r} \lambda_i c_i.$$

Here again, $\sum_{i=1}^r \lambda_i = 0$ and there exist i and j such that $\lambda_i \neq \lambda_j$. Since K includes the permutations between the $(c_i)_{i=1}^r$, without loss of generality we may again assume $\lambda_1 \neq \lambda_2$. Eventually, since $\langle x_0', k(x_0)\rangle = 0$ for all k in K, we get, in particular,

$$\lambda_1 \lambda_1' + \lambda_2 \lambda_2' + (\lambda_3 \lambda_3' + \cdots + \lambda_r \lambda_r') = 0,$$
$$\lambda_1 \lambda_2' + \lambda_2 \lambda_1' + (\lambda_3 \lambda_3' + \cdots + \lambda_r \lambda_r') = 0;$$

and the difference yields the contradiction $(\lambda_1 - \lambda_2)(\lambda_1' - \lambda_2') = 0$.

The second observation to be made is that e is an eigenvector of f. To see this, we write, for all k of K,

$$f(e) = (kfk^*)(e) = k(f(e)),$$

and Lemma 5.1 implies that there exists α in $\mathbb{R}$ such that $f(e) = \alpha e$.

Now, since f is symmetric, this last observation implies that E_0 is stable by f. Let F_λ be the eigenspace of f restricted to E_0 for the eigenvalue λ. Then F_λ is stable by K. To see this, we take h in F_λ and k in K. We obtain

$$\lambda h = f(h) = (k^* fk)(h).$$

Since $k^* = k^{-1}$, we have $\lambda k(h) = (fk)(h)$, and $k(h)$ is in F_λ. However, as we have seen, if $x_0 \in E_0 \setminus \{0\}$, then $\{k(x_0);\ k \in K\}$ generates E_0, that is, the only nonnull stable subspace of E_0 by K is E_0. We get $F_\lambda = E_0$, that is, f restricted to E_0 is $\lambda\,\mathrm{id}_E$. Finally, if $\pi\colon E \to E_0$ is the orthogonal projection, we write, for h in E,

$$h = \pi(h) + \frac{1}{r}\langle e, h\rangle e,$$

and we get

$$f(h) = \lambda\pi(h) + \frac{\alpha}{r}\langle e, h\rangle e = \lambda h + \frac{\alpha - \lambda}{r}\langle e, h\rangle e.$$

Thus $f = \lambda\,\mathrm{id}_E + [(\alpha - \lambda)/r] e \otimes e$ and the proof of Proposition 6.1 is complete. □

To apply Proposition 6.1, let us consider a quadratic polynomial Q on E and valued in the space $L_S(E)$ of symmetric endomorphisms on E. We further assume that Q is G-invariant, that is,

$$gQ(x)g^* = Q(gx) \quad \text{for all } (x, g) \text{ in } E \times G. \tag{6.1}$$

Two basic examples of such a Q are $Q(x) = x \otimes x\colon y \mapsto \langle x, y\rangle x$ (this is a standard fact of Euclidean spaces, that this Q is G-invariant) and $Q = P$, the quadratic map of the Jordan algebra E [see (3.8)].

For such a Q, since $\mathscr{L}(Z)$ is K-invariant, $f = \mathbb{E}(Q(Z))$ fulfills the hypothesis of Proposition 6.1, that is, there exists (λ_Q, μ_Q) in $\mathbb{R}^2$ such that

$$\mathbb{E}(Q(Z)) = \lambda_Q \operatorname{id}_E + \mu_Q e \otimes e. \tag{6.2}$$

We have now the following explicit result.

PROPOSITION 6.2. *Let Q: $E \to L_S(E)$ be a quadratic and G-invariant polynomial, and let λ_Q and μ_Q be defined by* (6.2). *Under the hypothesis of Theorem* 3.2, *we have the following for θ in $-E_+$, $c_1 e = \mathbb{E}(Z)$ and $\chi(\theta) = \log \mathbb{E}(\exp\langle \theta, U + V\rangle)$*:

$$c_1 Q\left(\frac{\partial}{\partial\theta}\right)\chi + c_1^2 Q(\chi') = \lambda_Q\left(P\left(\frac{\partial}{\partial\theta}\right)\chi + P(\chi')\right) + \mu_Q(\chi'' + \chi' \otimes \chi'). \tag{6.3}$$

PROOF. With the notation of Section 5 we write, for simplification, $F(0) = \exp\langle \theta, U + V\rangle$. Recall also that we have

$$P(\gamma(x)) = \gamma P(x)\gamma^* \quad \text{for all } (x, \gamma) \text{ in } E \times G. \tag{6.4}$$

We now have

$$\begin{aligned} \mathbb{E}\big(F(0)g^{-1}(U+V)(g^{-1}(U+V))^*\big) &\overset{(1)}{=} \mathbb{E}(F(0)P(U+V)) \\ &\overset{(2)}{=} e^{\chi}\left(P\left(\frac{\partial}{\partial\theta}\right)\chi + P(\chi')\right). \end{aligned} \tag{6.5}$$

In (6.5), (1) comes from (6.4) applied to $\gamma = g^{-1}(U+V)$ and to $x = e$, and from $P(e) = \operatorname{id}_E$; (2) is standard, since P is quadratic.

To obtain (6.3), we now compute $\mathbb{E}(F(0)Q(U))$ in two ways. We have first

$$\begin{aligned} \mathbb{E}(F(0)Q(U)) &\overset{(1)}{=} \mathbb{E}\big(F(0)Q(g^{-1}(U+V)(Z))\big) \\ &\overset{(2)}{=} \mathbb{E}\big(F(0)g^{-1}(U+V)Q(Z)(g^{-1}(U+V))^*\big) \\ &\overset{(3)}{=} \mathbb{E}\big(F(0)g^{-1}(U+V)\mathbb{E}(Q(Z))(g^{-1}(U+V))^*\big) \\ &\overset{(4)}{=} \lambda_Q \mathbb{E}\big(F(0)g^{-1}(U+V)(g^{-1}(U+V))^*\big) \\ &\qquad + \mu_Q \mathbb{E}(F(0)(U+V) \otimes (U+V)) \\ &\overset{(5)}{=} e^{\chi}\left[\lambda_Q\left(P\left(\frac{\partial}{\partial\theta}\right)\chi + P(\chi')\right) + \mu_Q(\chi'' + \chi' \otimes \chi')\right]. \end{aligned} \tag{6.6}$$

In (6.6), (1) comes from the definition of Z, (2) comes from the G-invariance (6.1) of Q, (3) comes from the independence of $U + V$ and Z, (4) is (6.2) and (5) is (6.5).

The second way is easier:

$$
\begin{aligned}
\mathbb{E}(F(0)Q(U)) &\overset{(1)}{=} \mathbb{E}(e^{\langle \theta, V\rangle})\mathbb{E}(e^{\langle \theta, U\rangle}Q(U))\\
&\overset{(2)}{=} \mathbb{E}(e^{\langle \theta, V\rangle})\left(Q\left(\frac{\partial}{\partial\theta}\right)k_{\mu} + Q(k'_U)\right)\mathbb{E}(e^{\langle \theta, U\rangle})\\
&\overset{(3)}{=} e^{\chi}\left(c_1 Q\left(\frac{\partial}{\partial\theta}\right)\chi + c_1^2 Q(\chi')\right).
\end{aligned}
\tag{6.7}
$$

In (6.7), (1) comes from the independence of U and V, (2) comes from the quadratic character of Q and (3) comes from (5.8). Combining (6.6) and (6.7), we get (6.3). □

We now apply (6.3) to $Q(x) = x \otimes x$ and to $Q = P$. Keeping the notation of Olkin and Rubin [(1962), formula (14)], we write

$$
\mathbb{E}(Z \otimes Z) = c_3 \,\mathrm{id}_E + c_2 e \otimes e,
\tag{6.8}
$$

that is, if $Q(x) = x \otimes x$, $\lambda_Q = c_3$ and $\mu_Q = c_2$. In this case, formula (6.3) becomes

$$
c_1 \chi'' + c_1^2 \chi' \otimes \chi' = c_3 P\left(\frac{\partial}{\partial\theta}\right)\chi + c_3 P(\chi') + c_2 \chi'' + c_2 \chi' \otimes \chi'.
\tag{6.9}
$$

Now, to apply (6.3) to $Q = P$, we have to compute $\mathbb{E}(P(Z))$ with respect to c_3 and c_2 as defined by (6.2), that is, to compute λ_P and μ_P in the sense of (6.2). For this, we need a lemma in linear algebra.

LEMMA 6.3. *Let E be a Euclidean space, and let Q be the space of homogeneous quadratic polynomials on E valued in the space $L_S(E)$ of symmetric endomorphisms of E. Let $L(L_S(E))$ be the space of endomorphisms of $L_S(E)$. Then there exists a unique map*

$$
Q \mapsto \psi_Q, \qquad Q \to L(L_S(E))
$$

such that, for all x in E, one has

$$
\psi_Q(x \otimes x) = Q(x).
\tag{6.10}
$$

Furthermore, if j: $E \to \mathbb{R}$ is a C^2 function, one has

$$
Q\left(\frac{\partial}{\partial\theta}\right)j = \psi_Q(j'').
\tag{6.11}
$$

Finally, if E is a Euclidean simple Jordan algebra and if $Q = P$ is the quadratic map, then

$$
\psi_P(\mathrm{id}_E) = \left(1 - \frac{d}{2}\right)\mathrm{id}_E + \frac{d}{2} e \otimes e.
\tag{6.12}
$$

PROOF. Here, the best way is simply to start from an orthonormal basis $(e_1, \dots, e_n)$ of E. Thus

$$
\left(e_i \otimes e_i,\ i = 1, \dots, n;\ e_i \otimes e_j + e_j \otimes e_i,\ 1 \le i < j \le n\right)
$$

is a basis of $L_S(E)$. We define ψ_Q by

$$\psi_Q(e_i \otimes e_i) = Q(e_i) \qquad \text{for } i = 1, \dots, n,$$

$$\psi_Q(e_i \otimes e_j + e_j \otimes e_i) = Q(e_i + e_j) - Q(e_i) - Q(e_j) \quad \text{for } 1 \le i < j \le n.$$

Clearly, this proves the existence and uniqueness of ψ_Q, and (6.11) is easily proved by watching the coordinates and the Hessian matrix of j.

To prove (6.12), we observe that, for all (x, g) in $E \times G$,

$$\psi_P(g(x \otimes x)g^*) = \psi(g(x) \otimes g(x)) = P(g(x)) = gP(x)g^*.$$

Thus, by linearity, since $\{x \otimes x;\ x \in E\}$ generates $L_S(E)$, we get, for all (f, g) in $L_S(E) \times G$,

$$\psi_P(gfg^*) = g\psi_P(f)g^*.$$

Taking $f = \mathrm{id}_E$ and g in K, and using Proposition 6.1, there exists (λ, μ) in $\mathbb{R}^2$ such that $\psi(\mathrm{id}_E) = \lambda\,\mathrm{id}_E + \mu e \otimes e$. To compute (λ, μ) we start from the Peirce decomposition (3.10) associated to a complete system of primitive orthonormal idempotents $\{c_1, \dots, c_r\}$. Denote by $(c_{ij}^k)_{k=1}^d$ an orthonormal basis of E_{ij} for $1 \le i < j \le r$. Then

$$\mathrm{id}_E = \sum_{i=1}^r c_i \otimes c_i + \sum_{k=1}^d \sum_{1 \le i < j \le r} c_{ij}^k \otimes c_{ij}^k.$$

Thus $\psi_P(\mathrm{id}_E)$ applied to c_1 gives

$$\begin{aligned}\psi_P(\mathrm{id}_E)(c_1) &= \sum_{i=1}^r P(c_i)(c_1) + \sum_{k=1}^d \sum_{1 \le i < j \le r} P\big(c_{ij}^k\big)(c_1)\\ &= c_1 + \sum_{k=1}^d \sum_{1 < j \le r} \frac{1}{2} c_j = \left(1 - \frac{d}{2}\right)c_1 + \frac{d}{2}e.\end{aligned}$$

For the computation, we have used the relations $c_{ij}^k \cdot c_1 = (\delta_{1i} + \delta_{1j})\frac{1}{2}c_{ij}^k$, $c_{ij}^k \cdot c_{ij}^k = \frac{1}{2}(c_i + c_j)$ and $c_i \cdot c_j = \delta_{ij}c_i$ [see Faraut and Koranyi (1994), Proposition 4.1.4, p. 65 and Theorem 4.2.1, p. 68]. Thus $\lambda = 1 - d/2$ and $\mu = d/2$. $\square$

This lemma enables us to compute λ_P and μ_P as follows. We have, from Lemma 6.3, with $G(\zeta) = \exp\langle \zeta, Z\rangle$,

$$\mathbb{E}(G(\zeta)P(Z)) = P\left(\frac{\partial}{\partial \zeta}\right)\mathbb{E}(G(\zeta)) = \psi_P(\mathbb{E}(G''(\zeta))).$$

Setting $\zeta = 0$ in this equality, we obtain

$$\mathbb{E}(P(Z)) = \psi_P(\mathbb{E}(Z \otimes Z)) = \psi_P(c_3\,\mathrm{id}_E + c_2 e \otimes e),$$

from (6.8). Using (6.12) we get

$$\mathbb{E}(P(Z)) = \left(c_3\left(1 - \frac{d}{2}\right) + c_2\right)\mathrm{id}_E + c_3\frac{d}{2}e \otimes e;$$

thus $\lambda_P = c_3(1 - d/2) + c_2$ and $\mu_P = c_3 d/2$.

We can now apply Proposition 6.2 to $Q = P$, yielding

$$
\begin{aligned}
&c_1 P\left(\frac{\partial}{\partial\theta}\right)\chi + c_1^2 P(\chi') \\
(6.13)\qquad &= \left(c_3\left(1 - \frac{d}{2}\right) + c_2\right)\left(P\left(\frac{\partial}{\partial\theta}\right)\chi + P(\chi')\right) + c_3\frac{d}{2}(\chi'' + \chi' \otimes \chi').
\end{aligned}
$$

We now put together identities (6.9) and (6.13) in order to eliminate $P(\partial/\partial\theta)\chi$. We skip this little computation. Writing, for simplicity,

$$c_4 = c_1 - c_2 + c_3\frac{d}{2},$$

we finally get the essential identity

$$
\begin{aligned}
(6.14)\qquad c_4(c_1 - c_2 - c_3)\chi'' &= c_3 c_1(1 - c_1)P(\chi') \\
&\quad + \left(c_4\left(c_2 - c_1^2\right) + c_3\left(c_4 + c_1^2 - c_2\right)\right)\chi' \otimes \chi'.
\end{aligned}
$$

We now spend some effort to prove that in (6.14) neither the coefficient $c_4(c_1 - c_2 - c_3)$ of χ'' nor the coefficient $c_3 c_1(1 - c_1)$ of $P(\chi')$ is 0.

First $c_1(1 - c_1) = 0$ is impossible. We have that

$$(6.15)\qquad Z = 0 \quad \Leftrightarrow \quad U = 0 \quad \text{and} \quad Z = e \quad \Leftrightarrow \quad V = 0.$$

Thus $c_1 = 0$ implies $\mathscr{L}(U) = \delta_0$, and $c_1 = 1$ implies $\mathscr{L}(V) = \delta_0$: this contradicts the hypothesis of Theorem 3.2.

We now show that $c_3 = c_1 - c_2 = 0$ is impossible. Assume the contrary; then

$$\mathbb{E}(Z \otimes Z) = c_1 e \otimes e \quad \text{and} \quad \mathbb{E}(Z) = c_1 e$$

imply that there exists a real random variable z such that $Z = ze$ [write $Z = ze + Z_0$, with Z_0 orthogonal to e, and observe that $\mathbb{E}(Z_0 \otimes Z_0) = 0$ and thus $Z_0 = 0$]. Since Z and $e - Z$ are in $\overline{E}_+$, we have $0 \le z \le 1$. Furthermore, $\mathbb{E}(z) = \mathbb{E}(z^2) = c_1$ implies $\mathbb{E}(z(1 - z)) = 0$ and $\mathscr{L}(z) = (1 - c_1)\delta_0 + c_1\delta_1$. From (6.15) we deduce that $\{U = 0\}$ and $\{V = 0\}$ are complementary events, both with positive probability (since $0 < c_1 < 1$): this contradicts the independence of U and V.

If now $c_4(c_1 - c_2 - c_3) = 0$, then $c_3 \neq 0$. If not, we have $c_3 = 0$ and $c_4(c_1 - c_2 - c_3) = (c_1 - c_2)^2 = 0$; as we have just seen, this is an impossibility. Thus $c_3 c_1(1 - c_1) \neq 0$ and from (6.14) there exists some α in $\mathbb{R}$ such that $P(\chi') = \alpha\chi' \otimes \chi'$. However, since $U + V$ is in E_+, χ' is also in E_+, and $P(\chi')$ is invertible and cannot be proportional to $\chi' \otimes \chi'$, which has rank 1. Thus $c_4(c_1 - c_2 - c_3) \neq 0$.

If now $c_3 = 0$, from (6.14) there exists β in $\mathbb{R}$ such that $\chi'' = \beta\chi' \otimes \chi'$. Since $\chi''(\theta)$ is the covariance of

$$(6.16)\qquad P_\theta(dx) = \exp(\langle\theta, x\rangle - \chi(\theta))\mathscr{L}_{U+V}(dx)$$

and since $\chi'(\theta)$ is its mean, this implies that P_θ is concentrated on $\mathbb{R}\chi'(\theta)$, and thus $\mathscr{L}(U + V)$ is concentrated on $\mathbb{R}\chi'(\theta)$ for all θ in $-E_+$, so that there exists x_0 in $E \setminus \{0\}$ such that $\mathscr{L}(U + V)$ is concentrated on $\mathbb{R}x_0$. Since U and V are independent, there exist x_1 in E and real independent random variables (u, v) such that $U = x_1 + ux_0$ and $V = -x_1 + vx_0$. Furthermore, since $c_3 = 0$, then $Z = ze$ for some real random variable z. Since $ze = g(U+V)(U)$, we have $z(U+V) = U$, and this implies, since $z \not\equiv 0$, that $x_1 = 0$; since U and V are now concentrated on $\mathbb{R}x_0$, the hypothesis of Theorem 3.2 is contradicted.

Thus the coefficients of χ'' and $P(\chi')$ in (6.14) are not zero, and we can claim that there exist λ in $\mathbb{R} \setminus \{0\}$ and β in $\mathbb{R}$ such that, for all θ in $-E_+$,

$$\chi''(\theta) = \frac{1}{\lambda} P(\chi'(\theta)) + \beta \chi'(\theta) \otimes \chi'(\theta). \tag{6.17}$$

We deduce first from (6.17) that $\mathscr{L}(U+V)$ is not concentrated on any affine hyperplane of E. Since $\chi'(\theta)$ is valued in E_+, $P(\chi'(\theta))$ is positive definite on E. Since $\chi' \otimes \chi'$ has rank 1, and since $\chi''(\theta)$ is positive, $1/\lambda > 0$. Now suppose that $\chi''(\theta)$ is positive definite for no θ in $-E_+$. Then (6.17) implies that P_θ, as defined by (6.16), is concentrated on $\chi'(\theta)^\perp$, the orthogonal complement of $\chi'(\theta)$, that is,

$$\langle x - \chi'(\theta), \chi'(\theta) \rangle = 0, \quad P_\theta\text{-almost everywhere.}$$

This implies that, for all θ and θ_0 in $-E_+$, we have $\langle x - \chi'(\theta), \chi'(\theta)\rangle = 0$, P_{θ_0}-a.e. Integrating this, we get $\langle \chi'(\theta_0) - \chi'(\theta), \chi'(\theta)\rangle = 0$. This leads by symmetry to $\langle \chi'(\theta_0) - \chi'(\theta), \chi'(\theta_0) - \chi'(\theta)\rangle = 0$ and $\theta \mapsto \chi'(\theta)$ is a constant: this contradicts (6.17). Then there exists θ_0 in $-E_+$ such that $\chi''(\theta_0)$ is positive definite, and $\mathscr{L}(U+V)$ is not concentrated on any affine hyperplane of E. Therefore, the natural exponential family F generated by $\mathscr{L}(U+V)$ exists. If M_F is its domain of the means ($\subset E_+$) and if V_F is its variance function, (6.17) can be rewritten as follows: for all m in M_F,

$$V_F(m) = \frac{1}{\lambda} P(m) + \beta(m \otimes m). \tag{6.18}$$

We now use (6.18) to prove that $\beta = 0$.

For this, we consider the set $\mathscr{V}$ of analytic maps $W\colon M_F \to L_S(E)$ such that, for all (m, x, y) in $M_F \times E^2$,

$$W'(m)(W(m)(x))(y) - W'(m)(W(m)(y))(x) = 0. \tag{6.19}$$

As we have seen in the basic Proposition 2.1, V_F is in $\mathscr{V}$. Since, for p in (3.13), $E_+ \to L_S(E)$ $m \mapsto (1/p)P(m)$ is the variance function of a Wishart family, Proposition 2.1 shows also that the restriction of P to M_F also belongs to $\mathscr{V}$. Finally, it is easy to check that $Q\colon m \mapsto m \otimes m$, restricted to M_F, is also in $\mathscr{V}$, since

$$Q'(m)(y) = m \otimes y + y \otimes m.$$

However, condition (6.19) is quadratic in W, not linear. This implies that $\mathscr{V}$ is not a vector space. Since P, Q and $V_F = (1/\lambda)P + \beta Q$ are in $\mathscr{V}$, condition (6.19), by polarization, implies that, for all (m, x, y) in $M_F \times E^2$,

$$\begin{aligned} \frac{\beta}{\lambda}\big[& P'(m)(Q(m)(x))(y) + Q'(m)(P(m)(x))(y) \\ & - P'(m)(Q(m)(y))(x) - Q'(m)(P(m)(y))(x)\big] = 0. \end{aligned} \tag{6.20}$$

Since the first member of (6.20) is a polynomial in m, this identity (6.20) is even true for (m, x, y) in E^3. We now take $m = e$ in (6.20) [recall that $(P'(e)(a))(b) = 2a \cdot b$] and obtain

$$\frac{\beta}{\lambda}(\langle e, x\rangle y - \langle e, y\rangle x) = 0,$$

which implies $\beta = 0$ as soon as $\dim E > 1$. Thus

$$V_F(m) = \frac{1}{\lambda} P(m).$$

Since V_F is a variance function, from Gindikin's theorem, λ must belong to (3.13), and F is a Wishart family with Jorgensen's parameter λ. The proof of Theorem 3.2 is now complete. □

7. Comments. Olkin and Rubin (1962) is an extraordinarily clever paper. We comment on certain parts of it.

1. The first bright idea is to realize that there exist several division algorithms in the space $S_r(\mathbb{R})$ of (r, r) symmetric matrices, and one does better to choose none of them in particular. However, they never make it clear that the factorization $WW^t = U + V$ should depend on $U + V$ only, that is, W has to be $(U + V)$-measurable. (For instance, if Γ is orthogonal such that $\Gamma^t U \Gamma$ is diagonal and $W = \sqrt{U + V}\,\Gamma$, then $WW^t = U + V$ but $Z = W^{-1}U(W^{-1})^t$ is not independent of $U + V$.)
2. The hypothesis of Theorem 1 about (n, m) is unclear. A proof of Theorem 1 appears in Olkin and Rubin (1964) using Jacobians and therefore assuming $n \geq p - 1$ and $m \geq p - 1$; presumably this is also assumed in the 1962 paper (our Theorem 3.1 actually proves it for $n + m \geq p - 1$ only).
3. The second bright idea is to include invariance of $\mathscr{L}(Z)$ by maps $Z \mapsto \Gamma Z \Gamma^t$, when Γ is orthogonal. This leads them to the essential Lemma 1, which we have imitated and split in two parts in our Lemma 5.1 and Proposition 6.1.
4. The statement of Theorem 2 is not quite correct without the assumption that $\mathscr{L}(U)$ and $\mathscr{L}(V)$ should not be concentrated on the same line $\mathbb{R}X_0$, where X_0 is a symmetric positive definite matrix. Actually, in this case, with
$$Z = \left(\sqrt{U + V}\right)^{-1} U \left(\sqrt{U + V}\right)^{-1}, \quad \text{say},$$
and $U = uX_0$ and $V = vX_0$ such that u and v are independent and gamma distributed, then Z and $U + V$ are independent and $\mathscr{L}(U)$ and $\mathscr{L}(V)$ are not Wishart.
5. The third bright idea is to consider the Laplace transform (2) of $(W, U + V)$ [we have adapted the idea in our Section 5 by considering $\mathbb{E}(F(a))$]. Note that the existence of $f(A, B)$ on a domain big enough to allow differentiation is not mentioned [our Section 5 between (5.1) and (5.2) addresses this problem].

6. Their formulae (11), (14) and (16) have lead us to our formula (6.9). However, we have not been able to understand, then to follow and to adapt, the remainder of their paper after their Lemma 2, and we have chosen the different and shorter road which computes not only $\mathbb{E}(Z \otimes Z)$ but also $\mathbb{E}(P(Z))$ and gives (6.13). Techniques inherited from exponential families lead us from there to the result.
7. Their Appendix contains a proof of the Gindikin theorem for $S_r(\mathbb{R})$. We have commented elsewhere [Casalis and Letac (1994)] on this appendix, explaining why this proof did not convince us.

REFERENCES

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

CASALIS, M. (1990). Familles exponentielles naturelles invariantes par un sous groupe affine. Thèse, Laboratoire de Statistique et Probabilités, Univ. Paul Sabatier, Toulouse.

CASALIS, M. (1991). Les familles exponentielles à variance quadratique homogène sont des lois de Wishart sur un cône symétrique. *C. R. Acad. Sci. Paris Sér. I Math.* **312** 537–540.

CASALIS, M. and LETAC, G. (1994). Characterization of the Jorgensen set in generalized linear models. *Test* **3** 145–162.

FARAUT, J. and KORANYI, A. (1994). *Analysis on Symmetric Cones*. Clarendon Press, Oxford.

GINDIKIN, S. G. (1975). Invariant generalized functions in homogeneous domains. *Functional Anal. Appl.* **9** 50–52.

LETAC, G. (1989). A characterization of the Wishart exponential families by an invariance property. *J. Theoret. Probab.* **2** 71–86.

LETAC, G. (1994). Les familles exponentielles statistiques invariantes par les groupes du cône et du paraboloïde de révolution. *J. Appl. Probab.* **31A** 71–95. (Special volume, *Studies in Probability*, edited by J. Galambos and J. Gani.)

LETAC, G. and MASSAM, H. (1995). Craig-Sakamoto's theorem for the Wishart distributions on symmetric cones. *Ann. Inst. Statist. Math.* **47** 785–799.

LUKACS, E. (1955). A characterization of the gamma distribution. *Ann. Math. Statist.* **26** 319–324.

MASSAM, H. and NEHER, E. (1996). On transformations and determinants of Wishart variables on symmetric cones. *J. Theoret. Probab.* To appear.

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.

MUIRHEAD, R. (1982). *Aspects of Multivariate Analysis*. Wiley, New York.

OLKIN, I. and RUBIN, H. (1962). A characterization of the Wishart distribution. *Ann. Math. Statist.* **33** 1272–1280.

OLKIN, I. and RUBIN, H. (1964). Multivariate beta distributions and independence properties of the Wishart distribution. *Ann. Math. Statist.* **35** 261–269.

SEBER, G. A. F. (1984). *Multivariate Observations*. Wiley, New York.

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

LABORATOIRE DE STATISTIQUE
ET PROBABILITÉS
UNIVERSITÉ PAUL SABATIER
118, ROUTE DE NARBONNE
31062 TOULOUSE CEDEX
FRANCE