Interpolation of the Wishart and non central Wishart distributions

Gérard Letac*

Angers, June 2016

*Laboratoire de Statistique et Probabilités, Université Paul Sabatier, 31062 Toulouse, France

Contents

I Introduction

The aim of these notes is to present with the most natural generality the Wishart and non central Wishart distributions on symmetric real matrices, on Hermitian matrices and sometimes on a symmetric cone. While the classical χ^2 square distributions are familiar to the statisticians, they quickly learn that these χ^2 laws as parameterized by n their degree of freedom, can be interpolated by the gamma distributions with a continuous shape parameter. A less familiar object is the non central chi square distributions obtained by considering the distribution of the sum of squares of n independent Gaussian real random variables with the same variance, but not with zero mean. Here again the integer n can be interpolated by a continuous parameter.

An other direction of generalization is the consideration of n iid random variables $Z_i \sim N(0, \Sigma)$ valued in $\mathbb{R}^d$. Writing Z_i as a column vector and Z_i^T as a line vector, then $Z_i Z_i^T$ is the simplest symmetric random matrix of order d and the distribution of $\sum_{i=1}^n Z_i Z_i^T$ is called a classical Wishart distribution with parameters depending on the parameters n and Σ. Suppose now that the independent Gaussian random variables Z_i's are not centered anymore, but that $Z_i \sim N(m_i, \Sigma)$ (the fact that the m_i's in $\mathbb{R}^d$ depend possibly on i is important). Then the distribution of $\sum_{i=1}^n Z_i Z_i^T$ i is called non central Wishart. Classical Wishart and non central Wishart are quite useful to practitionners, and problems of eigenvalues and moments have been widely developped in the literature.

However, we are going to push further the study of these objects by working on the problem of their interpolation as it has been done in the one dimensional case. While the Wishart distribution has been invented by John Wishart in 1929 and simplified by Maurice Bartlett in 1939, it happens that the problem of interpolation has been pending until 1975 where it was solved by Gindikin who is an analyst. Statisticians were not aware of his paper and Shanbhag gave in 1989 an another solution quite elementary and elegant. Not knowing these papers, Peddada and Richards gave an other proof in 1991. The case of the non central Wishart is even more recent, and E. Mayerhofer in 2013 has conjectured the proper necessary and sufficient conditions of the existence of these interpolated non central Wishart distributions. We will show here that his conjecture is indeed correct.

II Wishart distributions in the classical sense

The Gaussian distribution on $\mathbb{R}$ with mean m and variance $v > 0$ is the probability distribution

$$N_{m,v}(dz) = \frac{1}{\sqrt{(2\pi v)}} e^{-(z-m)^2/2} \, dz. \tag{2.1}$$

Similarly, the gamma distribution on $[0, +\infty)$ with shape parameter $p > 0$ and scale parameter σ is

$$\gamma_{p,\sigma}(du) = \exp(-\sigma^{-1}u)\sigma^{-p}u^{p-1}1_{(0,+\infty)}(u)(\Gamma(p))^{-1} \, du \tag{2.2}$$

Its Laplace transform is, for $1 - \theta\sigma > 0$,

$$\int_0^\infty e^{\theta u} \gamma_{p,\sigma}(du) = (1 - \theta\sigma)^{-p} \tag{2.3}$$

Note here that the image of $\gamma_{p,\sigma}$ by the map $u \mapsto u/\sigma$ is $\gamma_{p,1}$. If Z has distribution $N_{0,v}$, it is easily seen that the distribution of Z^2 is $\gamma_{1/2,2v}$: just compute the integral

$$\int_{\mathbb{R}} e^{\theta z^2} N_{0,v}(dz)$$

and use (1.3) and the fact that the Laplace transform characterizes the measure.

Consider now n independent real random variables $Z_1, ..., Z_n$, with the same distribution $N_{0,v}$, as well as the random variable $X_n = Z_1^2 + ... + Z_n^2$. The distribution of a sum of independent random variables is the convolution of the distributions of each, the Laplace transform of a convolution is the product of Laplace transforms; thus, using (1.3) again, we see that the distribution of X_n is $\gamma_{n/2,2v}$. For $v = 1$, this distribution is called the " χ^2 distribution with n degrees of freedom " by the statisticians. They use it in many circumstances, the simplest one being the following: suppose that for good reasons, you think that the sequence of n independent real random variables $Z_1, ..., Z_n$ that you have observed has common distribution $N_{m,v}$ where m is unknown, you estimate it by declaring that m is close to $\overline{Z}_n = (Z_1 + ... + Z_n)/n$, called the empirical mean. Furthermore, some linear algebra shows if the random variable, called the " empirical variance " is

$$S_n^2 = \frac{1}{n-1}((Z_1 - \overline{Z}_n)^2 + ... + (Z_n - \overline{Z}_n)^2), \tag{2.4}$$

then $(n-1)S_n^2/\sqrt{v}$ is χ^2 distributed with $n-1$ degrees of freedom. If the hypothesis that the distribution of the Z_j is $N_{m,v}$ for some m happens to be false, the size of the empirical variance will be disquietingly large, as it can be checked from tables of χ^2 distributions.

Now, we are going to extend the above process to several dimensions, where random variables are taken in a linear space V with dimension d. To simplify the matters, we assume that V has a Euclidean structure with scalar product $\langle z, z' \rangle$. We denote by E the space of symmetric endomorphisms of V, by E_+ the open cone of positive definite ones and by $\overline{E}_+$ its closure. E has also a natural Euclidean structure defined by the scalar product: $(x, y) \mapsto \text{Trace } xy = \langle x, y \rangle$.

If Σ is in $\overline{E}_+$, and if m is in V, the Gaussian distribution $N_{m,\Sigma}$ on V is shortly defined by its Laplace transform: for all θ in V

$$\int_V e^{\langle \theta, z \rangle} N_{m,\Sigma}(dz) = e^{\frac{1}{2}\langle \theta, \Sigma\theta \rangle + \langle \theta, m \rangle}.$$

Diagonalization of Σ shows that $N_{m,\Sigma}$ is a product of one dimensional Gaussian distributions of type (1.1). Assume from now to the end of the section that one has Σ in E_+; the explicit form is

$$N_{m,\Sigma}(dz) = (2\pi)^{-r/2}(\det \Sigma)^{-1/2} e^{-\langle z-m, \Sigma^{-1}(z-m) \rangle/2}\, dz. \tag{2.5}$$

Let us adopt the following notation : If a and b are in V, we consider the endomorphism of V: $z \mapsto (a \otimes b)(z) = a\langle b, z \rangle$. Thus $a \otimes a$ is in E. Consider now the random variable Z valued in V with distribution $N_{0,\Sigma}$. The random variable $X = Z \otimes Z$ is therefore valued in E. Actually, it is even valued in $\overline{E}_+ \setminus E_+$; it is our main character and its distribution is said to be Wishart with shape parameter $1/2$ and scale parameter 2Σ. We again denote it by $\gamma_{1/2,2\Sigma}$ (Here, for sake of future generalization, we have parted from the traditional notation and conventions of the statistical literature). Its Laplace transform is easily computed: taking $-\theta$ in E_+, we get:

$$\int_E e^{\langle\theta,x\rangle}\gamma_{1/2,2\Sigma}(dx) = (\det(\mathrm{id}_V - 2\Sigma\theta))^{-1/2}. \tag{2.6}$$

For Σ in E_+, the proof of (1.6) is obtained by observing that the trace of $\theta(Z\otimes Z)$ is $\langle Z,\theta(Z)\rangle$, and then by using the fact that (1.5) has mass 1. Use a density argument for the general case.

More generally, denote by $\gamma_{n/2,2\Sigma}$ the distribution of $Z_1\otimes Z_1 + ... + Z_n\otimes Z_n$, when $Z_1,...,Z_n$ are independent with the same distribution $N_{0,\Sigma}$. Thus $\gamma_{n/2,2\Sigma}$ is a convolution and, from (1.6), its Laplace transform is equal, for $-\theta$ in E_+, to $(\det(\mathrm{id}_V - 2\Sigma\theta))^{-n/2}$. This is the Wishart distribution with shape parameter $n/2$ and scale parameter 2Σ. One has to point out that, for $n<r$, $Z_1\otimes Z_1 + ... + Z_n\otimes Z_n$ has at most rank n, and its distribution cannot have a density with respect to the Lebesgue measure of E. For $n\geq r$, one has

$$\gamma_{n/2,2\Sigma}(du) = e^{-\langle\Sigma^{-1},u\rangle/2}(\det\ u)^{(n-1-r)/2}(\det(2\Sigma))^{-n/2}(\Gamma_r(n/2))^{-1}1_{E_+}du, \tag{2.7}$$

where du is the Lebesgue measure of the Euclidean space E and $\Gamma_r(n/2)$ is a constant with respect to u and Σ. To prove (1.7), from (1.6), enough is to show that for $-\theta$ in E_+ one has:

$$\int_E e^{\langle\theta,u\rangle}e^{-\langle\Sigma^{-1},u\rangle/2}(det\ u)^{(n-1-r)/2}(\det(2\Sigma))^{-n/2}(\Gamma_r(n/2))^{-1}1_{E_+}du = (\det(\mathrm{id}_V - 2\Sigma\theta))^{-n/2}. \tag{2.8}$$

The proof of (1.8) is not elementary: one has for instance to fix an orthonormal basis e of V, to write the representative matrix $[u]_e^e = U = T^*DT$, with D diagonal and T upper triangular with 1 on the diagonal, to compute the Jacobian of the map $(D,T)\mapsto U$ and to perform the integration in these new coordinates (D,T). This is sometimes called the Bartlett method (see Muirhead (1982)).

Similarly to (1.4), the random symmetric endomorphism of V equal to the "empirical covariance":

$$S_n^2 = \frac{1}{n-1}((Z_1-\overline{Z}_n)\otimes(Z_1-\overline{Z}_n) + ... + (Z_n-\overline{Z}_n)\otimes(Z_1-\overline{Z}_n))$$

is such that $(n-1)S_n^2$ has Wishart distribution $\gamma_{(n-1)/2,2\Sigma}$.

III Wishart distributions on symmetric cones

The more general Wishart distributions that we are going now to consider are extensions of the $\gamma_{n/2,2\Sigma}$ defined at the end of the previous section : first we replace the discrete parameter $n/2$ by a continuous one, exactly like for $r=1$, gamma distributions (1.2) extend the χ^2 distributions. Let us keep E to be the space of symmetric endomorphisms of the Euclidean space V for a while. For p in

$$\Lambda = \{1/2, 1, 3/2, ..., (d-1)/2\}\cup((d-1)/2,+\infty), \tag{3.1}$$

then for σ in E_+, the distribution $\gamma_{p,\sigma}$ is defined as before if p is a half integer, and is defined by

$$\gamma_{p,\sigma}(du) = e^{-\langle \sigma^{-1}, u\rangle}(\det\ u)^{p-(1+r)/2}(\det \sigma)^{-p}(\Gamma_r(p))^{-1}1_{E_+}du,$$

(where $\Gamma_r(p)$ is a constant), if $p > (r-1)/2$, as an extension of (1.7). Here again, one has for $-\sigma$ in E_+:

$$\int_E e^{\langle \theta, u\rangle}\gamma_{p,\sigma}(du) = (\det(\mathrm{id}_V - \sigma\theta))^{-p}. \tag{3.2}$$

The proof of (3.2) is the same as (1.8). Now, it has been considered as a challenging problem among statisticians until 1985 to prove that if $p > 0$ and if p does not belong to Λ as defined by (3.1), then no probability measure $\gamma_{p,\sigma}$ fulfilling (3.2) can exist (although actually the problem had been solved in 1975 by Gindikin in a different context). Thus

$$\{\gamma_{p,\sigma}; p \in \Lambda,\ \ \sigma \in E_+\}$$

provides the continuous extension of the classical Wishart distributions that we were looking for.

An other generalization of the classical Wishart distributions is offered when one looks at the Gaussian distributions on Hermitian spaces instead of Euclidean ones. Complex numbers can again be replaced by quaternions. We are going to consider now a common generalization of all these kinds of Wishart distributions, which is the topic of the present lectures. Let E be, from now on, a simple Euclidean Jordan algebra. Our reference is the book by Faraut and Koranyi (1994) " Analysis on symmetric cones ". The unit element is e. The rank of E is r, the Peirce constant is d and the dimension of E is $n = r + \frac{dr}{2}(r-1)$. The Jordan product is written $x.y$, and the scalar product is Trace $x.y =< x, y >$. We denote by $\overline{E}_+ = \{a.a; a \in E\}$ the cone of squares and by E_+ its interior. G denotes the group of automorphisms of the symmetric cone E_+, and K is the intersection of G with the orthogonal group of E. One also introduces the following extension of (3.1), which was associated to the unique E with structure constants r and $d = 1$:

$$\Lambda = \{d/2, d, 3d/2, ..., (r-1)d/2\} \cup ((r-1)d/2, +\infty). \tag{3.3}$$

and we state an important theorem:

Theorem 3.1: Let $p > 0$. Then there exists a positive measure μ_p on $\overline{E}_+$ such that for any θ in E_+ one has

$$\int_E e^{-<\theta,u>}\mu_p(du) = (\det \theta)^{-p} \tag{3.4}$$

if and only if p is in Λ defined by (3.3). Furthermore, μ_p is absolutely continuous if and only if $p > dr(r-1)/2$.

Comments: The part $(\Rightarrow)$ is the Gindikin theorem. Our proof is taken from Casalis and Letac (1994), although the idea is due to Shanbhag (1987). Note also that the μ_p appearing in the theorem are quasi invariant by G. More precisely, they are such that for all g in G, then the image of μ_p by g is $|Det\ g|^{-rp/n}\mu_p$. Since G is closed by $g \mapsto g^*$, this is a consequence of (3.4) and of the fact that $|\det\ g(\theta)| = |Det\ g|^{-r/n}|\det\theta|$. Note that in the last formula, the determinant is applied in two different contexts: the Jordan algebra E for θ and det, the space $L(E)$ of the endomorphisms of E for g and Det.

Proof of Th.3.1: ($\Leftarrow$) We show first the existence of $\mu_{d/2}$. Let c be a primitive idempotent of E. Consider the decomposition $E = E_1 \oplus E_{1/2} \oplus E_0$ in eigenspaces of the symmetric endomorphism $x \mapsto c.x$. Then there exists a constant C such that for $\theta = \theta_1 c + \theta_{12} + \theta_0$ in E_+, with $(\theta_1, \theta_{12}, \theta_0)$ in $\mathbb{R} \times E_{1/2} \times E_0$, one has:

$$\int_0^{\infty} (\int_{E_{1/2}} e^{-\langle \theta_{1/2}, b\rangle - a^{-1}\langle \theta_0, b.b\rangle} db) e^{-\theta_1 a} a^{d-1-dr/2}\, da = C(\det \theta)^{-d/2}. \tag{3.5}$$

(3.5) shows the existence of $\mu_{d/2}$, as the image of the measure on $(0, +\infty) \times E_{1/2}$ equal to

$$C^{-1} a^{d-1-dr/2}\, da\, db$$

by the map $(a, b) \mapsto ac + b + a^{-1}(e - c).(b.b)$. Thus the existence of μ_p is proved when $2p/d$ is an integer, by taking the powers of convolution of $\mu_{d/2}$. The proof of (3.5) is a standard computation, based on the fact that the Gaussian distribution (1.5) has mass one. To get the final determinant, one has to use a formula that I learnt from Massam and Neher (1997) : for $\theta \in E_+$:

$$\det \theta = (\det \theta_0)(\det(\theta_1 c - \mathbb{P}(\theta_{1/2})(\theta_0^{-1})).$$

where $\mathbb{P}$ is the quadratic map of the Jordan algebra E.

If $p > d(r-1)/2$, we imitate (1.8) and prove that there exists a constant C such that

$$\int_{E_+} e^{-\langle \theta, u\rangle} (\det\ u)^{p-n/r} du = C(\det \theta)^{-p}.$$

This is proved by the Bartlett method : see Faraut and Koranyi, Th.VI.4.9.

($\Rightarrow$) Suppose that there exists some $p \geq 0$ and such that there exists a positive measure μ_p on $\overline{E}_+$ with (for all $-\theta \in E_+$) :

$$\int_{\overline{E}_+} \exp\langle \theta, u\rangle \mu_p(du) = (Q(-\theta))^{-p} \tag{3.6}$$

where $Q = \det$. Since Q is a polynomial of degree r, we apply the differential operator $Q(\frac{d}{d\theta})$ to both sides of (3.6). The result for the second member is a polynomial in p of degree r (at most) multiplied by $(Q(-\theta))^{-p-r}$. Considering the result for the first member, we get that

$$R(p) = (Q(-\theta))^{p+r} \int_{\overline{E}_+} Q(u) \exp\langle \theta, u\rangle \mu_p(du) \tag{3.7}$$

is a polynomial in p with at most r roots. Actually, these r roots are $\{0, d/2, d, 3d/2, ..., (r-1)d/2\}$. The argument is the following : μ_p does exist for these values of p, as we have seen in the first part of the proof ; it is concentrated on $\overline{E}_+ \setminus E_+$ ($\mu_{d/2}$ was concentrated on the multiples of the primitive idempotents, thus on elements of rank 1 ; thus $\mu_{kd/2}$ is concentrated on elements of rank $\leq k$, for $k < r$, as a convolution). Now for such a $p = kd/2$, $Q(u) = 0$, for $\mu_{kd/2}$ almost all u, and (3.7) is 0.

Suppose now that μ_p exists for some $p = p_0$ which is not in Λ. Thus $R(p_0) = 0$ is impossible. Since Q is ≥ 0 on $\overline{E}_+$, $R(p_0) < 0$ is impossible. Since the roots of R are simple and since R is positive on $((r-1)d/2, +\infty)$ (because μ_p is concentrated on E_+) then R is negative on $((r-2)d/2,(r-1)/2)$. Therefore since $R(p_0)$ is positive, then $R(p_0+d/2)<0$. But $\mu_p * \mu_{d/2} = \mu_{p+d/2}$: a contradiction. This ends the proof of Th.3.1.

Finally we have the definition of a Wishart distribution.

Definition 3.1: Let E be a simple Jordan algebra as above. Let p be in Λ defined by (3.4), μ_p defined by Th.3.1, and σ in E_+. Then the Wishart distribution with shape parameter p and scale parameter σ is

$$\gamma_{p,\sigma}(du) = \exp(-\langle \sigma^{-1}, u\rangle)(\det \sigma)^{-p}\mu_p(du) \tag{3.8}$$

From Th.3.1, and the formula $\det(\mathbb{P}(y)x) = (\det\ y)^2 \det\ x$, it is clear that its Laplace transform is

$$\int_E e^{\langle \theta, u\rangle}\gamma_{p,\sigma}(du) = (\det(e-\mathbb{P}(\sigma^{1/2})(\theta))^{-p}. \tag{3.9}$$

Let us now comment on a special subclass:

Wishart distributions with Gaussian origin

E is still a simple Euclidean Jordan algebra with parameters r and d. Let F be a Euclidean space with dimension N and scalar product $\langle f, f'\rangle_F$. $L_s(F)$ is the space of symmetric endomorphisms of F. Consider a self adjoint representation ϕ of E on F, i.e. a linear map

$$\phi : E \to L_s(F)$$

such that $\phi(e) = \mathrm{id}_F$ and, for all x and y in E:

$$\phi(x.y) = \frac{1}{2}(\phi(x)\phi(y) + \phi(y)\phi(x)).$$

Then there exists a symmetric bilinear map

$$Q : F \times F \to E$$

such that

$$\langle \phi(x)f, f'\rangle_F = \langle x, Q(f, f')\rangle$$

We write $Q(f) = Q(f, f)$. Recall (Faraut and Koranyi, Prop.IV.5.2) that these hypothesis imply that ϕ is injective, that N is a multiple of r and that for all x in E one has $Det(\phi(x)) = (\det\ x)^{N/r}$.

Example: V is Euclidean with dimension r, $E = L_s(V)$ and $F = V^k$. For $f = (v_1, ...v_k)$, define

$$\phi(x)f = (x(v_1), ..., x(v_k))$$

Then $Q(f) = v_1 \otimes v_1 + ... + v_k \otimes v_k$.

Proposition 3.2 : Let Σ in $L_s(F)$ be in the image of ϕ and be positive definite. Then the image of the Gaussian distribution $N_{0,\Sigma}$ (see (1.5)) on F by $f \mapsto Q(f)$ in the Jordan algebra E is a Wishart distribution $\gamma_{p,\sigma}$ with $p = N/2r$ and $\sigma = 2\phi^{-1}(\Sigma)$.

Proof : It is standard to check that $\Sigma^{1/2}$ is also in the image of ϕ : use a Jordan basis $(c_1, ..., c_r)$ to represent σ and use the fact that $\phi(c_1), ..., \phi(c_r)$ are orthogonal projections on orthogonal subspaces $F_1, ..., F_r$ with common dimension N/r. We denote $g = \mathbb{P}(\sigma^{1/2})$.

We compute the Laplace transform of this image for $-\theta$ in E_+:

$$\int_F e^{\langle\theta,Q(f)\rangle} N_{0,\Sigma}(df) = \int_F e^{\langle\phi(\theta)f,f\rangle} N_{0,\Sigma}(df) = (\det(\mathrm{id}_F - 2\Sigma\phi(\theta)))^{-1/2},$$

from (1.6). But

$$Det(\mathrm{id}_F - 2\Sigma\phi(\theta)) = Det(\mathrm{id}_F - 2(\Sigma)^{1/2}\phi(\theta)(\Sigma)^{1/2}) = Det(\phi(e - g(\theta))).$$

But $g(e) = \sigma$. Thus

$$Det(\mathrm{id}_F - 2\Sigma\phi(\theta)) = Det\phi(g(\sigma^{-1} - \theta)) = (\det(e - \mathbb{P}(\sigma^{1/2})(\theta))^{N/r},$$

by using $Det(\phi(x)) = (\det\ x)^{N/r}$. With (3.9) the proof is complete.

When we apply this proposition to the previous example, we get the classical Wishart distribution as described in section 1 (but note that, for the example, Σ is in the image of ϕ if and only if it is made with k copies of some Σ_1 of $L_s(V)$, and $\sigma = 2\Sigma_1$ and $p = k/2$). Wishart distributions of Gaussian origin when the Jordan algebra is the space of Hermitian matrices on complex or quaternionic numbers have been considered by statisticians : Goodman (1963), Andersson (1975). Note that there is no representation, and hence no Wishart distribution of Gaussian origin on the exceptional Jordan algebra defined by $r = 3,\ d = 8$. The case where $r = 2$, corresponding to a symmetric cone of revolution in a Euclidean space, is quite interesting. The smallest N such that there exists a representation on $\mathbb{R}^N$ for this algebra E depends on the arithmetic properties of the dimension n of E, specially on the residue mod.8 of n. This is achieved by consideration of the Clifford algebra associated to a Euclidean space. The present author naively found (Letac (1994)) that the smallest N is the smallest of Λ (see (3.3)) if and only if n= 3, 4, 6 and 10, by using Hurwitz Theorem. However, he discovered that Jensen (1988) gives the complete story about N.

Wishart distributions with Gaussian origin occur also in statistics through the following problem. Suppose that you consider the Gaussian distribution $N_{0,\Sigma}$ on the Euclidean space F, where Σ is unknown, but supposed to be such that there exist two linear subspaces E and E' of $L_s(F)$, of same dimension, such that Σ is in E and $(\Sigma)^{-1}$ is in E'. Without loss of generality, we may assume that E contains id_F : change a bit the Euclidean structure of F to have this. Under these circumstances, then E is a subJordan algebra of $L_s(F)$- and injection is a representation of E, of course. It is not necessarily simple, but it is the sum of identical simple ones. And the classical Wishart distributions that you build from $N_{0,\Sigma}$ will be Wishart on E, obviously of Gaussian origin. The proof of the fact that E is a Jordan algebra is a consequence of the following (see Jensen (1988)).

Proposition 3.3 : Let F be a Euclidean space, S be the space of the symmetric endomorphisms of F, S_+ be the cone of positive definite ones, E be a linear subspace of S containing id_F. Then there exists a linear subspace E' of S such that the map $a \mapsto a^{-1}$ is a bijection between $E \cap S_+$ and $E' \cap S_+$ if and only if $a^2 \in E$ for all $a \in E$. In this case, $E = E'$.

Proof : $(\Rightarrow)$ If a is in E, then for small real t one has $\mathrm{id}_F - ta \in E \cap S_+$, $(\mathrm{id}_F - ta)^{-1} \in E' \cap S_+$, and

$$a = \lim_{t\to 0}((\mathrm{id}_F - ta)^{-1} - \mathrm{id}_F)/t \in E'.$$

Thus $E \subset E'$, and $E = E'$ by symmetry. Similarly

$$a^2 = \lim_{t\to 0}((\mathrm{id}_F - ta)^{-1} - \mathrm{id}_F - ta)/t^2$$

Since $(\mathrm{id}_F - ta)^{-1} \in E' \cap S_+ = E \cap S_+$, one has $a^2 \in E$.

($\Leftarrow$) For a in E, $2a^n = (a^{n-1} + a)^2 - a^{2n-2} - a^2$ implies that polynomials in a are in E. If $a \in E \cap S_+$, Cayley Hamilton theorem shows that $a^{-1} \in E' \cap S_+$.

IV Wishart distributions as natural exponential families

Let us first introduce some statistical objects. Let E be a n dimensional real linear space. E^* is its dual, $(\theta, x) \mapsto \langle \theta, x\rangle$ is the canonical bilinear map on $E^* \times E$. $\mathcal{M}$ is the set of positive measures μ on E which are not concentrated on some affine hyperplane and such that the set where the Laplace transform

$$L_\mu(\theta) = \int e^{\langle \theta, x\rangle} \mu(dx)$$

is finite has a non empty interior $\Theta(\mu)$. This is an easy exercise on Holder inequality to see that $\Theta(\mu)$ is convex and that $k_\mu(\theta) = \log L_\mu(\theta)$ is a strictly convex function on it. Then the set $F = F(\mu)$ of probability measures

$$P(\theta, \mu)(dx) = \exp(\langle \theta, x\rangle - k_\mu(\theta))\mu(dx), \qquad \theta \in \Theta(\mu) \tag{4.1}$$

is called the natural exponential family generated by μ (NEF). It is easily seen that

$$k'_\mu(\theta) = \int_E xP(\theta, \mu)(dx). \tag{4.2}$$

For this reason, the subset $M_F = k'_\mu(\Theta(\mu))$ of E is called the domain of the means of the NEF. Since k_μ is strictly convex, k'_μ is a diffeomorphism from $\Theta(\mu)$ onto M_F. We denote by ψ_μ its inverse from M_F onto $\Theta(\mu)$. Thus the map

$$m \mapsto P(\psi_\mu(m), \mu) = P(m, F) \tag{4.3}$$

is a new parametrization of the NEF F by its domain of the means.

An important object for a NEF is now its variance function. Given any probability P on the space E such that $E^* \subset L^2(P)$, the covariance operator is the linear map Σ from E^* to E defined by

$$\langle \alpha, \Sigma(\beta)\rangle = \int_E \langle \alpha, x - m\rangle\langle \beta, x - m\rangle P(dx),$$

where $m = \int_E xP(dx)$ is the expectation of P. Hence Σ belongs to the space $L_s(E^*, E)$ of symmetric linear maps. When we specialise P to be $P(\theta, \mu)$ in the NEF $F(\mu)$, then $\Sigma = k''_\mu(\theta)$. If we take $P = P(m, F)$ as defined by (4.3), then we denote $\Sigma = V_F(m)$. This function from M_F to $L_s(E^*, E)$ $m \mapsto V_F(m)$ is called the variance function of the NEF. It satisfies the identities

$$V_F(m) = (\psi'_\mu(m))^{-1}, \tag{4.4}$$

$$k''_\mu(\theta) = V_F(k'_\mu(\theta)). \tag{4.5}$$

Equality (4.5) shows that the knowledge of the variance function on some open part of M_F gives the knowledge of the NEF, since k'_μ is the solution of the differential equation $y'(\theta) = V_F(y(\theta))$.

A variance function has strong properties of regularity:

Proposition 4.1 : For α and β in E^*, then for m in M_F one has

$$V'_F(m)(V_F(m)(\alpha))(\beta) = V'_F(m)(V_F(m)(\beta))(\alpha)$$

Proof : Differentiate (4.4) : since $\psi''_\mu(m)$ is a Hessian, it must be symmetric.

To end up the generalities on NEF, we shall say that the NEF F is reducible if $E = E_1 \oplus E_2$ with dim $E_i > 0$, and there exists NEF F_i on E_i $(i = 1, 2)$ such that any element of F is the product of two elements of F_1 and F_2. If not, F is said to be irreducible.

Let us now apply these definitions to the case where E is a simple Euclidean Jordan algebra and to the $\mu = \mu_p$ of Th.3.1. From definition 3.1, we see that $F(\mu_p)$ is nothing but the set of Wishart distributions with fixed shape parameter p. Here E^* is identified to E through the Euclidean structure, the Laplace transform is given by Th.3.1, $\Theta(\mu)$ is $-E_+$, $k_\mu(\theta)$ is $-p \log \det(-\theta)$. Thus

$$k'_\mu(\theta) = p\theta^{-1}, \quad M_F = E_+, \quad \psi_\mu(m) = -pm^{-1}. \tag{4.6}$$

This computation of k'_μ appears in Faraut and Koranyi, Prop. III.5.3. Now, combining (4.4) and (4.6), we get the variance function of the Wishart NEF with shape parameter $p \in \Lambda$, just by differentiating the map $m \mapsto m^{-1}$. From Faraut and Koranyi, Prop. II.4.3, this differential is $-\mathbb{P}(m)^{-1}$, where $\mathbb{P}$ denotes the quadratic representation; finally, using (4.4), we get that the variance function is $p^{-1}\mathbb{P}(m)$. A remarkable feature of this result is that this variance function is a homogeneous polynomial of degree 3. Solving a conjecture of the author of these lectures, M.Casalis (1991) has proved the converse of this statement by the following beautiful result:

Theorem 4.2 : Let F be a natural exponential family on a real linear space E. Assume that the variance function V_F is the restriction to M_F of a homogeneous polynomial with degree 3. Assume also that F is irreducible. Then there exists a structure of simple Euclidean Jordan algebra on E such that F is a Wishart family for some p in the Λ of the algebra.

Proof : We fix $e \in M_F$ arbitrarily. For convenience, the canonical bilinear form on $E^* \times E$ is rather denoted by (θ, x). We give to E the Euclidean structure $\langle x, y\rangle = (V_F(e)^{-1}(x), y)$. Thus E and E^* are identified, and the variance function of F is now the symmetric endomorphism of E equal to $V(m) = V_F(m)V_F(e)^{-1}$. Prop.4.1 implies that $(V'(m)x)V(m)^{-1}y$ is symmetric in x and y. Write

$$V'(m)h = 2A(m, h)$$

Since V is quadratic, A is bilinear and symmetric. The previous symmetry in $x = V(m)u$ and $y = V(m)(v)$ becomes

$$A(m, A(m,m)u)v = A(m, A(m,m)v)u \tag{4.7}$$

We now equip E with a structure of algebra by the product $u.v = A(e,u)v$. It happens that we get a Jordan algebra, which admits e as unit element. This is proved by polarization of the identity (4.7) : see Casalis (1991) for the explicit calculation. It is Euclidean, since $A(e,x)$ is a symmetric operator. We have also $V(m) = \mathbb{P}(m)$. This proves in particular that the algebra is simple, since F is irreducible, thus its variance function cannot be diagonalised in several blocks. Finally, the exact value of p in Λ occurs in a subtle way: the scalar product did not fill the condition $< x, y >$=trace $x.y$. But define $p = \langle e, e\rangle / r$, where the rank r of E is trace e. Then $V_F(m) = \mathbb{P}(m)/p$ for all m in M_F. By definition, e belongs to both M_F and E_+. Thus, we have two variance functions which coincide on some non void open set, we have $M_F = E_+$, and this ends the proof.

V Properties of the Wishart distributions

They are obviously guessed from the properties of the classical ones. Then either the extension is plain, or the extension is more creative, or...the classical property was not well understood and the extension to Jordan algebras enlightens it and provides a simpler proof. The second situation may be challenging, but the third is pretty much what Pascual Jordan had in mind while creating his algebras. Let us give an example of the second situation (Letac and Massam (1995)) with an extension of the so called Craig's theorem.

Theorem 5.1 : Let X be Wishart distributed on the Jordan algebra E, with scale parameter e. Let a and b in E. Then $\langle a, X\rangle$ and $\langle b, X\rangle$ are independent if and only if $a.b = 0$ and $[L(a), L(b)] = 0$.

(Here, $L(a)$ is the endomorphism of E defined by $L(a)(x) = a.x$.)

Proof : If p is the scale parameter of the distribution of X, then from (3.9),

$$\mathbb{E}(e^{t\langle a,X\rangle + s\langle b,X\rangle}) = (\det(e - (ta + sb))^{-p},$$

for real numbers t and s small enough. Thus the problem is to show that

$$\det(e - (ta + sb)) = \det(e - ta)\det(e - sb) \tag{5.1}$$

for all $(t, s) \in \mathbb{R}^2$, if and only if $a.b = 0$ and $[L(a), L(b)] = 0$.

($\Leftarrow$) We write $a = t_1 c_1 + ... + t_k c_k$, where $(c_1, ..., c_k)$ is a sequence of orthogonal primitive idempotents and $t_1, ..., t_k$ are non zero real numbers. Write $c = c_1 + ... + c_k$, and $b = b_1 + b_{1/2} + b_0$ for the Peirce decomposition in the three eigenspaces of $L(c)$. Applying $[L(a), L(b)] = 0$ to c gives

$$0 = a.(b_1 + b_{1/2}/2) - (b_1 + b_{1/2} + b_0).a,$$

thus $a.b_{1/2} = 0$, and we get $b_{1/2} = 0$ (with some standard calculation using the fact that the t_j are not zero). Furthermore, by hypothesis $a.b = 0$, this leads to $a.b_1 = 0$, and finally to

$b_1 = 0$. We now split b itself into a sum of primitive idempotents : $b = t_{k+1}c_{k+1} + ... + t_r c_r$, with $e - c = c_{k+1} + ... + c_r$. Now we get (5.1) easily.

($\Rightarrow$) Taking the Taylor expansion of the log of both sides of (5.1), we get that for all integers n :

$$Trace((ta + sb)^n - t^n a^n - s^n b^n) = 0 \tag{5.2}$$

We take $t = 1$ and $n = 4$ in (5.2) and compute the coefficient of s^2. It is

$$0 = Trace(a.(a.b^2) + 2a.(b.(a.b)) + 2b.(a.(a.b)) + b.(b.a^2)) =$$

$$2Trace\mathbb{P}(a)(b^2) + 4Trace(a.b)^2$$

This leads to $a.b = 0$ and to $\mathbb{P}(a)(b^2) = 0$, and again with some standard manipulation, to $[L(a), L(b)] = 0$.

We now present an example of a situation which has been illuminated by the generalization to Jordan algebras. A classical result in statistics, due to Lukacs (1955), says that if U and V are independent positive random variables (non Dirac), then $Z = U/(U+V)$ is independent of $U + V$ if and only if there exist σ, p and $q > 0$ such that the distributions of U and V are respectively $\gamma_{p,\sigma}$ and $\gamma_{q,\sigma}$, as defined by (1.2). The extension of this to classical Wishart distributions (and a little bit more, by accepting the parameters p in Λ defined by (3.1)), was a challenge taken by Olkin and Rubin (1962). Their paper is very difficult to read, and even obscure ; but ideas which can be transposed in Jordan algebras pervade it. Casalis and Letac (1996) and Letac and Massam (1998) lead to the Th.5.2 below. Before stating it, recall that for a simple Euclidean Jordan algebra E, K denotes the intersection of G with the orthogonal group of E. In the Lukacs theorem, we are going to replace the positive line by $\overline{E}_+$, and the gamma distributions by the Wishart ones. But the definition of Z is more delicate ; we are not going to consider the Jordan product $U.(U + V)^{-1}$, which anyway is not necessarily in $\overline{E}_+$. After Olkin and Rubin, we take an arbitrary measurable map $x \mapsto g(x)$ from E_+ to G, and we call it a division algorithm if $g(x)(x) = e$ for all x. For instance, if $E = L_s(V)$, with V Euclidean,

$$g(x)(y) = x^{-1/2} y x^{-1/2}$$

defines one of the possible division algorithms on $L_s(V)$.

Theorem 5.2 : Let E be a simple Euclidean Jordan algebra, and U and V be two independent random variables valued in $\overline{E}_+$, and such that $U + V$ is not concentrated on a fixed half line. Consider the four statements :

(A) $U + V$ is almost surely in E_+, there exists a division algorithm g such that $Z = g(U + V)(U)$ is independent of $U + V$ and the distribution of Z is K invariant.

(B) There exist a in $\mathbb{R}$ and a $(2, 2)$ real matrix b such that

$$\mathbb{E}(U|U + V) = a(U + V) \tag{5.3}$$

and

$$\begin{aligned}\mathbb{E}(U \otimes U|U + V) &= b_{11}(U + V) \otimes (U + V) + b_{12}\mathbb{P}(U + V),\\ \mathbb{E}(\mathbb{P}(U)|U + V) &= b_{21}(U + V) \otimes (U + V) + b_{22}\mathbb{P}(U + V).\end{aligned} \tag{5.4}$$

(C) There exists σ in E_+, p and q in Λ as (3.3) such that U and V have Wishart distributions $\gamma_{p,\sigma}$ and $\gamma_{q,\sigma}$.

(C') Same as (C), with $p+q > (r-1)/3$.

Then (C') $\Leftrightarrow$ (A) $\Rightarrow$ (B) $\Leftrightarrow$ (C). Furthermore, under (C') the distribution of Z depends neither on σ, nor on the particular division algorithm. Under (C), the a and b of (B) are $a = p/(p+q)$ and $b = A(p)(A(p+q))^{-1}$, where

$$A(p) = \begin{bmatrix} p & p^2 \\ pd/2 & p(1-d/2)+p^2 \end{bmatrix}.$$

Proof : (sketch) (C') $\Rightarrow$ (A). One first uses the quasi invariance by G of the μ_p of Th.3.1 to show that for any g of G, the image by g of the conditional distribution K_s of U knowing that $U+V=s$ is $K_{g(s)}$. After this remark, computations are easy.

The remainder of the proof of Th.5.2 makes an essential use of the following :

Lemma 5.3 : 1) If x is in E, then $k(x) = x$ for all $k \in K$ if and only if x is a multiple of e.

2) 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

$$f = \lambda id_E + \mu e \otimes e.$$

We leave the proof to Jordan's aficionados.

(A) $\Rightarrow$ (B) : Since the distribution of Z is $K-$invariant, then $\mathbb{E}(Z)$ is equal to ae for some real a, from the lemma. Hence

$$ae = \mathbb{E}(g(U+V)(U)|U+V) = g(U+V)(\mathbb{E}(U|U+V))$$

Since $g(x)(x) = e$ for all x, we have $(g(x))^{-1}(e) = x$. Hence $a(U+V) = \mathbb{E}(U|U+V)$. The proof of the existence of b is similar, using the second part of the lemma.

(B) $\Rightarrow$ (C) : This is the heart of the proof. We use the notation k_U instead of k_μ used in section 3, with obvious meaning. It is easy to see that if we write $\chi = k_{U+V}$, then the first part of the hypothesis (B) implies that $k_U = a\chi$ and $k_V = (1-a)\chi$. Thus a consequence of the second part of the hypothesis (B) is that

$$a\chi'' + a^2\chi' \otimes \chi' = b_{11}(\chi'' + \chi' \otimes \chi') + b_{12}(\mathbb{P}(\frac{d}{d\theta})\chi + \mathbb{P}(\chi')) \tag{5.5}$$

and

$$a\mathbb{P}(\frac{d}{d\theta})\chi + a^2\mathbb{P}(\chi') = b_{21}(\chi'' + \chi' \otimes \chi') + b_{22}(\mathbb{P}(\frac{d}{d\theta})\chi + \mathbb{P}(\chi')) \tag{5.6}$$

Elimination of $\mathbb{P}(\frac{d}{d\theta})\chi$ between (5.5) and (5.6), with a tedious discussion showing that coefficients are not zero and that the distribution of $U+V$ is not concentrated on an affine hyperplane, shows that

$$\chi''(\theta) = \mathbb{P}(\chi'(\theta))/\lambda + \beta\chi'(\theta) \otimes \chi'(\theta). \tag{5.7}$$

for some λ and β in $\mathbb{R}$. An application of Prop.4.1 implies that $\beta = 0$. Thus the variance function of the natural exponential family generated by the law of $U+V$ is, from (5.7), the one of a Wishart family.

(C) $\Rightarrow$ (B) : The proof relies on the two identities :

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

$$\mathbb{E}(\mathbb{P}(U)) = (p - \frac{pd}{2} + p^2)\sigma \otimes \sigma + \frac{pd}{2}\mathbb{P}(\sigma). \tag{5.8}$$

The first one is a particular case of Huyghens theorem (the expectation of a square is the square of the expectation plus the variance), but the second one is more specific to Jordan algebras. Its proof relies on the lemma, but is a bit technical here. And from (5.8), standard probability theory leads to (B). Finally, note that (A)+(C) imply trivially (C'), and the proof of the Olkin and Rubin's result, which was (C') $\Leftrightarrow$ (A) for symmetric real matrices is now complete and extended to Jordan algebras.

Before stating the last property of Wishart distributions that we want to mention, let us observe that in general, if F is a NEF on some linear space E, and if g is an automorphism of E, then gF, defined as the set of images by g of the elements of F, is still a NEF, and that its variance function satisfies $gM_F = M_{gF}$ and

$$V_{gF}(m) = gV_F(g^{-1}m)g^* \tag{5.9}$$

We now characterize the Wishart NEF by an invariance property. Interestingly enough, the proof uses some elementary arguments of algebraic topology. See Letac (1989) and (1994) for $d = 1$ or $r = 2$, and Casalis (1990) for the general case and two distinct proofs.

Theorem 5.5 : Let F be a NEF on the simple Euclidean Jordan algebra E such that $gF = F$ for any g of the automorphism group G of the symmetric cone E_+. Then there exists p in Λ (as defined by (3.3)) such that if F_p is the Wishart NEF with shape parameter p, then either $F = F_p$ or $F = -F_p$.

Proof : (sketch) ($\Leftarrow$) The variance function of F_p is $p^{-1}\mathbb{P}$, as we have seen in section 4. We use the two classical identities (see Faraut and Koranyi, Prop.II.4.3,(iii) and III.5.3) :

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

and, for (g, y) in $G \times E$:

$$\det(gy) = |Det\ g|^{r/n} \det y,$$

and the polar decomposition of g in G as $g = \mathbb{P}(x)k$, with x in E_+ and k in K. Thus

$$\mathbb{P}(gm) = \mathbb{P}(\mathbb{P}(x)km) = \mathbb{P}(x)\mathbb{P}(km)\mathbb{P}(x) = \mathbb{P}(x)k\mathbb{P}(m)k^*\mathbb{P}(x) = g\mathbb{P}(m)g^*,$$

and from (5.9), since the variance function characterizes the NEF, this is enough to claim that $F_p = gF_p$. The case of $-F_p$ is easily obtained from this.

($\Rightarrow$) We fix a Jordan basis $(c_1, ..., c_r)$ and denote for $j = 0, 1, ..., r$, by f_j the element of E equal to $c_1 + \cdots + c_j - c_{j+1} - \cdots - c_r$. We denote $E_j = Gf_j$ and prove that

(i) there exists j such that $M_F = E_j$;

(ii) actually $j = 0$ or $j = r$;

(iii) there exists p in Λ such that $V_F = p^{-1}\mathbb{P}$.

To see (i), one observes that M_F is open and connected -as the image of an open convex set by a diffeomorphism, and that it is invariant by G, from the hypothesis. Thus it exists some j such that $M_F \supset E_j$. If it not equal to E_j, then M_F contains some x with $\det x = 0$, and one shows with this that there exists some m in M_F such that $V_F(m)$ is not positive definite : a contradiction.

(ii) uses the fact that the only open orbits E_j of G which are homeomorphic to E are $E_+ = E_0$ and $-E_+ = E_r$.

(iii) : Without loss, we take $j = 0$. Therefore we have

$$V_F(m) = \mathbb{P}(\sqrt{m})V_F(e)\mathbb{P}(\sqrt{m}) \tag{5.10}$$

and the delicate point is now to show that $V_F(e)$ is proportional to id_E. This is achieved by differentiating (5.10) and by using the symmetry condition of the variance functions appearing in Prop.4.1.

VI The one dimensional non central χ^2 distributions and their interpolation

If $Z \sim N(0,1)$ then the Laplace transform of $(Z+m)^2/2$ is

$$\mathbb{E}(e^{-\frac{1}{2}s(Z+m)^2}) = \frac{e^{-\frac{1}{2}sm^2}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{\frac{1}{2}z^2(s+1)-smz}dz = \frac{1}{\sqrt{1+s}}e^{-\frac{1}{2}\frac{sm^2}{1+s}}.$$

One observes that

$$e^{-\frac{1}{2}\frac{sm^2}{1+s}} = e^{-\frac{1}{2}m^2+\frac{1}{2}m^2\frac{1}{1+s}}$$

is the Laplace transform of $Y_1 + \ldots + Y_N$ when $(Y_i)_{i\geq 1}$ are iid exponential variables with mean 1 which are independent from the Poisson random variable N with mean $\frac{1}{2}m^2$ and that $\frac{1}{\sqrt{1+s}}$ is the Laplace transform of s gamma distribution with shape parameter $\frac{1}{2}$ and scale parameter 1. The law of $Y = Y_1 + \ldots + Y_N$ can be written as

$$e^{-\frac{1}{2}m^2}\delta_0(dy) + e^{-\frac{1}{2}m^2-y}\sum_{n=1}^{\infty}\frac{y^{n-1}}{(n-1)!)} \times (\frac{1}{2}m^2)^n\frac{1}{n!}$$

If we consider now n iid normal $N(0,1)$ random variables $Z_1, \ldots, Z_n$ and n numbers $m_1, \ldots, m_n$ then the Laplace transform of

$$W = (Z_1 + m_1)^2 + \cdots + (Z_n + m_n)^2$$

is

$$\frac{1}{(1+s)^{n/2}}e^{-\frac{1}{2}\frac{sm^2}{1+s}}$$

with the notation

$$m^2 = m_1^2 + \cdots + m_n^2$$

It can be seen as the sum of a gamma variable $G_{n/2}$ with shape parameter $n/2$ with the above Y with Y and $G_{n/2}$ independent.

At this point clearly one can replace the parameter $n/2$ by the continuous parameter $p > 0$ and we get the general non central chi square law. It should be more reasonable to call it the non central gamma distribution with Laplace transform

$$\frac{1}{(1+s)^p}e^{-\frac{1}{2}\frac{sm^2}{1+s}},$$

namely the distribution of $G_p + Y$ with obvious notations.

VII The classical non central Wishart distribution.

The non-central Wishart distribution is traditionally defined as the distribution of the random symmetric real matrix $X = Y_1 Y_1^* + \cdots + Y_n Y_n^*$ where $Y_i \in \mathbb{R}^d,\ i = 1, \ldots, n$ are independent Gaussian column vectors with the same non-singular covariance matrix Σ and respective means $m_i,\ i = 1, \ldots, n$ not necessarily equal (here $*$ means transposition). For s in the open cone $\mathcal{P}_d$ of positive definite symmetric matrices of order d and

$$w = m_1 m_1^* + \cdots + m_n m_n^*$$

in the closed cone of semi positive definite matrices $\overline{\mathcal{P}_d}$ one can readily derive the Laplace transform

$$\mathbb{E}(e^{-\operatorname{tr}(sX)}) = \frac{1}{\det(I_d + 2\Sigma s)^{n/2}} e^{-\operatorname{tr}(2s(I_d + 2\Sigma s)^{-1} w)}. \tag{7.11}$$

It is important to note that in this formula the rank k of w is less or equal to n and d.

Exactly like what we have done before by extending the familar chi square distribution with n degrees of freedom to the gamma distribution with a continuous shape parameter, one is tempted to extend the values that the power of $\det(I_d + 2\Sigma s)$ can take in (7.11). The question is then: given $\Sigma \in \mathcal{P}_d$ and $w \in \overline{\mathcal{P}_d}$, for which values of $p > 0$ does there exist a probability distribution on $\overline{\mathcal{P}_d}$ for X such that for all $s \in \mathcal{P}_d$ we have

$$\mathbb{E}(e^{-\operatorname{tr}(sX)}) = \frac{1}{\det(I_d + 2\Sigma s)^{p}} e^{-\operatorname{tr}(2s(I_d + 2\Sigma s)^{-1} w)}? \tag{7.12}$$

Call this hypothetic distribution for X satisfying (7.12) a non-central Wishart distribution with parameters $(2p, w, \Sigma)$, or $NCW(2p, w, \Sigma)$ for short.

VIII The non central Wishart exists only if p is in the Gindikin set

Proposition 8.1: If there exists a distribution on the semipositive definite real matrices of order d with Laplace transform (7.12) then p belongs to the Gindikin set

$$\Lambda = \left\{ \frac{1}{2}, 1, \ldots, (d-1)\frac{1}{2} \right\} \cup \left(\frac{1}{2}(d-1), \infty \right).$$

Proof: First we need a variation on the Leibnitz formula. Let $\theta \mapsto f(\theta)$ and $\theta \mapsto g(\theta)$ be sufficiently differentiable real functions defined on the same open subset of $\mathbb{R}^n$. For $j = 1, \ldots, n$ we write $D_j = \frac{\partial}{\partial \theta_j}$. Then for $a = (a_1, \ldots, a_n) \in \mathbb{N}^n$

$$D_1^{a_1} \ldots D_n^{a_n}(fg)(\theta) = \sum \begin{pmatrix} a_1 \\ i_1 \end{pmatrix} \cdots \begin{pmatrix} a_n \\ i_n \end{pmatrix} D_1^{i_1} \ldots D_n^{i_n}(f)(\theta) D_1^{a_1 - i_1} \ldots D_n^{a_n - i_n}(g)(\theta) \tag{8.13}$$

where the sum is taken for all $i = (i_1, \ldots, i_n) \in \mathbb{N}^n$ such that $i_j \leq a_j,\ j = 1, \ldots, n$.

Let us now prove the proposition. We imitate the proof of Gindikin's theorem due to Shanbhag (1987) that we have seen above. Let $a \in \overline{\mathcal{P}}_d$. Suppose that there exists $p > 0$ and a positive measure $\mu_p(dt)$ on $\overline{\mathcal{P}}_d$ such that for all $\theta \in -\mathcal{P}_d$ one has

$$\frac{1}{(-\theta)^p} e^{\operatorname{tr}(a(-\theta)^{-1})} = \int_{\overline{\mathcal{P}}_d} e^{\operatorname{tr}(\theta t)} \mu_p(dt). \tag{8.14}$$

We show that $p \in \Lambda$.

Let Q be any real polynomial on the space of real symmetric matrices of order d. Then we have

$$Q(\frac{\partial}{\partial \theta}) \frac{1}{(-\theta)^p} e^{\operatorname{tr}(a(-\theta)^{-1})} = \int_{\overline{\mathcal{P}}_d} Q(t) e^{\operatorname{tr}(\theta t)} \mu_p(dt).$$

Suppose that the maximal degree of Q is n. Then there exists a real polynomial P_Q on $\mathbb{R}$ with respect to p such that

$$Q(\frac{\partial}{\partial \theta}) \frac{1}{(-\theta)^p} e^{\operatorname{tr}(a(-\theta)^{-1})} = \frac{1}{(-\theta)^{n+p}} e^{\operatorname{tr}(a(-\theta)^{-1})} P_Q(p). \tag{8.15}$$

Let us insist on the fact that the coefficients of P depend on θ and a. This result can be shown by using Leibnitz formula (8.13) applied to the pair $f(\theta) = e^{\operatorname{tr}(a\sigma(\theta))} = e^{\langle a, \sigma \rangle}$ and $g(\theta) = (\det \sigma(\theta))^p$, using induction on n. We now apply (8.15) to the polynomial $Q(t) = \det t$ whose degree is d to obtain

$$\frac{1}{(-\theta)^{d+p}} e^{\operatorname{tr}(a(-\theta)^{-1})} P_Q(p) = \int_{\overline{\mathcal{P}}_d} (\det t) e^{\operatorname{tr}(\theta t)} \mu_p(dt). \tag{8.16}$$

Note that the right hand side of (8.16) is ≥ 0. Note also that this right hand side is 0 for $p = 0, 1/2, \ldots, (d-1)/2$ since $\mu_0 = \delta_0$ and since $\mu_p(dt)$ is concentrated on singular matrices in the $d-1$ other cases. Now the left hand side of (8.16) has the same sign as $P_Q(p)$ which is a polynomial of degree $\leq d$ with at least zeros at $p = 0, 1/2, \ldots, (d-1)/3$. Furthermore, Proposition 3.2 shows that $P_Q(p) > 0$ for $p > (d-1)/2$. Thus $\deg P_Q = d$, and the zeros of P_Q are all real and simple. Also $(-1)^i P_Q(p) > 0$ for $\frac{d-1-i}{2} < p < \frac{d-i}{2}$ and $i = 1, \ldots, d-1$. Now, assume that a positive measure μ_p exists and that $p \notin \Lambda$. Then we would have $P_Q(p) > 0$ and therefore there would exist an even $i \in \{1, \ldots, d-1\}$ such that $\frac{d-1-i}{2} < p < \frac{d-i}{2}$. For $d = 2$ this is impossible. For $d \geq 3$ we observe that if μ_p exists, then

$$\mu_{p+\frac{1}{2}} = \mu_p * \mu_{\frac{1}{2}}$$

exists also, as can be seen using the Laplace transform. But now $P_Q(p + \frac{1}{2}) < 0$ which is a contradiction.

To complete the proof, suppose that there exists $p \notin \Lambda$ such that a probability $\gamma(p, a)$ on $\overline{\mathcal{P}}_d$ exists and such that for $I_d + s \in \mathcal{P}_k$ one has

$$\int_{\overline{\mathcal{P}}_d} e^{-\operatorname{tr}(st)} \gamma(p, a)(dt) = \frac{1}{\det(I_d + s)^p} e^{-\operatorname{tr}((I_d+s)^{-1} sa)}.$$

Defining $\mu_p(dt) = e^{\operatorname{tr}(t+a)} \gamma(p, a)(dt)$ we see that (8.14) holds. This contradiction ends the proof. $\square$

IX The rank problem for the non central Wishart and the Mayerhofer conjecture.

This is Eberhart Mayerhofer in 2010 (private communication) and in Mayerhofer (2013) who has observed that the fact that p belongs to the Gindikin set does not garantee that the distribution (7.12) exists. More specifically Mayerhofer (2013) shows that, if $NCW(2p, w, \Sigma)$ exists, if $d \geq 3$ and if $n = 2p$ is in $\{1, 2, \dots, d-2\}$, then $\operatorname{rank} w \leq n+1$ (reproved below in a different form in Proposition 3.4). E. Mayerhofer even conjectures that $\operatorname{rank} w \leq n$ must hold. The aim of the remainder of these notes is to give a satisfactory necessary and sufficient condition of existence of $NCW(2p, w, \Sigma)$: in Section 10, we state, essentially through Proposition 10.5, that if $n = 2p$ is in $\{1, 2, \dots, d-2\}$ then $\operatorname{rank} w \leq n$, thus stating that the Mayerhofer conjecture is true.

Finally, in Proposition 10.6 below, we write explicitely the necessary and sufficient condition of existence of the noncentral Wishart distribution $NCW(2p, w, \Sigma)$. In short, the remainder of these notes is devoted to the proofs of Proposition 10.4 (thus equivalent to the main result of Mayerhofer (2013)) and Proposition 10.5. (thus equivalent to the proof of his conjecture). These questions are delicate and Mayerhofer (2013) uses a stochastic process valued in the set of symmetric matrices in order to prove his main statement, in a paper difficult to read. Our methods are simpler and rely on geometry and classical analysis on the space of symmetric real matrices.

Section 10 begins with the easy reduction of the problem of the existence of $NCW(2p, w, d)$ as defined by (7.12) to the equivalent problem of the existence of the unbounded measure $m(2p, k, d)$ defined by formula (10.17) below. But the basic tool of the paper (Lemma 14.2) is the following. Let M_b be the set of positive measures concentrated on the set S_b of the matrices of rank b of $\overline{\mathcal{P}_d}$. It is not generally true that if $a + b \leq d$, $\mu \in M_a$ and $\nu \in M_b$, then $\mu * \nu \in M_{a+b}$, but it is true if either μ or ν is invariant by $x \mapsto uxu^{-1}$ for any orthogonal matrix u of order d. This result is the subject of Section 14.

Lemma 14.2 is, however, not sufficient to prove Propositions 10.4 and 10.5: we need further information about the measure $m(d-1, d, d)$ on $\overline{\mathcal{P}_d}$ defined by its Laplace transform $(\det s)^{-(d-1)/2} \exp(\operatorname{tr}(s^{-1}))$. We need to show that it has an absolutely continuous part. For this reason, Propositions 11.1 and 11.4 give a description of $m(1, 2, 2)$ and $m(d-1, d, d)$. Such a description provides more details than strictly necessary for proving Propositions 10.4 and 10.5. But for the challenging computation of the singular and absolutely continuous parts of $m(d-1, d, d)$ in Section 12, we need a very careful use of zonal polynomials. The form of the singular part of $m(d-1, d, d)$ is very surprizing and we are led to a correct guess by an elementary study when $d \geq 2$ in Section 11 which uses the Faà di Bruno formula only.

X Reduction of the problem: the measures $m(2p, k, d)$

Let k be an integer such that $0 \leq k \leq d$. We consider the diagonal matrix $I(k, d)$ with its first $d-k$ diagonal terms equal to 0 and the last k equal to 1.

$$I(k, d) = \begin{bmatrix} 0_{d-k} & 0 \\ 0 & I_k \end{bmatrix}$$

For $p \in \Lambda_d$ we define the positive measure $m(2p, k, d)$ on $\overline{\mathcal{P}_d}$ such that for all $s \in \mathcal{P}_d$ we have

$$\int_{\overline{\mathcal{P}_d}} e^{-\operatorname{tr}(sx)} m(2p, k, d)(dx) = \frac{1}{(\det s)^p} e^{\operatorname{tr}(s^{-1} I(k,d))}. \tag{10.17}$$

Note that $m(2p, k, d)$ may or may not exist. For example formula (11.44) below shows that the density of $m(1, 1, 1)$ on $(0, \infty)$ is

$$\frac{\cosh 2\sqrt{x}}{\sqrt{\pi x}}.$$

More generally with $p > 0$

$$m(0, 0, 1) = \delta_0, \qquad m(0, 1, 1) = \delta_0 + \left(\sum_{n=1}^{\infty} \frac{x^{n-1}}{n!(n-1)!} \right) 1_{(0,\infty)}(x) dx$$

$$m(p, 0, 1) = \tfrac{x^{p-1}}{\Gamma(p)} 1_{(0,\infty)}(x) dx, \quad m(p, 1, 1) = \left(\sum_{n=0}^{\infty} \frac{x^{n+p-1}}{n!\Gamma(n+p)} \right) 1_{(0,\infty)}(x) dx$$

For $2p > d - 1$ formula (12.72) gives $m(2p, d, d)$. If k and n are integers such that $0 \leq k \leq n \leq d$, formula (11.25) gives $m(n, k, d)$. Finally $m(1, 2, 2)$ is computed in Section 11. The measure $m(d-1, d, d)$ is computed in Section 13.3 and details about $m(2, 3, 3)$ are given in Section 13.5. The paper will show that these examples are the only cases of existence. For instance the function $s \mapsto \exp \operatorname{tr}(s^{-1})$ on $\mathcal{P}_d$ is not the Laplace transform of a positive measure if $d \geq 3$.

The following proposition links this unbounded measure $m(2p, k, d)$ with our initial existence problem. It is of a great practical importance for the solution of the problem of the existence or non existence of the non central Wishart with continuous shape parameter p : it reduces the problem to its core by forgetting the normalization constant and the parameter Σ, and by reducing the parameter w to its most important characteristic, namely its rank k.

Proposition 10.1. Let $\Sigma \in \mathcal{P}_d$, $w \in \overline{\mathcal{P}_d}$ and $p \in \Lambda_d$. Suppose that $\operatorname{rank} w = k$. Then $NCW(2p, w, \Sigma)$ as defined by (7.12) exists if and only if $m(2p, k, d)$ exists.

Proof. Assume that $m(2p, k, d)$ exists and let us show that $NCW(2p, w, \Sigma)$ exists. The proof is based on the following principle. Let μ be a positive measure on a finite dimensional real linear space E such that its Laplace transform $L_\mu(s) = \int_E e^{-\langle s,x \rangle} \mu(dx)$ is finite on some convex subset $D(\mu)$ of the dual space E^* with a non empty interior. Let a be a linear automorphism of E^* and let $b \in E^*$ such that $L_\mu(a(b)) < \infty$. Then there exists a probability $P(a, b)$ on E with Laplace transform $L_{P(a,b)}(s) = L_\mu(a(s + b))/L_\mu(a(b))$. This probability $P(a, b)$ is obtained in two steps: first take the image $\nu(dy)$ of $\mu(dx)$ by the map $x \mapsto a^*(x) = y$ where a^* is the adjoint of a. Its Laplace transform is $L_\nu(s) = L_\mu(a(s))$. The second step constructs $P(a, b)(dy)$ as the probability $e^{-\langle b,y \rangle} \nu(dy)/L_\mu(a(b))$: it is a member of the exponential family generated by ν.

Let us apply this to the case where $E = E^*$ is the Euclidean space of real symmetric matrices of order d with scalar product $\langle x, y \rangle = \operatorname{tr}(xy)$ and where μ is $m(2p, k, d)$. Here

$D(\mu) = \mathcal{P}_d$. We take $b = (2\Sigma)^{-1}$ and a to be the linear transformation $s \mapsto a(s) = qsq^*$ where q is an invertible matrix of order d such that

$$2(2\Sigma)^{-1}w(2\Sigma)^{-1} = q^{-1}I(k,d)(q^*)^{-1}. \tag{10.18}$$

We have $a^*(x) = q^*xq$. The distribution $P(a,b)$ is the noncentral Wishart $NCW(2p, w, \Sigma)$ since

$$\frac{L_\mu(a(s+b))}{L_\mu(a(b))} = \frac{1}{\det(I_d + 2\Sigma s)^p} e^{-\operatorname{tr}(2s(I_d+2\Sigma s)^{-1}w)}. \tag{10.19}$$

The verification of (10.19) is done by a calculation of trace using $\operatorname{tr}(ab) = \operatorname{tr}(ba)$ and (10.18):

$$\begin{aligned} &\operatorname{tr}\left[((q^*)^{-1}(s+(2\Sigma)^{-1})^{-1}q^{-1} - (q^*)^{-1}(2\Sigma)q^{-1})I(k,d)\right] \\ =\ &\operatorname{tr}\left[((s+(2\Sigma)^{-1})^{-1} - 2\Sigma)q^{-1}I(k,d)(q^*)^{-1}\right] = -\operatorname{tr}(2s(I_d+2\Sigma s)^{-1}w) \end{aligned}$$

The only thing left to prove is the existence of q satisfying (10.18). To see this, since the matrix $2(2\Sigma)^{-1}w(2\Sigma)^{-1}$ of $\overline{\mathcal{P}}_d$ has rank k, we write $2(2\Sigma)^{-1}w(2\Sigma)^{-1} = u\Delta u^*$ where

$$\Delta = \operatorname{diag}(0,\dots,0,\lambda_1^2,\dots,\lambda_k^2)$$

with $\lambda_i > 0$ and where u is an orthogonal matrix of order d. Taking

$$q = \operatorname{diag}(1,\dots,1,\lambda_1^{-1},\dots,\lambda_k^{-1})\, u^*$$

provides a solution of (10.18).

The proof of the converse follows similar lines. □

Example: When $0 \le k \le n \le d$ we can use the above principle for constructing $NCW(n, 2I(k,d), I_d)$ from $m(n,k,d)$. We take $q = I_d$ and $b = I_d/3$. Since a is the identity we have therefore

$$m(n,k,d)(dx) = 2^{dn/2}e^{2k}e^{\operatorname{tr} x/2}NCW(n, 2I(k,d), I_d)(dx) \tag{10.20}$$

The next three propositions reformulate known facts in the langage of the measures $m(2p,k,d)$.

Proposition 10.3. Let n and k be integers such that $0 \le n,k \le d$. The measure $m(n,k,d)$ exists for $0 \le k \le n \le d$. Furthermore, the measure $m(d-1,d,d)$ exists.

Remark. The proof of the existence of $m(d-1,d,d)$ given below is easy, but, as said before, its explicit computation as done in Section 4 will be hard work.

Proof. Formula (11.25) provides an explicit form of $m(n,k,d)$. For $2p > d-1$ the probability $NCW(2p, I_d, I_d)$ exists as proved in Letac and Massam (2008) Proposition 3.3. This implies that

$$\lim_{p \searrow (d-1)/2} NCW(2p, I_d, I_d) = NCW(d-1, I_d, I_d)$$

exists by considering the Laplace transforms. From Proposition 10.1 we have the result. $\square$

Proposition 10.4. Suppose $d \geq 4$. If $m(d-2, d-1, d)$ does not exist then $m(n, k, d)$ exists for no pairs of integers (n, k) such that $0 \leq n < k < d$. If $m(d-2, d, d)$ does not exist then $m(n, d, d)$ exists for no integer n such that $0 \leq n \leq d-2$.

Proof. Suppose that $m(n, k, d)$ exists for some pair $0 \leq n < k < d$. We define $m'(dx)$ as the measure on $\overline{\mathcal{P}_d}$ with Laplace transform

$$\int_{\overline{\mathcal{P}_d}} e^{-\operatorname{tr}(sx)} m'(dx) = \frac{1}{(\det s)^{\frac{d-n-2}{2}}} e^{\operatorname{tr}[s^{-1}(I(d-1,d)-I(k,d))]}$$

Since the rank of $I(d-1, d)) - I(k, d)$ is equal to $d-1-k$ and less than or equal to $d-n-2$ then m' exists by Propositions 10.1 and 10.3. Now we write the convolution

$$m(n, k, d) * m' = m(d-2, d-1, d)$$

which contradicts the non-existence of $m(d-2, d-1, d)$. Similarly, suppose that $m(d-2, d, d)$ does not exist and that there exists n such that $0 \leq n \leq d-2$ and such that $m(n, d, d)$ exists. Then $m(n, d, d) * m(d-2-n, 0, d) = m(d-2, d, d)$ also leads to a contradiction. $\square$

The idea of the proof of Proposition 10.3 is essentially due to Mayerhofer (2013). Here is now his important main result:

Proposition 10.5. If $d \geq 3$ the measure $m(d-2, d, d)$ does not exist.

Here is our main result:

Proposition 10.5. If $d \geq 3$ the measure $m(d-2, d-1, d)$ does not exist.

We will prove Proposition 10.5 in Section 15. In the remainder of these notes we develop the tools that lead us to this proof. They will also enable us to give a quick proof of Proposition 10.5. Let us emphasize the fact that Propositions 10.1 to 10.5 lead to a necessary and sufficient condition of existence of the distribution $NCW(2p, w, \Sigma)$. It is worthwhile to make the following synthesis:

Proposition 10.6. Let $\Sigma \in \mathcal{P}_d$, $w \in \overline{\mathcal{P}_d}$ with rank $k = 0, 1, \ldots, d$ and $p > 0$. Then the non central Wishart distribution $NCW(2p, w, \Sigma)$ exists if and only

1. Either $2p \geq d-1$;
2. Or $2p = n \in \{0, 1, \ldots, d-2\}$ and $0 \leq k \leq n$.

In particular for $d = 2$ the probability $NCW(2p, w, \Sigma)$ exists if and only if $2p \geq 1$ and for arbitrary dimension $NCW(0, w, \Sigma)$ exists if only if $w = 0$, being the Dirac measure.

Proof. From Proposition 10.1, the existence of $NCW(2p, w, \Sigma)$ is equivalent to the existence of $m(2p, k, d)$ when rank $w = k$.

$\boxed{\Rightarrow}$ Proposition 8.1 has shown that p is in the Gindikin set.

If $2p = n \in \{0, 1, \ldots, d-2\}$ let us show that $0 \leq k \leq n$. Suppose the contrary $0 \leq n < k$. A reformulation of the first part of Proposition 10.3 is the following : if there exists (n, k) such that $0 \leq n < k < d$ and such that $m(n, k, d)$ exists then $m(d-2, d-1, d)$ exists. This contradicts the statement of the present Proposition 10.5. Thus the 'if' part of Proposition 10.6 is proved.

$\boxed{\Leftarrow}$ If $2p \geq d-1$ Proposition 3.2 of Letac and Massam (2008) proves the existence of $m(2p, k, d)$ without constraints on k. Passing to the limit when $2p = d-1$ show the existence of $m(d-1, k, d)$ also for any k. If $2p = n \leq d-2$ and if $0 \leq k \leq n$ Proposition 10.2 of the present paper shows that $m(n, k, d)$ exists. $\square$

XI Computation of $m(1, 2, 2)$

In this section we compute $m(1, 2, 2)$ (which exists, from Proposition 10.2) using only calculus. We parameterize the cone $\overline{\mathcal{P}_2}$ by the cone of revolution

$$C = \{(x, y, z) \in \mathbb{R}^3; x \geq \sqrt{y^2 + z^2}\}$$

using the map φ from C to $\overline{\mathcal{P}_2}$ defined by

$$(x, y, z) \mapsto \varphi(x, y, z) = \begin{bmatrix} x+y & z \\ z & x-y \end{bmatrix}. \tag{11.21}$$

Note that $\operatorname{tr}[\varphi(a, b, c)\varphi(x, y, z)] = 2ax + 2by + 2cz$.

Proposition 11.1: Consider the positive measure μ on C such that for $a > \sqrt{b^2 + c^2}$ we have

$$\frac{1}{\sqrt{a^2 - b^2 - c^2}} e^{\frac{2a}{a^2 - b^2 - c^2}} = \int_C e^{-2ax - 2by - 2cz} \mu(dx, dy, dz), \tag{11.22}$$

that is to say such that the image of μ by φ is $m(1, 2, 2)$. Then

$$\mu(dx, dy, dz) = r(dx, dy, dz) + f(x, y, z) \mathbf{1}_C(x, y, z) dx dy dz$$

where the singular part r is the image of the measure $g(2\sqrt{y^2 + z^2}) dy dz$ on $\mathbb{R}^2$ by the map $(y, z) \mapsto (x, y, z) = (\sqrt{y^2 + z^2}, y, z)$ with

$$g(t) = \frac{2}{\pi t} \cosh(2\sqrt{t})$$

and where for $(x, y, z) \in C$

$$f(x, y, z) = \frac{2}{\sqrt{\pi}} \sum_{k=0}^{\infty} \frac{(x^2 - y^2 - z^2)^k}{k!(k+1)!} \sum_{m=0}^{\infty} \frac{1}{\Gamma(m + 2k + \frac{5}{2})} \frac{(2x)^m}{m!}.$$

Proof: Let $D = \frac{\partial}{\partial x}$. Recall the following differentiation formula.

Faà di Bruno formula . If $f(t)$ and $g(x)$ are functions with enough derivatives, then

$$D^n f(g(x)) = \sum \frac{n!}{k_1! \cdots k_n!}(D^k f)(g(x)) \left(\frac{Dg(x)}{1!}\right)^{k_1} \cdots \left(\frac{D^n g(x)}{n!}\right)^{k_n}, \tag{11.23}$$

where $k = k_1 + \cdots + k_n$ and where the sum is taken on all integers $k_j \geq 0$ such that $k_1 + 2k_2 + \cdots + nk_n = n$.

For a reference see for instance Roman (1980). We apply (11.23) to g defined by $x \mapsto x^2 - y^2 - z^2$ for fixed y, z and to $f(t) = t^n$. Noting that $D^3 g = 0$, we obtain

$$\frac{\partial^n}{\partial x^n}(x^2 - y^2 - z^2)^n = n!^2 \sum_{k_2=0}^{[n/2]} \frac{1}{k_2!} \times \frac{(x^2 - y^2 - z^2)^{k_2}}{k_2!} \times \frac{(2x)^{n-2k_2}}{(n-2k_2)!}. \tag{11.24}$$

For simplication in the sequel we write

$$E = e^{-2ax-2by-2cz}, \quad F = e^{-2a\sqrt{y^2+z^2}-2by-2cz} \tag{11.25}$$

We now recall (see Letac and Wesołowski (2008) formula 4.24) that for $p > 1/2$ we have for $a > \sqrt{b^2 + c^2}$

$$\frac{1}{(a^2 - b^2 - c^2)^p} = \frac{2}{\sqrt{\pi}} \times \frac{1}{\Gamma(p)\Gamma(p - \frac{1}{2})} \int_C (x^2 - y^2 - z^2)^{p-\frac{3}{2}} E dx dy dz \tag{11.26}$$

Define

$$\begin{aligned} I_k(n) &= (2a)^k \int_C (x^2 - y^2 - z^2)^n E dx dy dz \\ &= \frac{\sqrt{\pi}}{2} n! \Gamma(n + \frac{3}{2}) \times \frac{(2a)^k}{(a^2 - b^2 - c^2)^{n+\frac{3}{2}}} \end{aligned}$$

where we apply (11.26) for $p = n + \frac{3}{2}$. The idea of the proof is to write the Laplace transform of μ as follows:

$$\frac{e^{\frac{2a}{a^2-b^2-c^2}}}{\sqrt{a^2 - b^2 - c^2}} = \frac{1}{\sqrt{a^2 - b^2 - c^2}} + \sum_{n=0}^{\infty} \frac{(2a)^{n+1}}{(n+1)!} \frac{1}{(a^2 - b^2 - c^2)^{n+\frac{3}{2}}} \tag{11.27}$$

$$= \frac{1}{\sqrt{a^2 - b^2 - c^2}} + \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{1}{(n+1)! n! \Gamma(n + \frac{3}{2})} I_{n+1}(n) \tag{11.28}$$

A first step is to observe that for $k = 0, 1, \ldots, n$ we have

$$I_k(n) = \int_C \frac{\partial^k}{\partial x^k}(x^2 - y^2 - z^2)^n E dx dy dz. \tag{11.29}$$

Let us prove it by induction on k. It is true for $k = 0$. Suppose that it is true for $k < n$ and let us show that (11.29) is true for $k + 1$. Observe that for fixed (y, z) the root $\sqrt{y^2 + z^2}$ of the polynomial $x \mapsto (x^2 - y^2 - z^2)^n$ has order n and this implies that $\frac{\partial^k}{\partial x^k}(x^2 - y^2 - z^2)^n$

is zero for $x = \sqrt{y^2+z^2}$. Using this remark and integration by parts with $V(x) = e^{-2ax}$ and $U(x) = \frac{\partial^k}{\partial x^k}(x^2-y^2-z^2)^n$, we compute the following integral:

$$\int_{\sqrt{y^2+z^2}}^{\infty} 2ae^{-2ax}\frac{\partial^k}{\partial x^k}(x^2-y^2-z^2)^n dx = \int_{\sqrt{y^2+z^2}}^{\infty} e^{-2ax}\frac{\partial^{k+1}}{\partial x^{k+1}}(x^2-y^2-z^2)^n dx \quad (11.30)$$

With (11.30) we are in position to prove (11.29). We have

$$\begin{aligned} I_{k+1}(n) &= 2a\int_C \frac{\partial^k}{\partial x^k}(x^2-y^2-z^2)^n E dxdydz \\ &= \int_{\mathbb{R}^2} e^{-2by-2cz}\left[\int_{\sqrt{y^2+z^2}}^{\infty} 2ae^{-2ax}\frac{\partial^k}{\partial x^k}(x^2-y^2-z^2)^n dx\right] dydz \\ &= \int_{\mathbb{R}^2} e^{-2by-2cz}\left[\int_{\sqrt{y^2+z^2}}^{\infty} e^{-2ax}\frac{\partial^{k+1}}{\partial x^{k+1}}(x^2-y^2-z^2)^n dx\right] dydz \\ &= \int_C \frac{\partial^{k+1}}{\partial x^{k+1}}(x^2-y^2-z^2)^n E dxdydz \end{aligned}$$

which proves (11.29). We will need (11.29) only for $k = n$.

The second step is to express $I_{n+1}(n)$ as the Laplace transform of a positive measure. We compute $I_n(n)$ as expressed (11.29) by using again an integration by parts. The new fact for $k = n$ is that the integrated part will not disappear and will provide a term for the singular measure s given in the statement of the theorem. This calculation of the integrated part will use (11.24). Taking $V(x) = -e^{-2ax}$ and $U(x) = \frac{\partial^n}{\partial x^n}(x^2-y^2-z^2)^n$, we write

$$\begin{aligned} I_{n+1}(n) = 2aI_n(n) &= \int_{\mathbb{R}^2} e^{-2by-2cz}\left[\int_{\sqrt{y^2+z^2}}^{\infty} 2ae^{-2ax}\frac{\partial^n}{\partial x^n}(x^2-y^2-z^2)^n dx\right] dydz \\ &= A_n + S_n \qquad (11.31) \end{aligned}$$

with

$$\begin{aligned} A_n &= \int_C \frac{\partial^{n+1}}{\partial x^{n+1}}(x^2-y^2-z^2)^n E dxdydz \qquad (11.32) \\ S_n &= n!\int_{\mathbb{R}^2} e^{-2by-2cz}\left[-e^{-2ax}(2x)^n\right]_{\sqrt{y^2+z^2}}^{\infty} dydz \\ &= n!\int_{\mathbb{R}^2}(2\sqrt{y^2+z^2})^n F dydz \qquad (11.33) \end{aligned}$$

where (11.33) comes from (11.24) by keeping only the term $k_2 = 0$. We will carry this value of $I_{n+1}(n) = A_n + S_n$ in (11.28). Doing this, we can guess that S_n will contribute to the singular part of $m(1,2,2)$. But the term $\frac{1}{\sqrt{a^2-b^2-c^2}}$ in (11.28) will also contribute to it.

More specifically, the third step of the proof is to represent the function on $C \setminus \partial C$ defined by $(a,b,c) \mapsto \frac{1}{\sqrt{a^2-b^2-c^2}}$ as a Laplace transform. Using the Gaussian integral in

(11.34) we obtain

$$\frac{1}{\sqrt{a^2-b^2-c^2}} = \frac{2}{\pi}\int_{\mathbb{R}^2} e^{-2a(u^2+v^2)-2b(u^2-v^2)-4cuv}dudv \tag{11.34}$$
$$= \frac{2}{\pi}\int_{\mathbb{R}^2} (2\sqrt{y^2+z^2})^{-1}Fdydz \tag{11.35}$$

To derive (11.35) observe that the map on $\{(u,v); u>0\}$ defined by $y=u^2-v^2,\ z=2uv$ is a bijection with $\mathbb{R}^2$; the same is true with $\{(u,v); u<0\}$. Furthermore $dydz = 4(u^2+v^2)dudv = 4\sqrt{y^2+z^2}dudv$ and therefore $dudv = \frac{dydz}{4\sqrt{y^2+z^2}}$. All of this leads to (11.35).

Now comes the fourth and final step. We use $\Gamma(n+\frac{1}{2}) = \frac{(2n)!}{4^n n!}\sqrt{\pi}$ and we consider the function

$$g(t) = \frac{2}{\pi t} + \frac{2}{\sqrt{\pi}}\sum_{n=0}^{\infty}\frac{1}{(n+1)!\Gamma(n+\frac{3}{2})}t^n = \frac{2}{\pi t}\cosh(2\sqrt{t})$$

We then define the measure $r(dx,dy,dz)$ concentrated on the boundary

$$\partial C = \{(x,y,z) : x = \sqrt{y^2+z^2})\}$$

of the cone C to be the image of the measure on $\mathbb{R}^2$

$$g(2\sqrt{y^2+z^2})dydz = \frac{1}{\pi\sqrt{y^2+z^2}}\cosh\left(2^{3/2}(y^2+z^2)^{1/4}\right)dydz \tag{11.36}$$

by the map $(y,z)\mapsto(x,y,z) = (\sqrt{y^2+z^2}, y, z)$.This r will be the singular part of the image μ of $m(1,2,2)$ by the reciprocal of φ defined by (11.21). Here is the details

$$\int_C Eds = \int_{\mathbb{R}^2} g(2\sqrt{y^2+z^2})Fdydz$$
$$= \int_{\mathbb{R}^2} F\frac{dydz}{\pi\sqrt{y^2+z^2}} + \frac{2}{\sqrt{\pi}}\sum_{n=0}^{\infty}\int_{\mathbb{R}^2}\frac{2^n(\sqrt{y^2+z^2})^n}{(n+1)!\Gamma(n+\frac{3}{2})}Fdydz$$
$$= \frac{1}{\sqrt{a^2-b^2-c^2}} + \frac{2}{\sqrt{\pi}}\sum_{n=0}^{\infty}\frac{S_n}{(n+1)!n!\Gamma(n+\frac{3}{2})} \tag{11.37}$$

Finally we focus on the absolutely continuous part of μ. We will need the following formula, similar to (11.24) and also obtained by the Faà di Bruno formula (11.23):

$$\frac{\partial^n}{\partial x^n}(x^2-y^2-z^2)^{n-1} = n!(n-1)!\sum_{k_2=1}^{[n/2]}\frac{1}{(k_2-1)!}\times\frac{(x^2-y^2-z^2)^{k_2-1}}{k_2!}\times\frac{(2x)^{n-2k_2}}{(n-2k_2)!}. \tag{11.38}$$

The absolutely continuous part of μ is given by (11.32) and (11.28). Its density is

$$f(x,y,z) = \frac{2}{\sqrt{\pi}}\sum_{n=0}^{\infty}\frac{1}{(n+1)!n!\Gamma(n+\frac{3}{2})}\frac{\partial^{n+1}}{\partial x^{n+1}}(x^2-y^2-z^2)^n \tag{11.39}$$

$$= \frac{2}{\sqrt{\pi}}\sum_{n=2}^{\infty}\frac{1}{(n-1)!n!\Gamma(n+\frac{1}{2})}\frac{\partial^{n}}{\partial x^{n}}(x^2-y^2-z^2)^{n-1}$$

$$= \frac{2}{\sqrt{\pi}}\sum_{n=2}^{\infty}\frac{1}{\Gamma(n+\frac{1}{2})}\sum_{k_2=1}^{[n/2]}\frac{1}{(k_2-1)!}\times\frac{(x^2-y^2-z^2)^{k_2-1}}{k_2!}\times\frac{(2x)^{n-2k_2}}{(n-2k_2)!}$$

$$= \frac{2}{\sqrt{\pi}}\sum_{k=0}^{\infty}\frac{(x^2-y^2-z^2)^k}{k!(k+1)!}\sum_{m=0}^{\infty}\frac{1}{\Gamma(m+2k+\frac{5}{2})}\frac{(2x)^m}{m!} \tag{11.40}$$

From (11.39), (11.28) and (11.32) the Laplace transform of f is

$$\int_C Ef dxdydz = \frac{2}{\sqrt{\pi}}\sum_{n=0}^{\infty}\frac{A_n}{(n+1)!n!\Gamma(n+\frac{3}{2})} \tag{11.41}$$

Combine (11.37) and (11.41) by adding them and use (11.28). This shows that the parameterization μ of $m(1,2,2)$ by φ is the sum of r and of the absolutely continous part with density f. Formula 11.41 shows that f has the form announced in the statement of Proposition 4.1. $\square$

Remark : This remark is essential for the understanding of Section 13. The image by φ of the measure $r(dx,dy,dz)$ is concentrated on the set $S_1 \subset \overline{\mathcal{P}}_2$ of matrices of rank one. Any element of S_1 can be written as $u\begin{bmatrix}\lambda_1 & 0\\ 0 & 0\end{bmatrix}u^*$ where u is an orthogonal matrix of $\mathbb{O}(2)$ and $\lambda_1 > 0$. We can compute the image of $r(dx,dy,dz)$ by the map

$$\begin{bmatrix}x+y & z\\ z & x-y\end{bmatrix} = \begin{bmatrix}\sqrt{y^2+z^2}+y & z\\ z & \sqrt{y^2+z^2}-y\end{bmatrix} \mapsto \lambda_1 = 2\sqrt{y^2+z^2}. \tag{11.42}$$

If $A_t = \{(x,y,z)\ ;\ 2\sqrt{y^2+z^2} < t\}$, then using polar coordinates $y = \lambda_1\cos\alpha, z = \lambda_1\sin\alpha$ with Jacobian equal to $\frac{\lambda_1}{2}$, we have

$$r(A_t) = \int_{A_t} r(dx,dy,dz) = \int_{2\sqrt{y^2+z^2}<t} g(2\sqrt{y^2+z^2})dydz = \frac{\pi}{2}\int_0^t g(\lambda_1)\lambda_1 d\lambda_1.$$

Since $g(\lambda_1) = \frac{2}{\pi\lambda_1}\cosh 2\sqrt{\lambda_1}$, the image of the measure r by the map (11.42) is

$$\cosh(2\sqrt{\lambda_1})\mathbf{1}_{(0,\infty)}(\lambda_1)d\lambda_1. \tag{11.43}$$

Now an important observation is the following: consider the measure $m(1,1,1)(d\lambda)$ on $(0,\infty)$ whose Laplace transform is for $s > 0$:

$$\frac{1}{\sqrt{s}}e^{1/s} = \sum_{n=0}^{\infty}\frac{1}{n!s^{n+\frac{1}{2}}} = \int_0^{\infty}e^{-s\lambda}\sum_{n=0}^{\infty}\frac{\lambda^{n-\frac{1}{2}}}{n!\Gamma(n+\frac{1}{2})}d\lambda = \frac{1}{\sqrt{\pi}}\int_0^{\infty}e^{-s\lambda}\frac{1}{\sqrt{\lambda}}\cosh(2\sqrt{\lambda})d\lambda.$$

As a consequence

$$m(1,1,1)(d\lambda) = \frac{1}{\sqrt{\pi}}\frac{1}{\sqrt{\lambda}}\cosh(2\sqrt{\lambda})\mathbf{1}_{(0,\infty)}(\lambda)d\lambda \tag{11.44}$$

and one observes that $m(1,1,1)$ is quite close to (11.43). To summarize this remark, the singular part of $m(1,2,2)$ can be seen as the image of $\sqrt{\pi\lambda_1}m(1,1,1)(d\lambda_1)\otimes du$ by the map $(u,\lambda_1)\mapsto u\begin{bmatrix}\lambda_1 & 0\\ 0 & 0\end{bmatrix}u^*$ from $(0,\infty)\times\mathbb{O}(2)$ where du is the uniform probability on $\mathbb{O}(2)$. This is the key to the generalization of the computation of $m(1,2,2)$ to the computation of $m(d-1,d,d)$ for $d\geq 2$ done in Proposition 5.4 below.

XII Computation of the measure $m(d-1,d,d)$

Before stating Proposition 13.4 which describes $m(d-1,d,d)$ we have to fix some notations, to recall a few facts about zonal functions and polynomials and to prove three lemmas. The Lebesgue measure dx on the space of real symmetric matrices of order d has the normalization associated to the Euclidean structure given by $\langle x,y\rangle = \mathrm{tr}\,(xy)$. Note that Muirhead (1983) has a different normalization. As mentioned in page ix of the introduction of Muirhead, zonal functions are the essential tool for the noncentral distribution theory.

XII.1 Zonal functions

Let $\mathcal{E}_d$ be the set of sequences $\kappa=(m_1,\ldots,m_d)$ of d integers such that

$$m_1\geq m_2\geq\ldots\geq m_d\geq 0.$$

If $\kappa\in\mathcal{E}_d$ we consider the two zonal polynomials

$$\Phi^{(d)}_\kappa(x)=\Phi^{(d)}_{m_1,\ldots,m_d}(x),\quad C^{(d)}_\kappa(x)=C^{(d)}_\kappa(I_d)\Phi^{(d)}_\kappa(x)$$

where $C^{(d)}_\kappa(I_d)$ is defined below in (12.46). In FK page 228 the (Φ_κ) are rather called spherical polynomials, and page 234 the notation Z_κ is used for the above $C^{(d)}_\kappa$. We use the definitions given in FK, while Muirhead (1983) and Takemura (1984) have other ways to introduce the zonal polynomials. To define $\Phi^{(d)}_\kappa$ we consider

$$\Delta_\kappa(x)=\Delta_1(x)^{m_1-m_2}\Delta_2(x)^{m_2-m_3}\ldots\Delta_{d-1}(x)^{m_{d-1}-m_d}\Delta_d(x)^{m_d}$$

where for $x=(x_{ij})_{1\leq i,j\leq d}$ a real symmetric matrix, $\Delta_k(x)=\det(x_{ij})_{1\leq i,j\leq k}$ is the principal determinant of x of order k. The function $\Phi^{(d)}_\kappa$ is defined by

$$\Phi^{(d)}_\kappa(x)=\int_{\mathbb{O}(d)}\Delta_\kappa(uxu^*)du \tag{12.45}$$

where du is the Haar probability on the orthogonal group $\mathbb{O}(d)$. When $x\in\mathcal{P}_d$ definition (12.45) makes sense even when $m_1,\ldots,m_d$ are complex numbers. In that case $\Phi^{(d}_{m_1,\ldots,m_d}(x)$

is no longer a polynomial and is called a zonal function. To give the value of the constant $C^{(d}_\kappa(I_d)$ we need the notations $\ell(\kappa)=\max\{j; m_j>0\}$, $|\kappa|=m_1+\ldots+m_d$ and

$$\Gamma_d(z_1,\ldots,z_d)=\prod_{j=1}^d\Gamma(z_j-\frac{j-1}{2})$$

defined for $z_j-\frac{j-1}{2}>0,\quad j=1,\ldots,d$. If p is a real number, we use the notational convention

$$\Gamma_d(z+p)=\Gamma_d(z_1+p,\ldots,z_d+p).$$

If $\kappa\in\mathcal{E}_d$ and if $p>(d-1)/2$ we define the Pochhammer symbol

$$(p)_\kappa=\Gamma_d(\kappa+p)/\Gamma_d(p).$$

If $\kappa\in\mathcal{E}_d$ the constant $C^{(d}_\kappa(I_d)$ is

$$\begin{aligned} C^{(d)}_\kappa(I_d) &= C^{(d)}_{m_1,\ldots,m_d}(I_d)=2^{2|\kappa|}|\kappa|!\left(\frac{d}{2}\right)_\kappa\frac{\prod_{1\leq i<j\leq \ell(\kappa)}(2m_i-2m_j-i+j)}{\prod_{i=1}^{\ell(\kappa)}(2m_i+\ell(\kappa)-i)!} &(12.46)\\ &= \frac{|\kappa|!}{(\frac{d+1}{2})_\kappa}\times\prod_{1\leq i\leq j\leq d}\frac{B(\frac{1}{2}(j-i+1),\frac{1}{2})}{B(m_i-m_j+\frac{1}{2}(j-i+1),\frac{1}{2})}. &(12.47)\end{aligned}$$

On the form (12.46) it is given in Muirhead page 237 formula (38) where he refers to Constantine (1963) for a proof. On the form (12.47) it is proved in detail in FK by combining Propositions XI. 5.1 (i) page 230 and XI.5.3 page 232 and the definition of $Z_\kappa(x)=C^{(d)}_\kappa(x)$ on the last line of page 235. We never consider $C^{(d}_\kappa(x)$ if $\kappa\notin\mathcal{E}_d$. The exact value of $C^{(d}_\kappa(I_d)$ will be crucial in the proof of Proposition 5.4 when we shall need the formula (3) page 259 of Muirhead (1983):

$$e^{\operatorname{tr}x}=\sum_{\kappa\in\mathcal{E}_d}\frac{1}{|\kappa|!}C^{(d}_\kappa(x) \tag{12.48}$$

It is worthwhile to perform the calculation of the zonal polynomials for $d=3$. One can consult FK Exercise 5 page 237. Define the Legendre polynomials $(P_k)_{k=0}^\infty$ by their generating function

$$\sum_{k=0}^\infty P_k(x)z^k=\frac{1}{\sqrt{1-2zx+z^2}}.$$

Let $x=\left[\begin{array}{cc} a+b & c\\ c & a-b\end{array}\right]$ in $\mathcal{P}_3$. Then for $(m_1,m_2)\in\mathcal{E}_2$ we have

$$\Phi_{m_1,m_2}(x)=(a^2-b^2-c^2)^{(m_1+m_2)/2}P_{m_1-m_2}(\frac{a}{\sqrt{a^2-b^2-c^2}}). \tag{12.49}$$

For convenience of the reader we detail the proof of (12.49). The Legendre polynomial P_k satisfies

$$P_k(\cosh t)=\frac{1}{\pi}\int_0^\pi(\cosh t+\cos u\sinh t)^k du. \tag{12.50}$$

To check this, call Q_k the right hand side of (12.50). The computation of $\sum_{k=0}^{\infty} Q_k z^k$ gives $(1-2zx+z^2)^{-1/2}$ for $|z|$ small enough. This proves (12.50). Denote

$$R(\theta)=\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},\quad J=\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

and observe that

$$\mathbb{SO}(2)=\{R(\theta);\theta\in\mathbb{R}\},\quad \mathbb{O}(2)\setminus\mathbb{SO}(2)=J\mathbb{SO}(2)$$

Writing $\begin{pmatrix} B \\ C \end{pmatrix}=R(2\theta)\begin{pmatrix} b \\ c \end{pmatrix}$ we have from a little calculation

$$R(\theta)\begin{bmatrix} a+b & c \\ c & a-b \end{bmatrix}R(-\theta) = \begin{bmatrix} a+B & C \\ C & a-B \end{bmatrix}, \tag{12.51}$$

$$JR(\theta)\begin{bmatrix} a+b & c \\ c & a-b \end{bmatrix}R(-\theta)J = \begin{bmatrix} a+B & -C \\ -C & a-B \end{bmatrix} \tag{12.52}$$

The two formulas (12.51) and (12.52) enable us to compute the zonal polynomial:

$$\begin{aligned}
\Phi_{m_1,m_2}(x) &= \int_{\mathbb{O}(2)} \Delta_{m_1,m_2}(uxu^*)du \\
&= \frac{1}{2}\int_{\mathbb{SO}(2)} \Delta_{m_1,m_2}(uxu^*)du+\frac{1}{2}\int_{\mathbb{O}(2)\setminus\mathbb{SO}(2)} \Delta_{m_1,m_2}(uxu^*)du \\
&= \frac{1}{4\pi}\int_0^{2\pi} (\Delta_{m_1,m_2}(R(\theta)xR(-\theta))+\Delta_{m_1,m_2}(JR(\theta)xR(-\theta)J))\,d\theta \\
&= (a^2-b^2-c^2)^{m_2}\frac{1}{2\pi}\int_0^{2\pi}(a+B)^{m_1-m_2}d\theta \\
&= (a^2-b^2-c^2)^{m_2}\frac{1}{\pi}\int_0^{\pi}(a+\sqrt{b^2+c^2}\cos\theta)^{m_1-m_2}d\theta \\
&= (a^2-b^2-c^2)^{\frac{m_1+m_2}{2}}\frac{1}{\pi}\int_0^{\pi}\left(\frac{a}{\sqrt{a^2-b^2-c^2}}+\frac{\sqrt{b^2+c^2}}{\sqrt{a^2-b^2-c^2}}\cos\theta\right)^{m_1-m_2}d\theta
\end{aligned}$$

The form given in (12.50) of the Legendre polynomial concludes the proof of (12.49). The application of (12.47) also provides

$$C^{(2)}_{m_1,m_2}(I_2)=\frac{(m_1+m_2)!}{(m_1-m_2)!m_2!}\times\frac{1}{(\frac{3}{2}+m_1-m_2)_{m_2}}. \tag{12.53}$$

XII.2 Three properties of zonal functions

We are indebted to Jacques Faraut for the next lemma:

Lemma 13.1: If $x=\begin{bmatrix} x_1 & x_{12} \\ x_{21} & x_2 \end{bmatrix}\in\mathcal{P}_d$ we denote $[x]_1=x_1\in\mathcal{P}_{d-1}$. Then for all complex numbers $m_1,\dots,m_d$ we have

$$\Phi^{(d)}_{m_1,\dots,m_d}(x)=(\det x)^{m_d}\int_{\mathbb{O}(d)}\Phi^{(d-1)}_{m_1,\dots,m_{d-1}}([uxu^*]_1)du.$$

Proof: Consider $v = \begin{bmatrix} v_1 & 0 \\ 0 & 1 \end{bmatrix} \in \mathbb{O}(d)$ where $v_1 \in \mathbb{O}(d-1)$. Observe that for any $y \in \mathcal{P}_d$ we have

$$[vyv^*]_1 = v_1[y]_1v_1^*. \tag{12.54}$$

We write

$$\begin{aligned}
\Phi^{(d)}_{m_1,\ldots,m_d}(x) &= \int_{\mathbb{O}(d)} \Delta_{m_1,\ldots,m_{d-1},m_d}(uxu^*)du && (12.55)\\
&= (\det x)^{m_d}\int_{\mathbb{O}(d)} \Delta_{m_1,\ldots,m_{d-1}}([uxu^*]_1)du && (12.56)\\
&= (\det x)^{m_d}\int_{\mathbb{O}(d)} \Delta_{m_1,\ldots,m_{d-1}}([vuxu^*v^*]_1)du && (12.57)\\
&= (\det x)^{m_d}\int_{\mathbb{O}(d)} \Delta_{m_1,\ldots,m_{d-1}}(v_1[uxu^*v^*]_1v_1^*)du && (12.58)\\
&= (\det x)^{m_d}\int_{\mathbb{O}(d)} \left(\int_{\mathbb{O}(d-1)} \Delta_{m_1,\ldots,m_{d-1}}(v_1[uxu^*v^*]_1v_1^*)dv_1\right) du && (12.59)\\
&= (\det x)^{m_d}\int_{\mathbb{O}(d)} \Phi^{(d-1)}_{m_1,\ldots,m_{d-1}}([uxu^*]_1)du && (12.60)
\end{aligned}$$

In this list (12.55) comes from the definition of $\Phi^{()}_{\kappa}(x)$, (12.56) separates the roles of $[uxu^*]_1$ and $\det(uxu^*) = \det x$ in the definition of $\Delta_\kappa(uxu^*)$, (12.57) uses the fact that du is the Haar probability, (12.58) applies (12.54) to $y = uxu^*$, (12.59) uses the fact that the Haar measure dv_1 of $\mathbb{O}(d-1)$ has mass 1, (12.60) comes from the definition of $\Phi^{()}_{m_1,\ldots,m_{d-1}}(x)$. $\square$

Lemma 13.2: If $x \in \mathcal{P}_d$ then $\Phi^{(d)}_{m_1,\ldots,m_d}(x^{-1}) = \Phi^{(d)}_{-m_d,\ldots,-m_1}(x)$

Proof: Define $p \in \mathbb{O}(d)$ by

$$p = \begin{bmatrix} 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & \ldots & 1 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & 0 & \ldots & 0 & 0 \end{bmatrix}$$

and define $\Delta^*_{m_1,\ldots,m_{d-1},m_d}(x) = \Delta_{m_1,\ldots,m_{d-1},m_d}(pxp^*)$. We can write $\Phi^{(d)}_{\kappa}(x^{-1})$ as

$$\begin{aligned}
\int_{\mathbb{O}(d)} \Delta_{m_1,\ldots,m_{d-1},m_d}(ux^{-1}u^*)du &= \int_{\mathbb{O}(d)} \Delta^*_{-m_d,\ldots,-m_1}(uxu^*)du && (12.61)\\
&= \int_{\mathbb{O}(d)} \Delta_{-m_d,\ldots,-m_1}(uxu^*)du && (12.62)
\end{aligned}$$

In this list (12.61) comes from the FK formula Proposition VII.1.5 (ii) page 127 which says

$$\Delta_{m_1,\ldots,m_{d-1},m_d}(x^{-1}) = \Delta^*_{-m_d,\ldots,-m_1}(x),$$

and (12.62) comes from the invariance of the Haar probability du on $\mathbb{O}(d)$ by $u \mapsto pu$. $\square$

Lemma 13.3: If $x \in \mathcal{P}_d$ then $\Phi^{(d)}_{m_1,\dots,m_d}(x)(\det x)^p = \Phi^{(d)}_{m_1+p,\dots,m_d+p}(x)$.

Proof: We have from the definition $\Phi^{(d)}_\kappa(x)(\det x)^p =$

$$\int_{\mathbb{O}(d)} \Delta_{m_1,\dots,m_d}(uxu^*)(\det uxu^*)^p du = \int_{\mathbb{O}(d)} \Delta_{m_1+p,\dots,m_d+p}(uxu^*)du = \Phi^{(d)}_{m_1+p,\dots,m_d+p}(x) \ \square$$

XII.3 The calculation of $m(d-1,d,d)$

Proposition 13.5. Define the singular measure $r(dt)$ on $\overline{\mathcal{P}_d}$ and concentrated on the set S_{d-1} of symmetric matrices of rank $d-1$ as the image of the product measure

$$\frac{(\pi \det x)^{1/2}}{\Gamma(d/2)} m(d-1,d-1,d-1)(dx) \otimes du$$

by the map from $\mathcal{P}_{d-1} \times \mathbb{O}(d)$ to $\overline{\mathcal{P}_d}$ which is $(x,u) \mapsto t = u \begin{bmatrix} x & 0 \\ 0 & 0 \end{bmatrix} u^* = u\tilde{x}u^*$. Define

$$f_d(t) = (\det t)^{-1} \left(\sum_{\kappa \in \mathcal{E}'_d} \frac{C^{(d)}_\kappa(t)}{|\kappa|!\Gamma_d(\kappa + \frac{d-1}{2})} \right). \tag{12.63}$$

where $\mathcal{E}'_d = \{\kappa \in \mathcal{E}_d \ ; \ m_d > 0\}$. Then

$$m(d-1,d,d)(dt) = r(dt) + f_d(t)\mathbf{1}_{\mathcal{P}_d}(t)dt.$$

Proof. The fonction $f_d(t)$ is a well defined analytic function around $t = 0$ since from the definition (12.45) of $\Phi^{()}_\kappa$ the polynomial $C^{(d)}_{m_1,\dots,m_d}(t)$ is divisible by $(\det t)^{m_d}$. Therefore $(\det t)^{-1} C^{(d)}_\kappa(t)$ is a polynomial when $\kappa \in \mathcal{E}'_d$. Recall the basic fact (see FK, Lemma XI.3.3 page 226 or Muirhead Theorem 7.3.7 page 248):

$$\int_{\mathcal{P}_d} e^{-\operatorname{tr}(sx)}(\det x)^{p-\frac{d+1}{2}} \frac{\Phi^{(d)}_\kappa(x)}{\Gamma_d(\kappa+p)} dx = \Phi^{(d)}_\kappa(s^{-1})(\det s)^{-p}. \tag{12.64}$$

Note that the choice of the suitable Lebesgue measure dx is crucial in (12.64). This formula (12.64) is correct for $p + m_d > (d-1)/3$. This comes from the references above for $m_d = 0$. If $m_d > 0$ we observe that

$$\Phi^{(d)}_\kappa(x) = \Phi^{(d)}_{m_1,\dots,m_d}(x) = (\det x)^{m_d} \Phi^{(d)}_{m_1-m_d,\dots,m_{d-1}-m_d,0}(x) = (\det x)^{m_d} \Phi^{(d)}_{\kappa-m_d}(x). \tag{12.65}$$

As a consequence

$$\begin{aligned} & \int_{\mathcal{P}_d} e^{-\operatorname{tr}(sx)}(\det x)^{p-\frac{d+1}{2}} \frac{\Phi^{(d)}_\kappa(x)}{\Gamma_d(\kappa+p)} dx \\ = & \int_{\mathcal{P}_d} e^{-\operatorname{tr}(sx)}(\det x)^{p+m_d-\frac{d+1}{2}} \frac{\Phi^{(d)}_{\kappa-m_d}(x)}{\Gamma_d(\kappa+p)} dx & (12.66) \\ = & \ \Phi^{(d)}_{\kappa-m_d}(s^{-1})(\det s)^{-p-m_d} & (12.67) \\ = & \ \Phi^{(d)}_\kappa(s^{-1})(\det s)^{-p}. & (12.68) \end{aligned}$$

where (12.66) and (12.68) come from (12.65) and (12.67) comes from (12.64) where p is replaced by $p + m_d$.

From (12.48) we know that, for $2p \geq d-1$, the Laplace transform of $m(2p, d, d)$ is

$$\int_{\mathcal{P}_d} e^{-\operatorname{tr}(sx)} m(2p, d, d)(dx) = (\det s)^{-p} \sum_{\kappa \in \mathcal{E}_d} \frac{C_\kappa^{(d)}(s^{-1})}{|\kappa|!} \tag{12.69}$$

Observe that the Laplace transform of $f_d(t)\mathbf{1}_{\mathcal{P}_d}(t)dt$ as defined by (12.63) is easily deduced from (12.64) and is

$$\int_{\mathcal{P}_d} e^{-\operatorname{tr}(st)} f_d(t)dt = (\det s)^{-\frac{d-1}{2}} \left(\sum_{\kappa \in \mathcal{E}'_d} \frac{C_\kappa^{(d)}(s^{-1})}{|\kappa|!} \right) \tag{12.70}$$

In (12.69) take $2p = d-1$. From the Laplace transform (12.70) our aim is therefore to prove that the Laplace transform of $r(dt)$ is

$$\int_{\mathcal{P}_d} e^{-\operatorname{tr}(st)} r(dt) = (\det s)^{-\frac{d-1}{2}} \left(\sum_{\kappa \in \mathcal{E}_d \setminus \mathcal{E}'_d} \frac{C_\kappa^{(d)}(s^{-1})}{|\kappa|!} \right) = (\det s)^{-\frac{d-1}{2}} \left(\sum_{\kappa \in \mathcal{E}_{d-1}} \frac{C_{(\kappa,0)}^{(d)}(s^{-1})}{|\kappa|!} \right) \tag{12.71}$$

To prove (12.71) we undertake the calculation of the Laplace transform of r from its definition. Observe first that (12.64) and (12.69) imply for $2p > d-1$

$$m(2p, d, d)(dx) = (\det x)^{p - \frac{d+1}{2}} \left(\sum_{\kappa \in \mathcal{E}_d} \frac{C_\kappa^{(d)}(x)}{|\kappa|! \Gamma_d(\kappa + p)} \right) \mathbf{1}_{\mathcal{P}_d}(x)dx. \tag{12.72}$$

In particular, in (12.72) let us replace d by $d-1$ and do $2p = d-1$. We get

$$(\det x)^{\frac{1}{2}} m(d-1, d-1, d-1)(dx) = \left(\sum_{\kappa \in \mathcal{E}_{d-1}} \frac{C_\kappa^{(d-1)}(x)}{|\kappa|! \Gamma_{d-1}(\kappa + \frac{d-1}{2})} \right) \mathbf{1}_{\mathcal{P}_{d-1}}(x)dx. \tag{12.73}$$

We can now write

$$\begin{aligned}
\int_{\overline{\mathcal{P}}_d} e^{-\operatorname{tr}(st)} r(dt) &= \frac{\pi^{1/2}}{\Gamma(d/2)} \int_{\mathbb{O}(d)} \left(\int_{\mathcal{P}_{d-1}} e^{-\operatorname{tr}(su\tilde{x}u^*)} \sum_{\kappa \in \mathcal{E}_{d-1}} \frac{C_\kappa^{(d-1)}(x)}{|\kappa|! \Gamma_{d-1}(\kappa + \frac{d-1}{2})} dx \right) du \qquad (12.74) \\
&= \frac{\pi^{1/2}}{\Gamma(d/2)} \sum_{\kappa \in \mathcal{E}_{d-1}} \frac{C_\kappa^{(d-1)}(I_{d-1})}{|\kappa|!} \int_{\mathbb{O}(d)} \left(\int_{\mathcal{P}_{d-1}} e^{-\operatorname{tr}([u^*su]_1 x)} \frac{\Phi_\kappa^{(d-1)}(x)}{\Gamma_{d-1}(\kappa + \frac{d-1}{2})} dx \right) du
\end{aligned}$$

This equality (12.74) comes from (12.73) and from the definition of r. Now we compute the last double integral as follows

$$
\begin{aligned}
&\int_{\mathbb{O}(d)}\left(\int_{\mathcal{P}_{d-1}} e^{-\operatorname{tr}([u^*su]_1x)}\frac{\Phi_\kappa^{(d-1)}(x)}{\Gamma_{d-1}(\kappa+\frac{d-1}{2})}dx\right)du\\
=&\int_{\mathbb{O}(d)}(\det[usu^*]_1^{-1})^{\frac{d}{2}}\Phi_\kappa^{(d-1)}([usu^*]_1^{-1})du && (12.75)\\
=&\int_{\mathbb{O}(d)}\Phi_{\kappa+\frac{d}{2}}^{(d-1)}([usu^*]_1^{-1})du && (12.76)\\
=&\int_{\mathbb{O}(d)}\Phi_{-m_{d-1}-\frac{d}{2},\ldots,-m_1-\frac{d}{2}}^{(d-1)}([usu^*]_1)du && (12.77)\\
=&\ \Phi_{-m_{d-1}-\frac{d}{2},\ldots,-m_1-\frac{d}{2},0}^{(d)}(s) && (12.78)\\
=&\ \Phi_{0,m_1+\frac{d}{2},\ldots,m_{d-1}+\frac{d}{2}}^{(d)}(s^{-1}) && (12.79)\\
=&\ \Phi_{-\frac{d}{2},m_1,\ldots,m_{d-1}}^{(d)}(s^{-1})(\det s^{-1})^{\frac{d}{2}} && (12.80)\\
=&\ \Phi_{m_1,\ldots,m_{d-1},0}^{(d)}(s^{-1})(\det s^{-1})^{\frac{d-1}{2}} && (12.81)
\end{aligned}
$$

In this list equality (12.75) comes from (12.64) by replacing (d,p) by $(d-1,d/2)$. Equalities (12.76) and (12.80) come from Lemma 5.4. In identities following (12.77) we have replaced κ by $(m_1,\ldots,m_{d-1})$ for clarity. Formulas (12.77) and (12.79) come from Lemma 13.2, and (12.78) comes from Lemma 13.1. The proof of (12.81) is more involved and is a consequence of formula (iii) of Theorem XIV 4.1 of FK where we replace (d,r,λ,μ) respectively by $1,d$ and

$$
\begin{aligned}
\lambda &= (m_1+\frac{d-1}{4},m_2+\frac{d-3}{4},\ldots,m_{d-1}-\frac{d-3}{4},-\frac{d-1}{4}),\\
\mu &= (-\frac{d-1}{4},m_1+\frac{d-1}{4},m_2+\frac{d-3}{4},\ldots,m_{d-1}-\frac{d-3}{4})
\end{aligned}
$$

The fact that μ is a permutation of λ and the reference above imply (12.81). Now we observe that

$$\frac{\pi^{1/2}}{\Gamma(d/2)}C_\kappa(I_{d-1})=C_{\kappa,0}(I_d) \qquad (12.82)$$

implied by a careful use of formula (12.47). Finally we gather (12.74), (12.81) and (12.82)

$$\int_{\overline{\mathcal{P}}_d}e^{-\operatorname{tr}(st)}r(dt)=(\det s)^{-\frac{d-1}{2}}\sum_{\kappa\in\mathcal{E}_{d-1}}\frac{C_{(\kappa,0)}(s^{-1})}{|\kappa|!}$$

which proves (12.71) and Proposition 13.4 itself. □

XII.4 Example: $m(2,3,3)$

For illustrating Proposition 13.4 we consider the function $f_3(t)$ defined on $\mathcal{P}_3$ by (12.63). More specifically we have

$$f_3(t)=\frac{1}{\det t}\sum_{m_1\geq m_2\geq m_3>0}\frac{C^{(3)}_{m_1,m_2,m_3}(t)}{(m_1+m_2+m_3)!m_1!\Gamma(m_2+\frac{1}{2})(m_3-1)!}.$$

We also consider the measure $m(2,2,2)(dt)$ on $\mathcal{P}_2$ parameterized by $(a,b,c) \mapsto t = \varphi(a,b,c)$ as in (11.21). From formula (12.73) it is

$$\begin{aligned} m(2,2,2)(da,db,dc) &= \frac{1}{(\det t)^{1/2}} \sum_{m_1 \geq m_2 \geq 0} \frac{C^{(2)}_{m_1 m_2}(t)}{(m_1+m_2)! m_1! \Gamma(m_2+\frac{1}{2})} \\ &= \frac{1}{\sqrt{\pi}} \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \frac{(a^2-b^2-c^2)^{n+\frac{k-1}{2}} 2^{2n}}{(\frac{3}{2}+k)_n (n+k)! k! (2n)!} P_k\left(\frac{a}{\sqrt{a^2-b^2-c^2}}\right) \end{aligned}$$

This last formula is obtained by using the calculations done in Section 13.1 for $\Phi^{(2)}_{m_1,m_2}$ in (12.49), for $C^{(2)}_{m_1,m_2}(I_2)$ in (12.53) and an easy manipulation replacing (m_1-m_2, m_1) by (k,n). Finally the measure $r(dt)$ concentrated on the set of matrices of rank 2 in the cone $\overline{\mathcal{P}_3}$ of semi positive definite matrices of order 3 is the image of $m(2,2,2)(da,db,dc)$ times the uniform probability measure du on the orthogonal group $\mathbb{O}(3)$ by the map

$$(a,b,c,u) \mapsto t = u \begin{bmatrix} a+b & c & 0 \\ c & a-b & 0 \\ 0 & 0 & 0 \end{bmatrix} u^*$$

Proposition 13.4 says that the measure $m(2,3,3)(dt)$ on the set $\overline{\mathcal{P}}_3$ of semipositive definite matrices of order 3 defined by the Laplace transform $(\det s)^{-1} \exp \operatorname{trace}\ s^{-1}$ is $r(dt) + f_3(t) 1_{\mathcal{P}_3}(t) dt$.

XIII Convolution lemmas in the cone $\overline{\mathcal{P}_d}$

We give the proof of Lemma 14.1 below although this intuitively obvious fact is certainly a proof somewhere in the literature.

Lemma 14.1: In a Euclidean space E of dimension d let us fix a linear subspace F of dimension n. We choose randomly a linear subspace G of dimension $k \leq d-n$ with the uniform distribution, that is the unique distribution on G such that $G \sim uG$ for all $u \in \mathbb{O}(d)$. Then $\Pr(G \cap F \neq \{0\}) = 0$.

Proof: It is enough to prove the lemma for $E = \mathbb{R}^d$, $F = \{0\} \times \mathbb{R}^{d-n}$ and $k = d-n$. Let $Z_1, \ldots, Z_n$ be i.i.d. random variables in $\mathbb{R}^d$ following the standard Gaussian distribution $N(0, I_d)$. Let G be the random linear subspace of E generated by the vectors $Z_1, \ldots, Z_n$. Since for all $u \in \mathbb{O}(d)$ we have $(uZ_1, \ldots, uZ_n) \sim (Z_1, \ldots, Z_n)$ clearly $G \sim uG$ and G has the uniform distribution. Introduce the matrix

$$M = [Z_1, \ldots, Z_n] = (Z_{ij})_{1 \leq i \leq d,\ 1 \leq j \leq n}$$

whose columns are the vectors $Z_1, \ldots, Z_n$. Then $x_1 Z_1 + \cdots + x_n Z_n = MX$ where $X = (x_1, \ldots, x_n)^*$. Now $G \cap F \neq \{0\}$ implies that there exists a non-zero X such that the n first elements of MX are zero. In other terms, considering the square matrix M_1 of order n defined by $M_1 = (Z_{ij})_{1 \leq i,j \leq n}$, we have that $G \cap F \neq \{0\}$ implies that there exists a non-zero X such that $M_1 X = 0$. This happens if and only if $\det M_1 = 0$. Since the

n^2 entries of the matrix M_1 are independent $N(0,1)$ variables, the event $\det M_1 = 0$ has probability zero and this proves the lemma. $\square$

For the sequel we recall that we denote by S_b the set of $x \in \overline{\mathcal{P}_d}$ with $b = \operatorname{rank} x = 0, \ldots, d$.

Lemma 14.2: Let Y be a random variable in S_b and assume that $uYu^* \sim Y$ for all u in the orthogonal group $\mathbb{O}(d)$. Let $x_0 \in S_a$. Then $x_0 + Y$ is concentrated on S_{a+b} if $a + b \leq d$ or on $S_d = \mathcal{P}_d$ if $a + b \geq d$. Furthermore if $x_0 \in S_c$, and if $x_0 + Y$ is concentrated on S_{a+b} with $a + b < d$ then $c = a$.

Remark: If $a + b = d$ and if $x_0 + Y$ is concentrated on $S_{a+b} = S_d$, x_0 could be on any S_c with $a \leq c \leq k$.

Proof: Apply Lemma 14.1 to $F = x_0\mathbb{R}^d$ and to $G = Y\mathbb{R}^d$. Then almost surely we have $\dim(F + G) = a + b$ if $a + b \leq d$. Furthermore we have always $\operatorname{rank}(x_0 + Y) \leq a + b$. To see that $\operatorname{rank}(x_0 + Y) = a + b$ almost surely, let us suppose that $(x_0 + Y)\mathbb{R}^d \neq E = F + G$. Let x_0' and Y' be the restrictions of the endomorphisms x_0 and Y to the linear space E. Since x_0 and Y are symmetric, this implies that $x_0'E = F$ and $Y'E = G$. Since $(x_0' + Y')E \neq E$ there exists $v \in E \setminus \{0\}$ which is orthogonal to $(x_0' + Y')E$ and this implies that $(x_0' + Y')v = 0$. Since $x_0'v \in F$ and $Y'v \in G$ and since $F \cap G = \{0\}$ this implies that $x_0'v = Y'v = 0$, and v is in $\operatorname{Ker}(x_0') \cap \operatorname{Ker}(Y')$. But since we have almost surely $F \oplus G = E$ (a direct sum, not necessarily an orthogonal one) we have also almost surely $\operatorname{Ker}(x_0') \oplus \operatorname{Ker}(Y') = E$ which implies $\operatorname{Ker}(x_0') \cap \operatorname{Ker}(Y') = \{0\}$. Thus almost surely $v = 0$, a contradiction. Finally $(x_0 + Y)\mathbb{R}^d = E = F + G$ and $\operatorname{rank}(x_0 + Y) = a + b$. If $a + b > d$ then F contains a subspace of dimension $d - b$ and $\dim(F + G) = d$.

Finally, suppose now that $x_0 \in S_c$, and that $x_0 + Y$ is concentrated on S_{a+b} with $a + b < k$. If $c + b < d$ then, by the first part of the lemma, $x_0 + Y$ is concentrated on S_{c+b}. But $S_{a+b} = S_{c+b}$ implies $c = a$. If $c + b \geq k$ then, by the first part of the lemma again, $x_0 + Y$ is concentrated on S_d. This is impossible since $S_{a+b} \neq S_d$. $\square$

Lemma 14.3: Let μ and ν be positive measures on $\overline{\mathcal{P}_d}$ such that ν is concentrated on S_b and ν is invariant by the transformations $x \mapsto uxu^*$, $u \in \mathbb{O}(d)$. Let $a = 0, \ldots, d-2$ such that $a + b < d$. Then $\mu * \nu$ is concentrated on S_{a+b} if and only if μ is concentrated on S_a. Furthermore $\mu * \nu$ is concentrated on $S_d = \mathcal{P}_d$ if μ is concentrated on S_{d-b}.

Proof: $\Rightarrow$ For $y_0 \in S_b$ consider the distribution $K_{y_0}(dy)$ on S_b of the random variable Uy_0U^* where U is uniformly distributed on the orthogonal group $\mathbb{O}(k)$. Let D_b the set of diagonal elements y_0 of S_b of the form $y_0 = \operatorname{diag}(\lambda_1, \ldots, \lambda_b, 0, \ldots, 0)$ such that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_b > 0$. Then there exists a unique positive measure ν_0 on D_b such that the following desintegration holds

$$\nu(dy) = \int_{D_b} \nu_0(dy_0) K_{y_0}(dy).$$

It follows that

$$\begin{aligned}(\mu * \nu)(dx) &= \int_{S_b} \nu(dy)\mu(dx-y) = \int_{S_b} \mu(dx-y)\int_{D_b} \nu_0(dy_0)K_{y_0}(dy) \\ &= \int_{D_b} \nu_0(dy_0)\int_{S_b} \mu(dx-y)K_{y_0}(dy)\end{aligned}$$

Therefore the measure $\mu * K_{y_0}$ is concentrated on S_{a+b} for ν_0 almost all $y_0 \in D_b$. From Lemma 14.2 this implies that μ is concentrated on S_a.

$\Leftarrow$ If μ is concentrated on S_a with $a+b \leq k$ it is an easy consequence of Lemma 14.2 that $\mu * \nu$ is concentrated on S_{a+b}. $\square$

Lemma 14.4: Let $a, b \in \{1, \ldots, d-1\}$ such that $a+b < d$. If $m(a, a+b, d)$ exists, it is concentrated on S_a.

Proof: From the Laplace transforms we know that

$$m(a, a+b, d) * m(b, 0, d) = m(a+b, a+b, d).$$

From Proposition 10.2 we know that $m(a+b, a+b, d)$ is concentrated on S_{a+b}. Since the Laplace transform of $m(b, 0, d)$ is $(\det s)^{-b/2}$ we know that $m(b, 0, d)$ is invariant by the transformations $x \mapsto uxu^*$ for any $u \in \mathbb{O}(d)$. By Lemma 14.3 we deduce that $m(a, a+b, d)$ is concentrated on S_a if it exists. $\square$

XIV $m(d-2, d-1, d)$ and $m(d-2, d, d)$ do not exist for $d \geq 3$

In this section we prove Propositions 10.4 and 10.5.

Proof of Proposition 10.5. Suppose that $m(d-2, d, d)$ exists. By Lemma 14.4 the measure $m(d-2, d, d)$ is concentrated on S_{d-2}. By Lemma 14.3 the convolution

$$m(d-2, d, d) * m(1, 0, d) = m(d-1, d, d)$$

is concentrated on S_{d-1}. This contradicts Proposition 13.4, where it has been shown that $m(d-1, d, d)$ has an absolutely continuous part. This shows in a different way the main result of Mayerhofer. $\square$

Proof of Proposition 10.5. Suppose that $m(d-2, d-1, d)$ exists. By Lemma 14.4 the measure $m(d-2, d-1, d)$ is concentrated on S_{d-2}. Therefore there exists a positive measure $m(dy) = m(dy_1, \ldots, dy_{d-2})$ on $\mathbb{R}^{d(d-2)} = \mathbb{R}^d \times \ldots \times \mathbb{R}^d$ such that for all $s \in \mathcal{P}_d$ we have

$$\int_{\mathbb{R}^{d(d-2)}} e^{-(y_1^* s y_1 + \cdots + y_{d-2}^* s y_{d-2})} m(dy) = \frac{1}{\det s^{(d-2)/2}} e^{\operatorname{tr}(s^{-1} I(d-1,d))}. \tag{14.83}$$

We write more conveniently the elements $y = (y_1, \ldots, y_{d-2})$ with the help of the transposed matrix $y^* = (y_{i,j})$ with $d-2$ rows $y_1^*, \ldots, y_{d-2}^*$ and d columns $c_1, \ldots, c_d$

$$y^* = \begin{bmatrix} y_1^* \\ \ldots \\ y_{d-2}^* \end{bmatrix} = \begin{bmatrix} y_{1,1} & \cdots & y_{1,d} \\ \ldots & \ldots & \ldots \\ y_{d-2,1} & \cdots & y_{d-2,d} \end{bmatrix} = [c_1, \ldots, c_d].$$

With this notation introduce the Gram matrix

$$G(c) = G(c_1, \dots, c_d) = (\langle c_j, c_k \rangle)_{1\leq j,k\leq d}$$

and denote by $m(dc)$ what we denoted by $m(dy)$ before. We get

$$\int_{\mathbb{R}^{d(d-2)}} e^{-\operatorname{tr}(sG(c))} m(dc) = \frac{1}{\det s^{(d-2)/2}} e^{\operatorname{tr}(s^{-1}I(d-1,d))}. \tag{14.84}$$

Equality (14.84) means that $m(d-2, d-1, d)(dx)$ is the image of $m(dc)$ by $c \mapsto x = G(c)$.

Now in (14.84) we choose $s = \operatorname{diag}(1, s_1)$ where s_1 is a symmetric positive definite matrix of order $d-1$. We also desintegrate $m(dc)$ by introducing a probability kernel $K(c_2, \dots, c_d; dc_1)$ and a positive measure $m_1(dc_2, \dots, dc_d)$ such that

$$e^{-\|c_1\|^2} m(dc_1, dc_2, \dots, dc_d) = m_1(dc_2, \dots, dc_d) K(c_2, \dots, c_d; dc_1)$$

With these notations we can write

$$\begin{aligned}
\frac{1}{\det s_1^{(d-2)/2}} e^{\operatorname{tr}(s_1^{-1})} &= \int_{\mathbb{R}^{d(d-2)}} e^{-\operatorname{tr}(sG(c))} m(dc) \\
&= \int_{\mathbb{R}^{d(d-2)}} e^{-\|c_1\|^2} e^{-\operatorname{tr}(s_1 G(c_2,\dots,c_d))} m(dc_1, dc_2, \dots, dc_d) \\
&= \int_{\mathbb{R}^{(d-1)(d-2)}} e^{-\operatorname{tr}(s_1 G(c_2,\dots,c_d))} \left(\int_{\mathbb{R}^{d-2}} K(c_2, \dots, c_d; dc_1) \right) m_1(dc_2, \dots, dc_d) \\
&= \int_{\mathbb{R}^{(d-1)(d-2)}} e^{-\operatorname{tr}(s_1 G(c_2,\dots,c_d))} m_1(dc_2, \dots, dc_d)
\end{aligned}$$

since K is a probability kernel. The last equality says that the image of $m_1(dc_2, \dots, dc_d)$ by the map $(c_2, \dots, c_d) \mapsto x = G(c_2, \dots, c_d)$ is nothing but $m(d-2, d-1, d-1)(dx)$. Denote $G_2 = G(c_2, \dots, c_d)$ for simplicity. Since $c_2, \dots, c_d$ are vectors of a Euclidean space of dimension $d-2$ the rank of G_2 is less than or equal to $d-2$. To prove this elementary fact of linear algebra we use $G_2 \in \overline{\mathcal{P}_{d-1}}$. This implies that if $x = (x_2, \dots, x_d)^*$ then $G_2 x = 0$ if and only if $x^* G_2 x = 0$. Since $x^* G_2 x = \| \sum_{i=2}^d x_i c_i \|^2$ the linear space of $x \in \mathbb{R}^{d-1}$ such that $\sum_{i=2}^d x_i c_i = 0$ has at least dimension 1, the kernel of the endomorphism of $\mathbb{R}^{d-1}$ with matrix G_2 has at least dimension 1 and its image has at most dimension $d-3$. This contradicts Proposition 14.4 which says that $m(d-2, d-1, d-1)$ has an absolutely continuous part and therefore charges matrices with rank $d-1$. This proves the conjecture of Mayerhofer. $\square$

XV References

S.A. Andersson (1975) Invariant normal models, *Ann. Statist.* **3,** 132-155.

M. Casalis (1990). Familles exponentielles naturelles invariantes par un groupe, Thèse, Université Paul Sabatier, Toulouse.

M. Casalis (1991). Les familles exponentielles à variance quadratique sont des lois de Wishart sur un cône symétrique, *C. R. Acad. Sci. Paris,* **312,** Série I, 537-540.

M. Casalis and G. Letac (1994). Characterization of the Jorgensen set in the generalized linear model, *Test,* **3** 145-161.

M. Casalis and G. Letac (1996). The Lukacs-Olkin-Rubin characterization of Wishart distributions on symmetric cones, *Ann. Statist.* **34,** 763-786.

A.G. Constantine 'Some noncentral distribution problems in multivariate analysis.' *Ann. Math. Statist.* **34**, 1270-1285.

J. Faraut and A. Korányi (1994) *Analysis on Symmetric Cones.* Oxford University Press.

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

P. Graczyk, G. Letac, G. and H. Massam (2003) 'The complex Wishart distribution and the symmetric group.' *Ann. Statist.* , **31** no 1, 287-309.

S. Gindikin (1975) 'Invariant generalized functions in homogeneous spaces', *J. Funct. Anal. Appl.,* **9**, 50-53.

P. Graczyk, G. Letac, G. and H. Massam (2005) 'The hyperoctahedral group, symmetric group representations and the moments of the real Wishart distribution.' *J. Theoret. Probab.*, **18**, 1-43.

S.T. Jensen (1988). Covariance hypotheses which are linear in both the covariance and the inverse covariance, *Ann.Statist.* **16,** 302-323.

G. Letac (1989). A characterization of the Wishart exponential families by an invariance property, textitJ.Theoret.Prob. **2,** 71-86.

G. Letac (1994). Les familles exponentielles statistiques invariantes par les groupes du cône et du paraboloide de révolution, *J.Appl.Prob,* **31A,** 71-95. Volume in honor of Lajos Takács, J.Galambos and J.Gani editors.

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

G. Letac and H. Massam (1998). Quadratic and inverse regression of Wishart distributions, *Ann. Statist.* **36,** 763-786.

G. Letac and H. Massam (2008) 'The noncentral Wishart as an exponential family and its moments.' *J. Multiv. Anal.*, **99**, 1393-1417.

G. Letac and J.Wesołowski (2008) 'Laplace transforms which are negative powers of quadratic polynomials' *Trans. Amer. Math. Soc.*, **360**, 6475-6496.

H. Massam and E. Neher (1997). On transformations and determinants of Wishart variables in symmetric cones, *J. Theoret. Prob.***10,** 867-903.

E. Mayerhofer (2013) 'On the existence of non-central Wishart distributions.' *J.Multivariate Anal.*, **114**, 448-456. *ArXiv* 1009-0473

R.J. Muirhead (1982) *Aspects of Multivariate Analysis.* Wiley, New York.
vspace4mmI. Olkin and H. Rubin (1962). A characterization of the Wishart distribution, *Ann. Math. Stat.*,**33,** 1272-1280.

S.D. Peddada and D. St. P. Richards (1991) Proof of a conjecture of M.L. Eaton on the characteristic function of the Wishart distribution', *Ann. Probab.* **19**, 869-875. 272-280.

S. Roman (1980) 'The formula of Faà di Bruno.' *Amer. Math. Monthly*, **87**, 806-809.

D.N. Shanbhag (1988) 'The Davidson Kendall problem and related results on the structure of the Wishart distribution.' *Austr. J. Statist.*, **30A** 272-280.

A. Takemura. (1984) *Zonal Polynomials.* Lecture notes-Monograph series, Institute of Mathematical Statistics, Hayward, California.